

Pre-Stimulus hemodynamic activity in dorsal attention network is negatively associated with decision confidence in visual perception

Dobromir Asenov Rahnev, Linda Bahdo, Floris Pieter De Lange and Hakwan Lau

J Neurophysiol published online 20 June 2012

You might find this additional info useful...

Updated information and services including high resolution figures, can be found at:
<http://jn.physiology.org/content/early/2012/06/13/jn.00184.2012.full>

Additional material and information about *Journal of Neurophysiology* can be found at:
<http://www.the-aps.org/publications/jn>

This information is current as of July 29, 2012.

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32

Pre-Stimulus hemodynamic activity in dorsal attention network is negatively associated with decision confidence in visual perception

Dobromir A. Rahnev^{1,2}, Linda Bahdo¹, Floris P. de Lange², Hakwan Lau^{1,2}

¹ Columbia University, Department of Psychology, New York, USA

² Radboud University Nijmegen, Donders Institute for Brain, Cognition and Behaviour, Netherlands

Abbreviated title: Pre-stimulus BOLD activity and decision confidence

Manuscript information: 198 words (Abstract), 419 words (Introduction), 1264 words (Discussion), 24 pages, 3 Figures, 1 Table.

Keywords: Attention, confidence, fMRI, spontaneous fluctuation, perceptual decision making

Acknowledgements: HL and FdL together received funding from the Human Frontiers Science Project (HFSP Short-Term Fellowship), the Templeton Foundation (grant number 21569) and the Netherlands Organization for Scientific Research (NWO) (grant number 2300164119).

Correspondence:

Dobromir Rahnev
Columbia University
Department of Psychology
406 Schermerhorn Hall
1190 Amsterdam Ave, MC 5501
New York, NY 10027
Tel: 1 212 854 7169
E-mail: drahnev@gmail.com

33 **Abstract**

34 Attention is thought to improve most aspects of perception. However, we recently showed that,
35 somewhat surprisingly, endogenous attention can also lead to low subjective perceptual ratings
36 (Rahnev et al., 2011). Here we investigated the neural basis of this effect and tested whether
37 spontaneous fluctuations of the attentional state can lead to low confidence in one's perceptual
38 decision. We measured pre-stimulus functional magnetic resonance imaging activity in the dorsal
39 attention network and used that activity as an index of the level of attention involved in a motion
40 direction discrimination task. Extending our previous findings, we showed that low pre-stimulus
41 activity in the dorsal attention network, which presumably reflected low level of attention, was
42 associated with higher confidence ratings. These results were explained by a signal detection
43 theoretic model in which lack of attention increases the trial-by-trial variability of the internal
44 perceptual response. In line with the model, we also found that low pre-stimulus activity in the
45 dorsal attention network was associated with higher trial-by-trial variability of post-stimulus
46 peak activity in motion-sensitive region MT+. These findings support the notion that lack of
47 attention may lead to liberal subjective perceptual biases, a phenomenon we call "inattentional
48 inflation of subjective perception" (IISP).

49

50 **Introduction**

51 A large number of studies have clearly demonstrated that endogenous attention improves
52 perception (for review see Carrasco, 2011). However, the exact relationship between the two can
53 be fairly complex (Reynolds and Heeger, 2009). We recently showed that attention can lead to
54 low subjective visibility ratings (Rahnev et al., 2011). This is in line with another study that
55 demonstrated that attention can improve accuracy without influencing confidence (Wilimzig et
56 al., 2008). Together, these studies suggest that accuracy can dissociate from confidence under
57 attention. In Rahnev et al. (2011) we proposed a formal model that accounts for this effect.

58 According to the model, lack of attention increases the variability of the internal perceptual
59 response and confidence ratings are given using the same unified criteria for both attended and
60 unattended stimuli (Figure 1). Thus, the “low attention” distributions are marked by high
61 variability. This makes these distributions, as compared to the “high attention” distributions,
62 extend further into the high confidence regions and thus producing more high confidence trials.

63
64 Here we used fMRI to investigate the neural bases of the above effects (Rahnev et al., 2011). In
65 particular, we tested whether spontaneous fluctuations of attention would lead to differential
66 effects on accuracy and confidence ratings, as well as whether such effects would be
67 accompanied by a decrease in variability of the perceptual signal (as predicted by our model).

68 We measured the pre-stimulus BOLD activity in the dorsal attention network (Corbetta and
69 Shulman, 2002) which is implicated in directed attention and working memory and is part of the
70 task-positive system in the brain. The name “dorsal attention network” distinguishes it from the
71 “ventral attention network” which is specialized for the detection of behaviorally relevant
72 stimuli, particularly when they are salient or unexpected (Corbetta and Shulman, 2002). The
73 fluctuations of BOLD activity in dorsal attention network have been investigated in a number of

74 previous resting-state studies (Fox et al., 2005, 2006, 2007; Sadaghiani et al., 2009, 2010) and
75 have been theorized to reflect the attentional level of the subject (Eichele et al., 2008; Sapir et al.,
76 2005). As in previous studies (Hesselmann et al., 2008a,b; Sadaghiani et al., 2009), we focused
77 on pre-stimulus activity in the dorsal attention network because it is not contaminated by
78 stimulus related activity and reflects ongoing neural fluctuations.

79

80 Based on our previous findings, we hypothesized that high pre-stimulus BOLD activity in the
81 dorsal attention network would be associated with low confidence ratings and higher accuracy in
82 the motion discrimination task. Further, we predicted that these effects would be accompanied by
83 an increase in the variability of the perceptual signal as measured by the variance in the evoked
84 responses in the motion sensitive region MT+.

85

86

87 **Methods**

88 *Participants*

89 Fifteen volunteers (9 women; mean age = 22; range = 19-26 years) participated in the
90 experiment. All subjects were naive regarding the purposes of the experiments, had normal or
91 corrected-to-normal vision, and signed an informed-consent statement approved by the local
92 ethics committee (CMO region Arnhem-Nijmegen, The Netherlands).

93

94 *Stimulus and Procedure*

95 In each trial (Figure 2), subjects were required to indicate the overall direction of motion of
96 white dots (density = 2.4 dots/degree²; speed = 6 degrees/sec) presented inside a black annulus
97 (outer circle radius = 10°; inner circle radius = 1°). The motion direction was either contracting
98 or expanding. Incoherent dots moved randomly with the same speed as coherent dots. All dots
99 had infinite lifetime: they were never removed from the screen during the motion presentation. A
100 small fixation square was presented for the duration of the trial and subjects were required to
101 maintain fixation on it. The stimuli were presented on gray background, and were generated
102 using the Psychophysics Toolbox (Brainard, 1997) in MATLAB.

103

104 We employed a sparse event-related design (Hesselmann et al., 2008a,b; Sadaghiani et al., 2009)
105 where each trial began with 20-40 seconds of no motion stimuli. This period allowed for the
106 blood-oxygenated level dependent (BOLD) signal to return to baseline after each trial. We then
107 presented 100 ms of coherent motion and followed it with a “mask” of 400 ms of random
108 motion. The masking was employed to increase task difficulty and to avoid retrospection based
109 on iconic memory (Sperling, 1960). We elected to use motion stimuli because of ease of control
110 of the stimulus properties and localization of the relevant brain regions. These have also been

111 used in previous studies of spontaneous fluctuations of BOLD activity (Hesselmann et al.,
112 2008a; Sapir et al., 2005). Subjects were required to indicate the direction of motion (expanding /
113 contracting), and give a confidence rating with separate button presses made with their right
114 hand. Confidence was judged on a 4-point scale, where 1 represented low and 4 represented high
115 confidence. Subjects were instructed to use the scale as evenly as possible. Nevertheless, since
116 some subjects had a bias toward using one side of the confidence scale, and in order to maximize
117 power, we defined what constituted high and low confidence for each subject individually using
118 a median split. Overall low confidence was defined as a rating of 1 or 2 for 9 subjects and as a
119 rating of 1 for the other 6 subjects. If both button presses were not completed within 6 seconds,
120 the trial was marked as unanswered and excluded from further analyses. Subjects completed 104
121 trials separated in two scans of 52 trials. Each scan took about 26 minutes. After these two scans,
122 we acquired a 307-volume “resting state” scan (Fox and Raichle, 2007) that took about 10
123 minutes.

124
125 Prior to the fMRI experiment, each subject took part in a one-hour training session on a separate
126 day, in which the subject practiced 520 trials that were identical to the trials in the scanner but
127 for a shorter ITI of 1-3 seconds. During the first half of the training session subjects were given
128 trial-by-trial feedback. The feedback was discontinued during the second part of the training in
129 order to prepare subjects for the experiment in the scanner environment. Based on the data from
130 the training session, motion coherence levels were chosen for each subject to produce ~65%
131 correct responses (mean motion coherence = 30.6%, SD = 27.5%). Just before the fMRI
132 experiment, subjects practiced for additional 5 minutes (68 trials) to remind them of the task.

133

134

135 *Behavioral analyses*

136 We analyzed our data to check for the existence of trial-to-trial contingencies. For each subject
137 we investigated whether a correct or high confidence response on the current trial predicted
138 correct or high confidence response on the subsequent trials. To do that, we estimated the
139 proportion of correct and high confidence trials following correct, error, high confidence, and
140 low confidence trials. We then compared these proportions for correct and error trials, as well as
141 for high and low confidence trials. The comparison was done between subjects using a paired-
142 sample t-test.

143

144 *fMRI acquisition*

145 Images were acquired on a 3 Tesla Trio MRI system (Siemens, Erlangen, Germany). Functional
146 images were acquired using a 32-channel coil, with a single shot gradient echo-planar imaging
147 (EPI) sequence (repetition time: 1950 ms; echo time: 30 ms; 31 ascending slices; voxel size: 3 x
148 3 x 3 mm; flip angle = 80 degrees; field of view = 192 mm). A high resolution anatomical image
149 was acquired using a T1-weighted MP-RAGE sequence (repetition time: 2300 ms; echo time:
150 3.03 ms; voxel size: 1 x 1 x 1 mm).

151

152 *fMRI preprocessing*

153 Analysis was performed using SPM5 (Wellcome Department of Imaging Neuroscience, London,
154 UK). The first 6 volumes of each scan were discarded to allow for scanner equilibration.
155 Preprocessing consisted of realignment through rigid-body registration to correct for head
156 motion, slice timing correction to the onset of the first slice, co-registration of the functional and
157 anatomical images, segmentation of the anatomical image, normalization to Montreal
158 Neurological Institute (MNI) space using the grey matter image obtained from the segmentation,

159 interpolation of functional images to 2 x 2 x 2 mm, and smoothing with a Gaussian kernel with a
160 full-width at half-maximum of 8 mm.

161

162 *Definition of brain networks*

163 We defined the attention, the default, and the alertness networks using a separate 10-minute
164 resting state scan and employing seed-based analyses as in previous studies (Fox et al., 2005,
165 2006; Sadaghiani et al., 2009, 2010). Briefly, using the MarsBaR toolbox for MATLAB, we
166 extracted the time courses for the resting state scan for spheres with 10 mm radius. We used
167 seeds in the right intraparietal sulcus (IPS; 27, -58, 49) for the attention network (Fox et al.,
168 2006; Sadaghiani et al., 2009), the posterior cingulate cortex (PCC; -5, -49, 40) for the default
169 network (Fox et al., 2005; Sadaghiani et al., 2009), and dorsal ACC (dACC; 0, 15, 40) for the
170 alertness network (Sadaghiani et al., 2010). Each time course was high-pass filtered (1/128 Hz),
171 and the time courses for all grey matter voxels, all white matter voxels, and all cerebral spinal
172 fluid (CSF) voxels were regressed out in a multiple regression (Sadaghiani et al., 2009). This
173 step was similar to the “global signal regression” that is typically performed in such experiments
174 (Fox et al., 2005) whereby the average signal from the whole brain is regressed out. This pre-
175 processing step is useful in that it removes BOLD activity from physiological (i.e., non-neuronal)
176 origin (Fox et al., 2009) but it has been criticized for inducing artificial anti-correlations between
177 networks (Anderson et al., 2011; Murphy et al., 2009). Nevertheless, the primary interest in this
178 study was investigating the influence on behavior of the pre-stimulus activity in a specific
179 network (the dorsal attention network) and thus we wanted to isolate its influence independent of
180 physiological and other global influences. On the other hand, we were not concerned with
181 potential anticorrelations that could result from this processing step.

182

183 The resulting time courses were used as regressors in separate general linear models (GLM) for
184 each subject; the three networks were defined as the set of voxels correlated at $p < 0.001$ with the
185 respective regressor. Each of the three networks was then combined at the second level. Each
186 subject-specific network was a combination of that subject's network masked with the
187 corresponding group network.

188

189 To test for the robustness of the findings, the dorsal attention network was alternatively
190 identified using independent components analysis (ICA). We employed the GIFT toolbox
191 (Calhoun et al., 2001) for MATLAB to extract the networks consistent across all subjects
192 (Damoiseaux et al., 2006). We used the Infomax algorithm to find 16 components. The attention
193 network was identified among these 16 components and was defined using a z-score threshold of
194 1. The region MT+ was removed from the attention networks defined either using either seed-
195 based methods or ICA.

196

197 *Pre-stimulus activity analyses*

198 After the networks were defined for each subject, we analyzed the data from the first two runs in
199 which subjects engaged in the motion discrimination task. We extracted time courses for each
200 network using the MarsBaR toolbox (<http://marsbar.sourceforge.net/>) and regressed out nuisance
201 variables such as head motion and global signal fluctuations as above. Pre-stimulus activity was
202 defined as the average of the two volumes preceding the onset of the motion stimulus. We used
203 paired sample t-tests and repeated measures ANOVAs to test whether pre-stimulus activity
204 differed significantly between high and low confidence trials, as well as between correct and
205 incorrect trials. Even though we tested for the effects in three different networks (see above), we

206 did not correct for multiple comparisons because our main hypothesis was about the effects in
207 the dorsal attention network, and the other networks were just included for completeness.
208
209 We also investigated whether pre-stimulus activity in the dorsal attention system influenced the
210 variability in MT+. For each subject, we categorized trials as either having low or high pre-stim
211 dorsal attention network activity if the pre-stimulus activity in the network was lower or higher
212 than the average activity for that subject. We then computed the Fano factor of the MT+ evoked
213 activity for these two types of trials. The evoked activity was defined as the maximum BOLD
214 activity of the 3 scans that had highest activity on average (across all subjects and all trials). The
215 Fano factor was defined as the variance divided by the mean for the evoked activity for each of
216 the two types of trials for each subject (Carandini, 2004; Churchland et al., 2010; Gur et al.,
217 1997).

218

219 *MT+ localizer*

220 We identified MT+ using a separate localizer scan after the main experiment and the resting state
221 scan. Sixty blocks of moving dots (block duration of 16 s) were alternated with 10 blocks of
222 stationary dots (block duration of 16 s), resulting in ~19 min scan duration. Motion-sensitive
223 areas were obtained using the contrast moving dots > stationary dots. Local maxima near the
224 ascending limb of the inferior temporal sulcus were defined bilaterally and combined to form
225 area MT+ separately for each subject (Hesselmann et al., 2008a). MT+ defined in this was used
226 as a region of interest (ROI) in a separate GLM analysis (see below).

227

228 *Standard GLM analysis*

229 Regressors for the first-level analysis of evoked activity were obtained by convolving the unit
230 impulse time series for each condition with the canonical hemodynamic response function. The
231 motion stimulus was modeled with duration of 500 ms. Our model included 4 regressors,
232 reflecting the combination of 2 levels of accuracy (correct/error), and 2 levels of confidence
233 (high/low). We included 12 nuisance regressors related to head motion: three regressors related
234 to translation and three regressors related to rotation of the head, as well as their derivatives
235 (Lund et al., 2005).

236

237 *Simulations*

238 We performed simple computer simulations in order to investigate whether our model (Figure 1)
239 can provide an adequate explanation of the results, especially given that we did not observe a
240 statistically significant effect of pre-stimulus BOLD signal in the dorsal attention network on
241 motion discrimination accuracy (see Results). The main idea of the model is that attention
242 decreases the variability of the perceptual signal. We implemented this intuition in the equations
243 below. We formalized that:

244

$$245 \quad S(t) = E(t) + N_p(t, A(t))$$

246

247 where S is the evidence available to the perceptual system, E is the evidence present in the
248 stimulus, N_p is the amount of physiological noise inherent in the system, A is the level of
249 attention in the current trial, and t indicates the trial. The pre-stimulus BOLD signal in the dorsal
250 attention network was modeled as:

251

$$252 \quad B(t) = e_A * A(t) + N_B(t)$$

253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276

where B is the BOLD signal, N_B is the noise present in the BOLD signal, and e_A controls the effect of attention on the BOLD signal.

We modeled the level of attention present in each trial (A) using a uniform distribution in the interval $(0, 1)$. The evidence present on each trial (E) was sampled from a normal distribution with a mean μ_E and standard deviation σ_E in order to reflect the fluctuations in the presentation of the random dot motion. N_P and N_B were modeled as Gaussian distributions with means μ_P and μ_B , and standard deviations σ_P and σ_B , respectively. Since the absolute amount of evidence was not of interest, μ_E was set to 1 and μ_P was set to 0. For simplicity we assumed a linear influence of attention on the decrease of variability in the perceptual signal and therefore modeled σ_P as $\sigma_P = m - d \cdot A$, where, m is the mean value of the noise in N_P while d controls the degree of attentional decrease of the variability in N_P . Finally, since the random variables E and N_P are both normally distributed, their sum will also be normally distributed and have a standard deviation that is the sum of the standard deviations of E and N_P . Therefore, rather than fitting σ_P and σ_E separately, we fit only their sum.

The perceptual decision was made by comparing $S(t)$ to 0: positive values indicated correct responses, while non-positive values were coded as incorrect response. On the other hand, confidence was determined by comparing $S(t)$ to a set of criteria (c_1, c_2). High confidence ratings were given if $S(t) < c_1$ or if $S(t) > c_2$. Because of lack of significant bias for expanding or contracting motion in our experiment, for simplicity c_1 was set to $-c_2$. Thus, the set of confidence criteria we used was $(-c, c)$.

277 We fitted the above equations with the average data from all 15 participants using a simulated
278 annealing procedure (Kirkpatrick et al., 1983). Each iteration of the fitting procedure generated
279 1,500,000 trials (roughly corresponding to 1000 times the data from our experiment) based on
280 the current values of the parameters. The fitting function attempted to minimize the error in
281 overall percent of correct answers, overall confidence, the mean pre-stimulus BOLD activity for
282 all high confidence, low confidence, correct, and error trials. To test for the robustness of our
283 findings, the fitting was done four different times using different starting values for the
284 parameters. We varied the starting values of the parameters in order to avoid getting “stuck” in
285 local minima. Each of the four fits provided a good fit to the data. Here we report the results
286 from the first fit which produced the best match to the observed values; the results of the other
287 fits were consistent with the best one. Using the parameters from the best fit, we generated
288 10,000 experiments that consisted of 15 subjects each completing 100 trials (15,000,000 trials
289 total) and checked the power to detect significant effects on confidence and accuracy.

290 **Results**

291 We first checked for the existence of trial-to-trial contingencies in accuracy or confidence
292 ratings. Paired-sample t-tests demonstrated that accuracy on the current trial predicted neither
293 accuracy ($t(14) = -0.83, p = 0.42$) nor confidence ($t(14) = -1.3, p = 0.21$) on the subsequent trial.
294 Similarly, confidence on the current trial predicted neither accuracy ($t(14) = 0.2, p = 0.85$) nor
295 confidence ($t(14) = -1.04, p = 0.32$) on the subsequent trial. Thus, it appeared that there were no
296 reliable trial-to-trial contingencies indicating that our subjects were generating stochastic
297 behavioral responses.

298
299 We identified the dorsal attention network using standard seed-based methods (Fig. 3a, see
300 Methods; other methods of identifying the network, such as independent component analysis,
301 ICA, gave similar results, see Table 1). Confirming our prediction, we found that high pre-
302 stimulus BOLD activity in the dorsal attention network was associated with low confidence
303 ratings ($F(1,13) = 6.68, p = 0.02$; Fig. 3b,c). On the other hand, pre-stimulus activity in this
304 network was not reliably associated with accuracy ($F(1,13) = 0.45, p = 0.52$). There was also no
305 interaction between confidence and accuracy ($F(1,13) = 0.02, p = 0.89$). The above analyses
306 were performed using repeated measures ANOVA. However, one subject needed to be excluded
307 from that analysis because he did not have any high confidence error trials (i.e., one of the cells
308 in the 2 x 2 design was empty). To include that subject one could perform a simple t-test on the
309 main effect of confidence and accuracy. This led to the same pattern of results for both
310 confidence ($t(14) = 2.69, p = 0.02$) and accuracy ($t(14) = 0.55, p = 0.59$).

311
312 Since the present experiment is based on detection of coherent motion, we also examined the
313 BOLD signal in motion sensitive area MT+. Although this area is sometimes considered to be

314 part of the dorsal attention network, in the above analysis we delineated it as a distinct region in
315 order to look at its effects separately. Unlike the effect in the dorsal attention network, pre-
316 stimulus activity in MT+ was not associated with subjects' confidence ratings ($t(14) = 1.31, p =$
317 0.21). This lack of significant effect on confidence suggests the difference between high and low
318 confidence trials in dorsal attention network was not driven by fluctuations of activity in MT+.
319 Similarly, pre-stimulus activity in MT+ for error trials (mean activity = -0.03) and correct trials
320 (mean activity = -0.01) was not significantly different ($t(14) = 0.39, p = 0.7$). Further, standard
321 GLM analyses showed that for post-stimulus evoked activity, high activity in MT+ was
322 associated with high level of confidence ratings ($t(14) = 2.98, p = 0.01$). Thus, in this study, the
323 activity in MT+ played a bigger role in reflecting stimulus processing rather than pre-stimulus
324 fluctuation of attentional states.

325
326 We further tested the relationship between confidence and pre-stimulus activity in two other
327 networks. We found that pre-stimulus activity in neither the default mode network (Damoiseaux
328 et al., 2006; Eichele et al., 2008; Fox et al., 2005; Sadaghiani et al., 2009) nor the alertness
329 network (Sadaghiani et al., 2009, 2010) predicted confidence or accuracy on the task (all p 's >
330 0.1; Table 1).

331
332 The results on the dorsal attention network (DAN) corroborate our computational model (Figure
333 1). One of the crucial assumptions of the model is that lack of attention increases the trial-by-trial
334 variability of the perceptual signal. We provided an indirect test of this assumption by
335 investigating whether pre-stimulus activity in DAN affected the variability in the evoked activity
336 in MT+, a region that likely codes the perceptual signal. We computed the Fano factor of the
337 distributions of evoked MT+ activity, a measure that has been used extensively to characterize

338 neural variability (for example, see Carandini, 2004; Churchland et al., 2010; Gur et al., 1997).
339 Confirming our model's prediction, high pre-stimulus activity in DAN led to a lower Fano factor
340 in MT+ ($t(14) = 2.56, p = 0.02$; Figure 3d). No such relationship was found for the default
341 network ($t(14) = -0.71, p = 0.49$) or the alertness network ($t(14) = 0.4, p = 0.7$). We considered
342 the alternative interpretation that the result for DAN was due to a ceiling effect: if the pre-
343 stimulus BOLD activity is already high in DAN then pre-stimulus activity may be relatively high
344 in MT+, and thus evoked activity in MT+ could potentially show ceiling effects. Thus we
345 computed the skewness of the distributions of evoked activity in MT+. A ceiling effect would
346 manifest itself as negative skewness. Nevertheless, we found that skewness was positive for
347 trials with either low (skewness = 0.30) or high (skewness = 0.28) pre-stimulus DAN activity
348 and there was no significant difference in skewness between the two distributions ($p = .92$).
349
350 Finally, in order to test whether our model (Figure 1) in which attention decreases the variability
351 of the perceptual signal can explain the observed data, we carried out simple computational
352 simulations in which attention modulated linearly the standard deviation of the signal detection
353 distributions (see Methods). These simulations were largely motivated by the fact that we did not
354 observe a statistically significant effect of pre-stimulus BOLD signal in the dorsal attention
355 network on motion discrimination accuracy, yet this was one of the predictions of our detection
356 theoretic model (Figure 1). We performed the computational simulations four different times and
357 obtained good fits of the data all four times (Table 2 reports the fit of the first simulation that
358 provided the best fit). Further, the parameter that controlled the attentional decrease of the
359 variability of the perceptual signal was consistently positive across all four fits (see Methods)
360 suggesting that attention indeed decreased perceptual noise (note that the parameter was not
361 constrained and that it could have taken negative values). For the fitted values of the parameters

362 we generated 10,000 replications of our experiment (by generating 15 sets of 100 trials for each
363 replication). We compared the simulated pre-stimulus BOLD responses and found that simulated
364 BOLD activity was significantly higher for low (compared to high) confidence trials on 7,188
365 trials (i.e., 72% of all simulations), while it was significantly higher for correct (compared to
366 error) trials on 1,306 trials (i.e., 13% of all simulations). Thus our simulations suggest that there
367 is a large difference in our statistical power to detect effects on confidence vs. effects on
368 accuracy, which may explain why in our dataset only the effect on confidence was significant.
369 Therefore, the lack of a significant effect of pre-stimulus BOLD signal in the dorsal attentional
370 network on accuracy may well be due to the lack of statistical power (as estimated from these
371 simulations), and thus does not necessarily contradict our detection theoretic model (Figure 1).

372

373

374 **Discussion**

375 We found that low pre-stimulus dorsal attention network (DAN) activity was associated with
376 high confidence in a motion discrimination task, as well as higher trial-by-trial variability in
377 post-stimulus peak activity in MT+. Our signal detection theoretic model (Figure 1) predicted
378 this pattern of results though its last prediction – that low pre-stimulus DAN activity should lead
379 to lower accuracy – was not confirmed. Overall, our data provide preliminary support of our
380 noise-based account of the phenomenon of “inattentive inflation of subjective perception” (Lau
381 and Rosenthal, 2011; Rahnev et al., 2011).

382

383 Our results are novel in a number of ways. First, they demonstrate that attention may lead to
384 dissociations between accuracy and confidence ratings. This is especially surprising since many
385 previous studies have reported that accuracy and confidence typically increase or decrease
386 together (e.g., Busey et al., 2000). Second, our analysis on the influence of pre-stimulus activity
387 in the dorsal attention network (DAN) on the evoked activity in MT+ suggests that attention
388 decreases the trial-by-trial variability of the perceptual signal. Our model implies that this noise
389 reduction is a critical feature that explains the observed dissociation between accuracy and
390 confidence. On the other hand, most previous models of attention focus primarily (or
391 exclusively) on the attentional increase on gain and have placed less focus on attention’s
392 influence on the trial-by-trial variability of the internal perceptual response (e.g., Desimone and
393 Duncan, 1995).

394

395 One may worry that our finding of pre-stimulus BOLD in the dorsal attention network being
396 negatively correlated with confidence could seem contradictory to previous single-neuron
397 recording research. For example, Kiani and Shadlen (2009) investigated the response of LIP

398 neurons when monkeys were given the opportunity of choosing a safe option (thus indicating
399 low confidence in their decision). The researchers found that the monkeys chose the safe option
400 when the activity in LIP neurons was at an intermediate level and therefore the activity in the
401 recorded neurons did not provide strong evidence for either decision option. Thus Kiani and
402 Shadlen's results could be interpreted as predicting that the population response of LIP neurons
403 would not distinguish between high and low confidence responses. Nevertheless, though Kiani
404 and Shalden found neurons that commonly code for both accuracy and confidence, this finding
405 does not necessarily imply that we cannot find dissociation elsewhere in the brain, or even within
406 the same region. Also, Kiani and Shadlen's study was mainly concerned with post-stimulus
407 neuronal activity, and did not investigate pre-stimulus activity, which is the focus of the current
408 study. Finally, the relationship between measures of individual neurons and a population
409 measure like fMRI can be complex, and higher activity for low compared to high confidence has
410 indeed been reported in previous fMRI studies for parietal and frontal areas (Fleming et al.,
411 2012). Thus, we believe that Kiani and Shadlen's work is not incompatible with our current
412 results reporting higher pre-stimulus DAN activity for low confidence trials.

413
414 In another study on the spontaneous fluctuation of activity in the dorsal attention network,
415 Sadaghiani et al. (2009) reported that low pre-stimulus DAN activity led to more hits in an
416 auditory detection task. Sadaghiani et al.'s findings may suggest that low level of attention
417 improves detection capacity, at least for auditory stimuli. However, the increase in detection rate
418 can be attributed to either an enhancement of capacity, or a change of detection criterion towards
419 the liberal direction (responding "yes" more often). Compatible with this interpretation, we
420 previously reported empirical and theoretical data that demonstrated that there is a close
421 relationship between high subjective perceptual ratings and more liberal detection criteria

422 (Rahnev et al., 2011). Thus our results that high pre-stimulus activity in DAN leads to low
423 confidence ratings suggest that Sadaghiani et al.'s findings may be partially explained by a
424 liberal detection bias caused by lack of attention.

425
426 Alternatively, Sadaghiani and colleagues (2009) suggest that the negative relationship between
427 pre-stimulus DAN activity and detection performance may be due to the fact that sounds are not
428 spatial (but see Tark and Curtis, 2009). In contrast with our results, in their study DAN did not
429 show considerable evoked activity. Given the differences in task (discrimination vs. detection)
430 and stimuli used (motion patches vs. auditory tones) between our study and that of Sadaghiani et
431 al. (2009), it is possible that the influence of DAN depends on the specific context of the
432 experiment as well. We should note that our detection theoretic account of Sadaghiani et al.'s
433 findings, and their own interpretation, are not necessarily mutually exclusive.

434
435 In a related study, Hesselmann et al. (2008) investigated the influence of pre-stimulus activity in
436 MT+ on the perception of coherent motion. They found that high pre-stimulus activity in right
437 MT+ biases subjects to perceive coherent motion. Here we used motion that was always coherent
438 and subjects simply needed to identify the direction of that motion. We did not find significant
439 differences in pre-stimulus activity in MT+ for error and correct trials. It is likely that pre-
440 stimulus activity in MT influences the perception of coherence but does not necessarily lead to
441 more accurate identification of the direction of motion.

442
443 Though the finding that attention decreased the variability of the sensory responses (Figure 3d) is
444 compatible with previous research (Mitchell et al., 2007; Bressler and Silver, 2010; Cohen and
445 Maunsell, 2009), it needs to be interpreted with caution. The two categories of motion stimuli

446 used in this study (contracting and expanding) are both expected to activate MT+. Thus, the
447 average activity in MT+ does not directly reflect the evidence for one motion direction or the
448 other. Also, BOLD activity is dominated by scanner and physiological noise (Fox and Raichle,
449 2007) that are not necessarily directly meaningful to the perceptual decision itself.

450

451 Another limitation of the study is that if our model (Figure 1) is correct, one may expect the
452 increase in noise associated with lack of attention to be reflected by higher discrimination
453 accuracy as well, something that we did not find in our dataset. We note that the relative sizes of
454 the impact of an increase in noise on accuracy and confidence may not be the same, and depend
455 on factors such as how the confidence criteria are set. We also had limited statistical power since
456 due to the sparse event-related design employed we only had a limited number of trials (~100)
457 for each subject. Indeed our computational simulations suggest that even if our model (Figure 1)
458 is correct, we may have nevertheless had limited statistical power to detect a significant positive
459 association between pre-stimulus dorsal attention network activity and discrimination accuracy.
460 The simulations were intended as a proof of concept and were not meant as a formal model of
461 how attention influences pre-stimulus BOLD or the internal perceptual distributions.

462 Nevertheless, our simulations demonstrate that the negative finding on the relationship between
463 pre-stimulus BOLD activity on accuracy does not necessarily contradict our model (Figure 1).

464

465 It is also important to note that our signal detection theoretic model is not intended as a general
466 mechanism about how attention, accuracy, and confidence are related since this relationship is
467 likely complex and could depend on the exact setting of the experiment (Kiani and Shadlen,
468 2009; Reynolds and Heeger, 2009; Wilimzig et al., 2008). Our model is also largely agnostic
469 about the specific neural mechanisms that support the generation of confidence ratings. In

470 particular, we do not believe that pre-stimulus dorsal attention network activity codes confidence
471 ratings; instead confidence is likely coded in structures related to metacognition (Fleming et al.,
472 2010, 2012) or self-referential evaluative operations (Northoff and Bermpohl, 2004). Rather, the
473 model depicted in Figure 1 attempts to give a formal explanation of the surprising negative
474 relationship between attention and confidence in certain circumstances such as when attended
475 and unattended stimuli are considered together in the same context.

476

477 We also have to be cautious not to assume that pre-stimulus activity in the dorsal attention
478 network can be absolutely equated with subject's attentional level. There are many other factors
479 that contribute to activity in any large brain network (Fox and Raichle, 2007). Nevertheless, our
480 interpretation of the current data is compatible with previous studies have shown that the brain
481 modulates the attentional level by changing the state of cortical networks (Harris and Thiele,
482 2011) and that the dorsal attention network in particular reflects positively on one's attentional
483 state (Eichele et al., 2008; Sapir et al., 2005).

484

485 Another limitation of the current study is the lack of behavioral effects between conditions,
486 independent of the fMRI results. This is because in this study, unlike in many other behavioral
487 studies of attention, we did not try to experimentally manipulate subjects' attention (Rahnev et
488 al., 2011; Wilimzig et al., 2008). Instead, we relied on spontaneous fluctuations of the neural
489 activity in the dorsal-attention network to indicate different presumed attentional states. Future
490 studies should explore novel ways to determine the attentional state of subjects without explicitly
491 manipulating attention, as well as to directly manipulate attention while measuring neural
492 activity, in order to further investigate the neural mechanism involved in the putative negative
493 relationship between attention and confidence.

494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512

Finally, even though one interpretation of our results is that pre-stimulus activity in the dorsal attention network influences the perception of the stimulus, our data are also consistent with the interpretation that what is being influenced is a late, cognitive stage of the information processing. Nevertheless, we favor the former interpretation because in our previous study (Rahnev et al., 2011) we showed that attention not only led to lower visibility ratings, it also led to more conservative detection biases. Also, these effects were resistant to feedback and payoff manipulations, as if they were automatic and part of the perceptual processes themselves.

To sum up, our results corroborate our previous psychophysical findings (Rahnev et al., 2011) in that lack of attention may lead to liberal subjective biases such as higher subjective ratings, a phenomenon that we call “inattentive inflation of subjective perception” (IISP). Such liberal subjective biases for unattended objects may partially explain why we find it surprising that we fail to perceive unattended objects (such as in inattentive and change blindness experiments) and can perhaps shed some light on why we think we see the whole visual scene in front of us, despite the fact that we seem to only be able to process a few objects effectively within our focus of attention (Kim and Blake, 2005).

513 **Figure Legends**

514

515 **Figure 1. Signal detection model based on Rahnev et al. (2011).** According to signal detection
516 theory (Green and Swets, 1966), stimuli with contracting (red curve) or expanding (blue curve)
517 motion produce overlapping distributions on a single decision dimension. Based on previous
518 research (Gorea and Sagi, 2000) decisions criteria are assumed to be the same for the different
519 attention conditions. The model postulates that lower attention leads to increased variability of
520 the perceptual signal (hence the wider distributions in the lower panel). Due to the higher
521 variability of the “low attention” distributions, they produce a higher percentage of high
522 confidence trials, because the tails of the distributions extend further into the “high confidence”
523 regions.

524

525 **Figure 2. Task design.** We employed a sparse design with inter-trial intervals (ITI) of 20 to 40
526 seconds to allow for the BOLD signal to return to baseline. During that period subjects viewed
527 stationary dots. Trials consisted of 100 ms coherent dot motion that was followed by a “mask” of
528 random motion (400 ms). Subjects were asked to judge the motion direction (expanding /
529 contracting) and then provide a confidence rating on a 1-4 scale.

530

531 **Figure 3. Pre-stimulus dorsal attention network activity was negatively associated with**
532 **confidence.** We identified the dorsal attention network (a) using standard seed-based methods
533 (see Methods). The network included the frontal eye fields and intraparietal regions. (b) Counter-
534 intuitively, the pre-stimulus activity was higher for low confidence than for high confidence
535 trials. There was no reliable difference in pre-stimulus activity for correct compared to error
536 trials. The vertical dashed line shows the average time of the second button press (the confidence

537 response). To facilitate our statistical analysis, we separated **(c)** the trials by confidence and
538 accuracy and plotted the pre-stimulus activity, which was defined as the average BOLD activity
539 in the two scans prior to stimulus onset. **(d)** To check if attention changed the variability of the
540 internal perceptual response, we computed the Fano factor for the distributions of evoked
541 responses of MT+ for the trials that had either low or high pre-stimulus DAN activity. Fano
542 factor was lower in the latter distribution, suggesting that lack of attention increased the
543 variability of the perceptual signal. The error bars represent the standard error of the mean. DAN
544 = dorsal attention network.
545

546

Network	% signal change				Statistical effects		
	High conf, correct	High conf, incorrect	Low conf, correct	Low conf, incorrect	Effect on confidence	Effect on accuracy	Confidence x accuracy interaction
Dorsal attention network	-0.221	-0.272	-0.098	-0.133	F(1,13)=6.68 <i>p</i> =0.02	F(1,13)=0.45 <i>p</i> =0.52	F(1,13)=0.02 <i>p</i> =0.89
Default network	0.175	0.153	0.144	0.076	F(1,13)=0.58 <i>p</i> =0.49	F(1,13)=0.24 <i>p</i> =0.64	F(1,13)=0.14 <i>p</i> =0.71
Alertness network	-0.23	-0.275	-0.215	-0.17	F(1,13)=2.17 <i>p</i> =0.16	F(1,13)=0 <i>p</i> =0.99	F(1,13)=1.31 <i>p</i> =0.27
Dorsal attention network (ICA)	-0.157	-0.244	-0.052	-0.117	F(1,13)=5.61 <i>p</i> =0.03	F(1,13)=1.48 <i>p</i> =0.25	F(1,13)=0.07 <i>p</i> =0.8

547

548 **Table 1. Pre-stimulus activity in other common brain networks.** We computed the average
549 pre-stimulus activity for each trial type (see Figure 3) in different brain networks. The dorsal
550 attention network was also defined in two different ways using either seed-based methods or
551 independent component analysis (ICA, see Methods). One subject was excluded from this
552 statistical analysis because he did not have any high confidence error trials.

553

554

Measure	Pre-stim DAN in low confidence trials	Pre-stim DAN in high confidence trials	Pre-stim DAN in correct trials	Pre-stim DAN in error trials	% correct trials	% high confidence trials
observed value	-0.129	-0.246	-0.172	-0.206	0.626	0.528
fitted value	-0.126	-0.242	-0.173	-0.212	0.624	0.528

555

556 **Table 2. Fit from our computational simulations.** We carried out simple computer simulations

557 (see Methods) to investigate if our model (Figure 1) can fit the observed data from the

558 experiment. The fitting procedure was performed four different times to ensure that the fitting

559 procedure did not get “stuck” in local minima. This table reports the fitted values produced by

560 the first fit which provided the best match to the observed values. The values were generated by

561 simulating 15,000,000 trials.

562

563 **References**

- 564 Anderson JS, Druzgal TJ, Lopez-Larson M, Jeong EK, Desai K, Yurgelun-Todd D (2011)
565 Network anticorrelations, global regression, and phase-shifted soft tissue correction. *Hum*
566 *Brain Mapp* 32:919-934.
- 567 Brainard DH (1997) The Psychophysics Toolbox. *Spat Vis* 10:433-436.
- 568 Bressler DW, Silver MA (2010) Spatial attention improves reliability of fMRI retinotopic
569 mapping signals in occipital and parietal cortex. *Neuroimage* 53:526-533.
- 570 Busey TA, Tunnickliff J, Loftus GR, Loftus EF (2000) Accounts of the confidence–accuracy
571 relation in recognition memory. *Psychonomic Bulletin and Review* 7:26-48.
- 572 Calhoun VD, Adali T, Pearlson GD, Pekar JJ (2001) A method for making group inferences
573 from functional MRI data using independent component analysis. *Hum Brain Mapp*
574 14:140-151.
- 575 Carandini M (2004) Amplification of trial-to-trial response variability by neurons in visual
576 cortex. *PLoS Biol* 2:e264.
- 577 Carrasco M (2011) Visual attention: the past 25 years. *Vis Res* 51:1484-1525.
- 578 Churchland MM, Yu BM, Cunningham JP, Sugrue LP, Cohen MR et al. (2010) Stimulus onset
579 quenches neural variability: a widespread cortical phenomenon. *Nat Neurosci* 3:369-378.
- 580 Cohen MR, Maunsell JH (2009) Attention improves performance primarily by reducing
581 interneuronal correlations. *Nat Neurosci* 12:1594-1600.
- 582 Corbetta M, Shulman GL (2002) Control of goal-directed and stimulus-driven attention in the
583 brain. *Nat Rev Neurosci* 3:201-215.
- 584 Damoiseaux JS, Rombouts SA, Barkhof F, Scheltens P, Stam CJ, Smith SM, Beckmann CF
585 (2006) Consistent resting-state networks across healthy subjects. *Proc Natl Acad Sci U S*
586 *A* 103:13848-13853.

587 Desimone R, Duncan J (1995) Neural Mechanics of Selective Visual Attention. *Ann Rev*
588 *Neurosci* 18:193-222.

589 Eichele T, Debener S, Calhoun VD, Specht K, Engel AK, Hugdahl K, von Cramon DY,
590 Ullsperger M (2008) Prediction of human errors by maladaptive changes in event-related
591 brain networks. *Proc Natl Acad Sci U S A* 105:6173-6178.

592 Fleming SM, Huijgen J, Dolan RJ (2012) Prefrontal contributions to metacognition in
593 perceptual decision-making. *J Neurosci* 32:6117-6125.

594 Fleming SM, Weil RS, Nagy Z, Dolan RJ, Rees G (2010) Relating introspective accuracy to
595 individual differences in brain structure. *Science*, 329:1541-1543.

596 Fox MD, Snyder AZ, Vincent JL, Corbetta M, Van Essen DC, Raichle ME (2005) The human
597 brain is intrinsically organized into dynamic, anticorrelated functional networks. *Proc*
598 *Natl Acad Sci U S A* 102:9673–9678.

599 Fox MD, Corbetta M, Snyder AZ, Vincent JL, Raichle ME (2006) Spontaneous neuronal activity
600 distinguishes human dorsal and ventral attention systems. *Proc Natl Acad Sci U S A*
601 103:10046-10051.

602 Fox MD, Raichle ME (2007) Spontaneous fluctuations in brain activity observed with functional
603 magnetic resonance imaging. *Nat Rev Neurosci* 8:700-711.

604 Fox MD, Snyder AZ, Vincent JL, Raichle ME (2007) Intrinsic fluctuations within cortical
605 systems account for intertrial variability in human behavior. *Neuron* 56:171–184.

606 Fox MD, Zhang D, Snyder AZ, Raichle ME (2009) The global signal and observed
607 anticorrelated resting state brain networks. *J Neurophysiol* 101:3270-3283.

608 Gorea A, Sagi D (2000) Failure to handle more than one internal representation in visual
609 detection tasks. *Proc Natl Acad Sci U S A* 97:12380–12384.

610 Green DM, Swets JA (1966) *Signal detection theory and psychophysics*. New York: Wiley.

611 Gur M, Beylin A, Snodderly DM (1997) Response variability of neurons in primary visual cortex
612 (V1) of alert monkeys. *J Neurosci* 17:2914–2920.

613 Harris KD, Thiele A (2011) Cortical state and attention. *Nat Rev Neurosci* 12:509-523.

614 Hesselmann G, Kell CA, Kleinschmidt A (2008a) Ongoing activity fluctuations in hMT+ bias
615 the perception of coherent visual motion. *J Neurosci* 28:14481–14485.

616 Hesselmann G, Kell CA, Eger E, Kleinschmidt A (2008b) Spontaneous local variations in
617 ongoing neural activity bias perceptual decisions. *Proc Natl Acad Sci U S A* 105:10984-
618 10989.

619 Kiani R, Shadlen MN (2009) Representation of confidence associated with a decision by neurons
620 in the parietal cortex. *Science* 324:759-764.

621 Kim CY, Blake R (2005) Psychophysical magic: rendering the visible ‘invisible’. *Trends Cogn*
622 *Sci* 9:381-388.

623 Kirkpatrick S, Gelatt CD Jr, Vecchi, MP (1983) Optimization by simulated annealing. *Science*
624 220:671–680.

625 Lau H, Rosenthal D (2011) Empirical support for higher-order theories of conscious awareness.
626 *Trends Cogn Sci* 15:365-373.

627 Lu, Z-H, Doshier, BA (2008) Characterizing observers using external noise and observer models:
628 Assessing internal representations with external noise. *Psychological Review* 115:44-82.

629 Lund TE, Norgaard MD, Rostrup E, Rowe JB, Paulson OB (2005) Motion or activity: Their role
630 in intra- and inter-subject variation in fMRI. *Neuroimage* 26:960–964.

631 Macmillan N, Creelman C (2005) *Detection theory: A user’s guide*. New York: Cambridge UP.

632 Maier A, Wilke M, Aura C, Zhu C, Ye FQ, Leopold DA (2008) Divergence of fMRI and neural
633 signals in V1 during perceptual suppression in the awake monkey. *Nat Neurosci* 11:1193-
634 2000.

635 Mitchell JF, Sundberg KA, Reynolds JH. (2007) Differential attention-dependent response
636 modulation across cell classes in macaque visual area V4. *Neuron* 55:131-141.

637 Murphy K, Birn RM, Handwerker DA, Jones TB, Bandettini PA (2009) The impact of global
638 signal regression on resting state correlations: are anti-correlated networks introduced?
639 *Neuroimage* 44:893-905.

640 Northoff G, Bermpohl F (2004) Cortical midline structures and the self. *Trends Cogn Sci* 8:102-
641 107.

642 Posner MI, Snyder CR, Davidson BJ (1980) Attention and the detection of signals. *J Exp*
643 *Psychol* 109:160-174.

644 Rahnev D, Maniscalco B, Graves T, Huang E, de Lange FP, Lau H (2011) Attention induces
645 conservative subjective biases in visual perception. *Nature Neuroscience* 14:1513-1515.

646 Reynolds JH, Heeger DJ (2009) The normalization model of attention. *Neuron* 61:168-185.

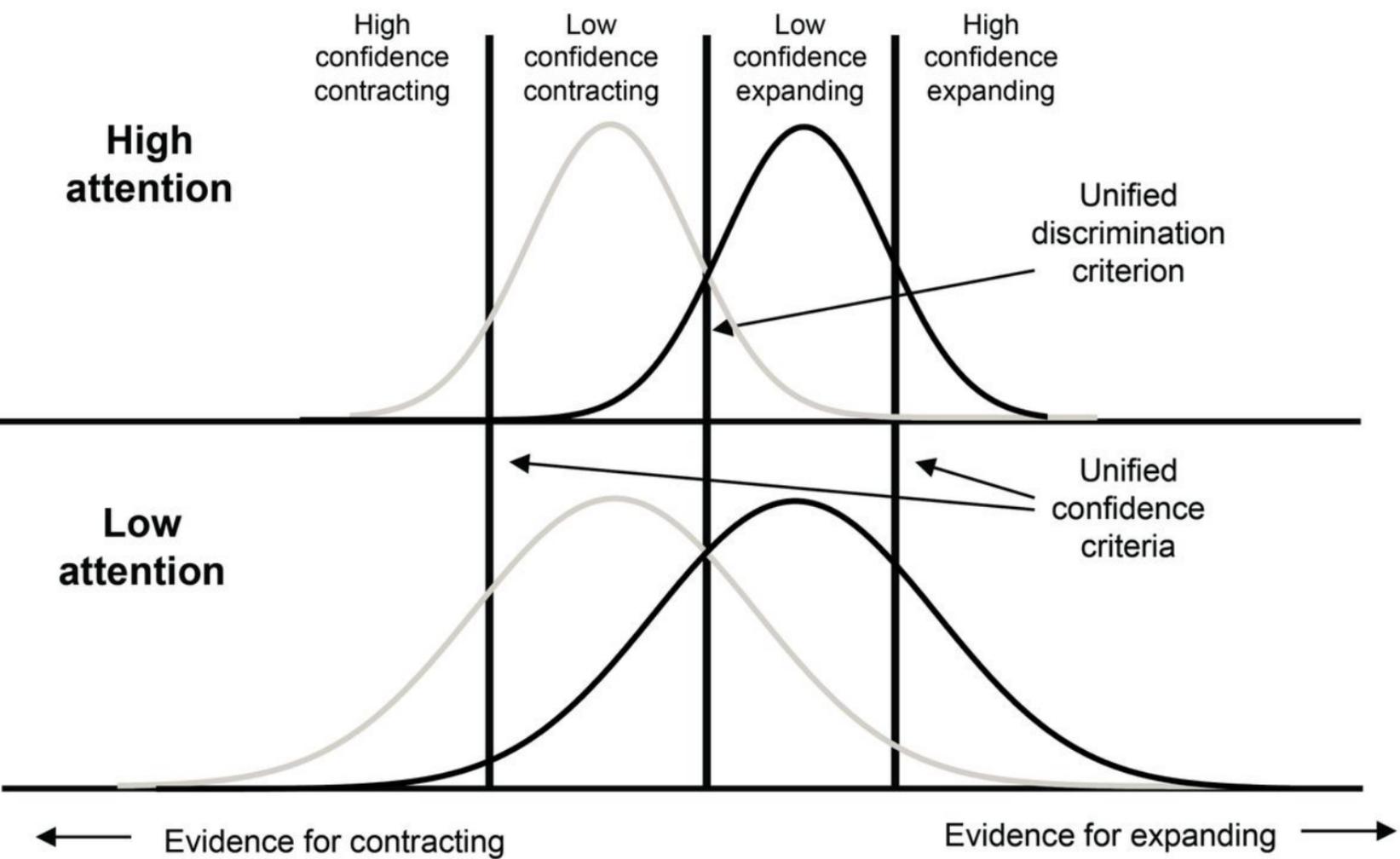
647 Sadaghiani S, Hesselmann G, Kleinschmidt A (2009) Distributed and antagonistic contributions
648 of ongoing activity fluctuations to auditory stimulus detection. *J Neurosci* 29:13410-
649 13417.

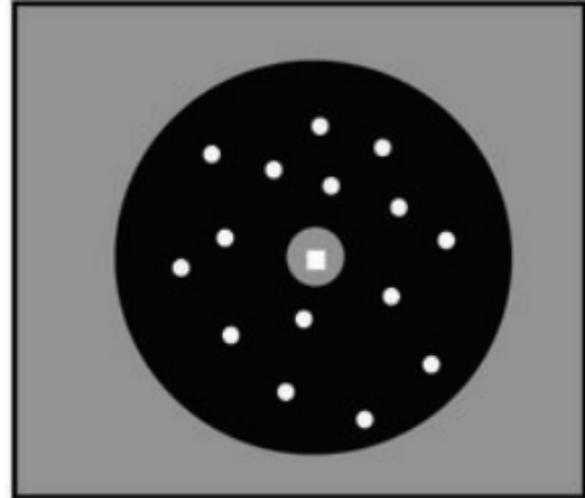
650 Sadaghiani S, Scheeringa R, Lehongre K, Morillon B, Giraud AL, Kleinschmidt A (2010)
651 Intrinsic connectivity networks, alpha oscillations, and tonic alertness: a simultaneous
652 electroencephalography/functional magnetic resonance imaging study. *J Neurosci*
653 30:10243-10250.

654 Sapiro A, d'Avossa G, McAvoy M, Shulman GL, Corbetta M (2005) Brain signals for spatial
655 attention predict performance in a motion discrimination task. *Proc Natl Acad Sci U S A*
656 102:17810 –17815.

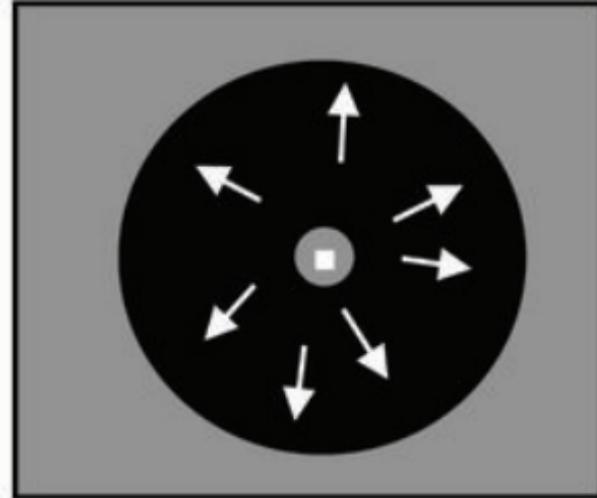
657 Sperling G (1960) The information available in brief visual presentations. *Psychol Monographs*
658 74:1-29.

659 Tark KJ, Curtis CE (2009) Persistent neural activity in the human frontal cortex when
660 maintaining space that is off the map. Nat Neurosci 12:1463-1468.
661 Wilimzig, C., Tsuchiya, N., Fahle, M., Einhäuser, W., & Koch, C. (2008). Spatial attention
662 increases performance but not subjective confidence in a discrimination task. J Vis 8:1-
663 10.
664

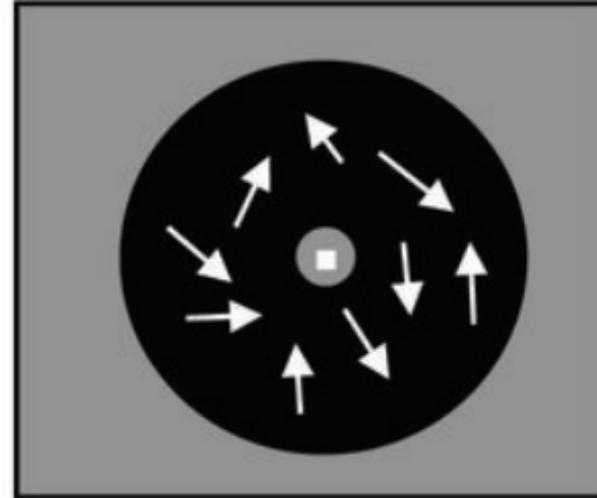




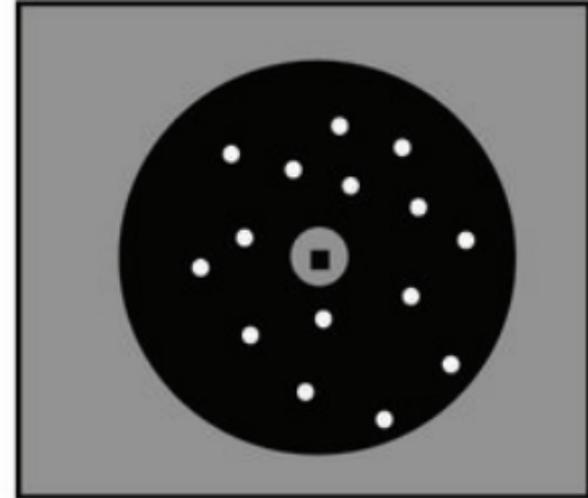
ITI
(20-40 s)



Coherent
Motion
(100 ms)



Random
Motion
(400 ms)



Response
cue
(up to 6 s)

Time



