1
2 **Simultaneous multi-area recordings suggest a novel hypothesis about how attention improves performance**

3 Douglas A. Ruff and Marlene R. Cohen
4 Department of Neuroscience and Center for the Neural Basis of Cognition, University of Pittsburgh, Pittsburgh, Pennsylvania 15260; email:
5 ruffd@pitt.edu

6

7 **Abstract -** Visual attention dramatically improves subjects' ability to see and also modulates the

8 responses of visual and oculomotor neurons. Despite hundreds of studies demonstrating the co-

9 occurrence of behavioral and neuronal effects of attention [1], the relationship between neuronal

10 modulations and improved performance remains unknown. There are three dominant hypotheses

11 (Figure 1A): that attention 1) improves visual information coding [2,3], 2) improves communication

12 between visual and oculomotor brain areas [4], or 3) reduces uncontrolled fluctuations in the

13 animals' cognitive state, thereby improving performance and rendering observed effects of

14 attention on response variability as epiphenomena [5]. Here we show that none of these hypotheses

15 account for observed perceptual improvements and we provide evidence for a novel hypothesis:

16 that the well known effects of attention on firing rates and shared response variability in visual

17 cortex reshape the representation of attended stimuli such that they more effectively drive

18 downstream neurons and guide decisions without explicitly changing the weights relating

19 sensory responses to downstream neurons or behavior (Figure 1B). We assessed each

20 hypothesized mechanism by recording simultaneously from groups of neurons in two stages of

21 visuomotor processing: area MT, which encodes motion information [6] and the superior colliculus

22 (SC), where neuronal responses are either visual, oculomotor, or intermediate [7-9] and have been

23 hypothesized to be involved in computing perceptual decisions [10]. Constraining our analyses by

24 the animals' behavior and the simultaneous recordings from both areas allowed us to reject each

25 prior hypothesis and greatly clarified the relationship between attention, neuronal responses and

26 behavior. Our results suggest a path toward understanding the neural underpinnings of

27   perception and cognition in health and disease by analyzing neuronal responses in ways that are

28   constrained by behavior and interactions between brain areas.

29

30   **Main -** We compared evidence for and against three hypothesized attention mechanisms using

31   neuronal responses collected while two rhesus monkeys performed a widely studied motion

32   direction change-detection task Figure 1C; [11,12,13]. The animals' performance was greatly affected

33   (Figure 1D) by a cue instructing them to shift spatial attention between a stimulus within the

34   same or opposite hemifield as the joint receptive fields of several dozen neurons that were

35   recorded on multielectrode probes in MT (Figure 1E, red points) and the SC (blue points). MT

36   and the SC represent different stages of perceptual decision-making and therefore provide the

37   opportunity to evaluate each hypothesized attention mechanism. MT contributes to motion

38   perception [6,14]. The SC is thought to play many roles in visually guided tasks [7-9], decision-making
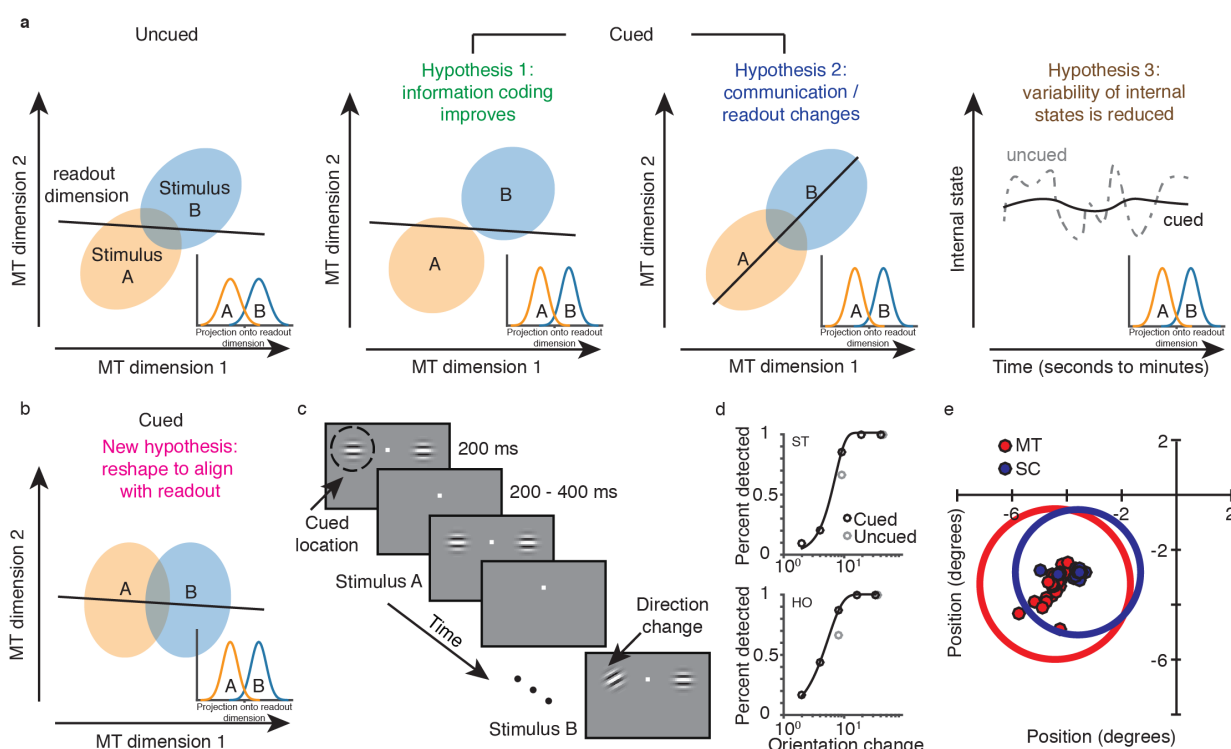
39   [10,15] and attention[16].

40

Figure 1

Figure 1. Hypotheses and methods. (A) Schematics describing predominant hypotheses about links between attention, visual cortical activity, and behavior. The left plot depicts MT population responses to two visual stimuli plotted along two dimensions in population response space (e.g. the first two principal components; see Methods) and a readout dimension which represents the visual information that is communicated to neuronal populations involved in planning behavior during the uncued condition. The insets depict projections of the population responses onto the readout dimension. Hypothesis 1 is that the MT representations of the two stimuli become more easily distinguishable (e.g. by separating the distributions of responses to the two stimuli). In this scenario, the distributions of projections along even a suboptimal readout axis may also be more separable. Hypothesis 2 suggests that attention changes the communication or readout dimensions, such that projections of responses to the two stimuli are more separable. Hypothesis 3 proposes that attention-related changes in visual cortex are epiphenomenal. Instead, it suggests that attention reduces uncontrolled variability in the animals' internal states, which might produce less variable neuronal population responses and therefore more separable projections onto the readout dimensions. (B) Our new hypothesis: attention reshapes population responses so they are better aligned with relatively static readout dimensions. (C) Direction change-detection task with cued attention. The stimuli before the change can be thought of as stimulus A and the changed stimulus can be thought of as stimulus B in the schematics in (A). (D) Psychometric curves from two example sessions (monkey ST, top, monkey HO, bottom) with best-fitting Weibull functions. Attention improved detection of median difficulty trials by 25% on average across all experiments. (E) Receptive field (RF) centers of recorded units from the same example session as in the top plot in (D). Dots represent the RF center (red, MT; blue, SC). The circle represents the size and location of the median RF from each area.

67   The three predominant attention hypotheses make different predictions about how attention

68   should affect MT and the SC in our task. The first (information coding) hypothesis predicts that

69   attention improves the motion direction information encoded in MT. The second

70   (communication) hypothesis posits that attention improves the stimulus information that is

71   shared between MT and the SC. The third (internal states) hypothesis suggests that attention

72   reduces uncontrolled fluctuations in the animals' cognitive states, allowing them to focus on their

73   psychophysical task. In this scenario, the widely observed attention-related reduction in response

74   variability in visual cortex [11,13,17] would be an epiphenomenon of decreased variability in

75   cognitive states.

76

77   Past studies have evaluated these hypotheses by analyzing the responses of individual neurons or

78   pairs of neurons, which typically lack the statistical power to reveal a strong link to behavior.

79   Using our data set, we replicated the observations that have been used as evidence in favor of all

80   three hypotheses. Consistent with previous studies evaluating the information coding hypothesis

81   [2,3,18], we found that attention increased the trial-averaged responses of neurons in both MT and

82   the SC (Supplemental Figure 1A and B). Consistent with studies evaluating both the information

83   coding and internal states hypotheses [5], we found that attention decreased the extent to which the

84   trial to trial fluctuations in neuronal responses to repeated presentations of the same stimulus are

85   shared between pairs of MT neurons[11,13,17] (quantified as the average spike count or noise

86   correlation, or r_{SC}[19]; Supplementary Figure 1C). Consistent with studies evaluating the

87   communication hypothesis, attention increases correlated variability between the two areas [12,20]

88   (Supplemental Figure 1C). This attention-related increase was weakly dependent on the visual

89   responsivity of SC neurons (Supplemental Figure 2).

90

91   Most previous studies aim to connect the observed attention-related changes in single neurons or

92   pairs of neurons to behavior via an explicit or implicit model. Recordings from neuronal

93   populations provide the opportunity to do so directly. We reasoned that analyzing the

94   relationship between populations of simultaneously recorded neurons in multiple brain areas

95   with the animals' behavior would provide insight into the relative importance of each

96   hypothesized mechanism.

97

98   The internal states hypothesis provides a challenge for the approach of linking populations of

99   visual neurons with behavior because it proposes that there is no link between attention-related

100   changes in MT or other visual areas and performance. Instead, it posits that attention limits mind

101   wandering, and the changes in performance and neuronal responses simply reflect improved

102   stability in internal states [5]. There are two reasons this hypothesis seems unlikely to account for

103   our data. First, the effects of spatial attention are spatially specific (e.g. correlated variability

104   increases in one hemisphere while decreasing in the other, even when neurons in the two

105   hemispheres are simultaneously recorded[11]), meaning that reductions in the variability of global

106   cognitive processes like arousal and motivation are unlikely to account for the attention-related

107   changes in visual cortex. Further, it is not obvious how reductions in fluctuations in internal

108   states could account for the attention-related increases in firing rates observed in spatial attention

109   studies like ours (Supplementary Figure 1), let alone the more complex firing rate changes

110   associated with feature attention [21].

111

112   This hypothesis can also be addressed using a population-analysis approach, by using the

113   responses of MT and SC neurons to attempt to quantify the variability in the animals' internal

114   states. We reasoned that fluctuations in internal states would 1) often occur at timescales longer

115   than the 400-600 ms between stimulus presentations in our task and 2) affect the covariability of

116   neurons in many brain areas, including both MT and the SC. We performed principal

117   components analysis on population responses to the identical visual stimuli that occurred before

118   the direction change on each trial (e.g. stimulus A in Figure 1; see Methods) in each attention

119   condition. Because the only variability in those population responses is internally generated, the

120   first principal component (PC) represents the axis of greatest shared variability in the population

121   of neurons in each brain area.
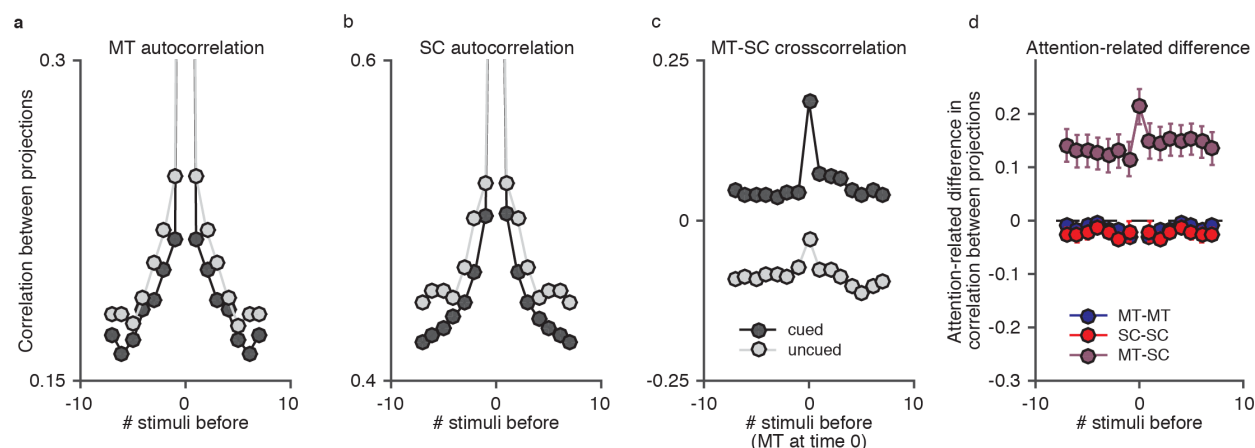
122

Figure 2

124   *Figure 2. Attention has opposite effects on slow fluctuations in neuronal population responses*
125   *within and across areas. (A, B) Autocorrelations between projections onto the first PCs of*
126   *population responses to repeated presentations of the same visual stimulus in (A) MT, and (B)*
127   *the SC. The x-axis plots time lag in units of stimulus presentations (400-600 ms; see Methods).*
128   *(C) Cross correlation between projections onto the first PCs in MT and the SC (same data and*
129   *plotting conventions as in A and B). (D) Attention-related difference in autocorrelation or cross*
130   *correlations between the projections in the previous plots. Error bars represent standard error*
131   *of the mean. Attention was associated with a statistically significant decrease in autocorrelation*
132   *overall (t-tests, p<.05) in both areas and in 11/15 individual MT data sets and 9/15 SC data sets*
133   *(t-tests, p<0.05 with a Bonferroni correction) and a significant increase in cross correlation*
134   *overall (t-test, p<.001) and in 11/15 individual data sets.*

135

136   The autocorrelation functions of projections onto the first PC show that there is indeed response

137   variability in each area that fluctuates slowly and is reduced by attention (Figures 2A, B), which

138   is not as readily observable by computing noise correlations between pairs of neurons

139   (Supplementary Figure 1C). However, the cross-correlation of projections onto the first PC in

140   MT and the SC showed a qualitatively different time course than the autocorrelation function.

141   Furthermore, not only did attention not reduce the covariability of these signals between these

142   two areas (and presumably brain wide), attention *increased* the slow variability that is shared

143   between areas (Figure 2C and Supplementary Figure 1C). These results are in conflict with the

144   idea that the attention-related decrease in covariability within each area is a byproduct of a

145   decrease in uncontrolled fluctuations in internal states, because such a decrease should be brain-
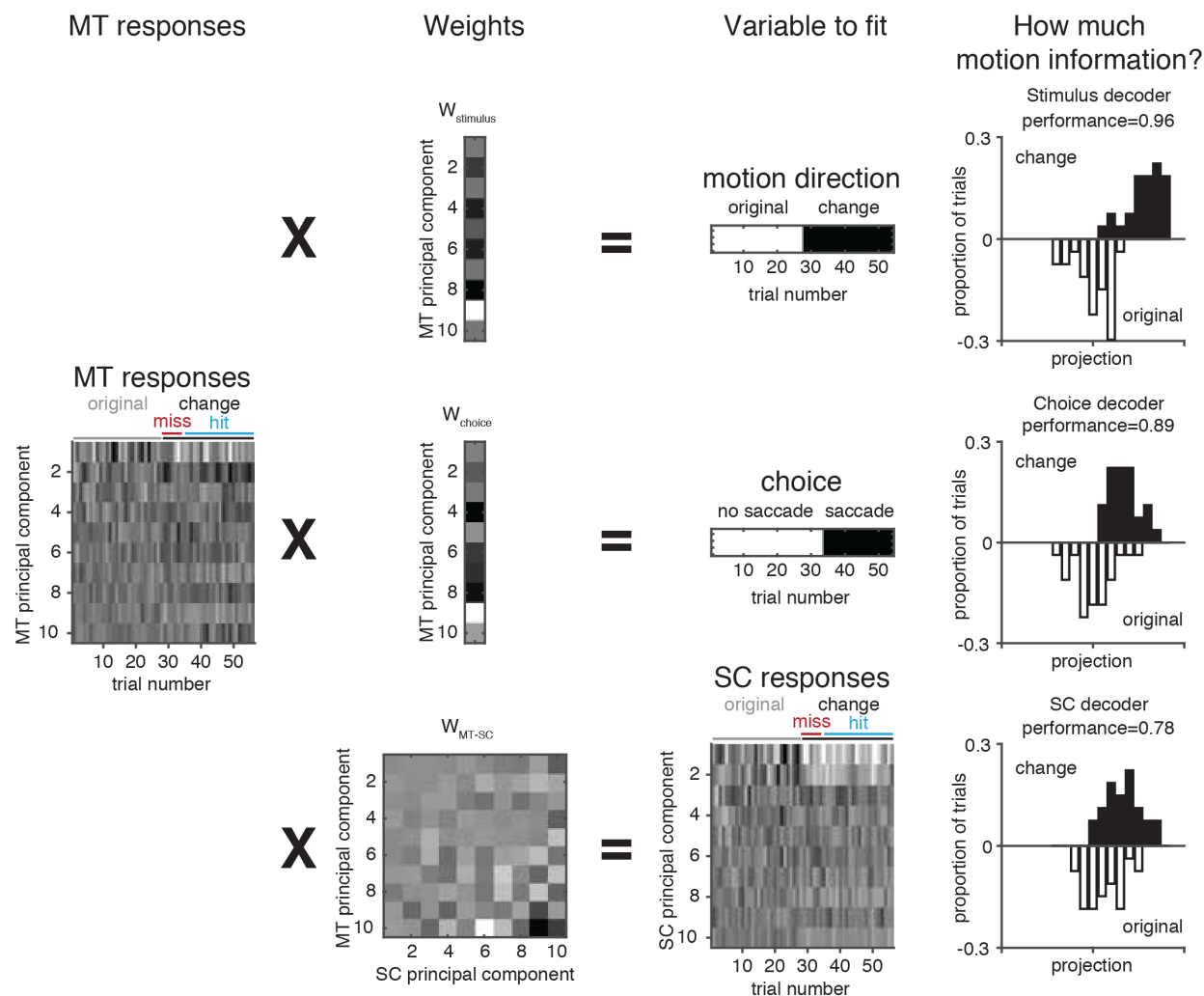
146   wide.

147

Figure 3

*Figure 3 – Schematic of our decoding procedure. We used linear regression to find the weights (second column) that best relate the **first ten principal components of the MT population's response** (left) to the visual stimulus (Stimulus decoder; top row), the animal's choice (Choice decoder; middle row), or the projections of the responses of the population of simultaneously recorded SC neurons (SC decoder; bottom row). We assessed the performance of each decoder by decoding stimulus information from MT responses on a separate set of trials using each set of weights (right column). See methods for detailed decoding and cross validation procedures.*

To investigate the remaining hypotheses, we determined whether attention affects the amount of

stimulus information that can be decoded from the population of MT neurons using linear

decoders that are optimized to a) dissociate between the original and changed stimuli (Stimulus

decoder in Figure 3), b) predict the animals' choices (whether or not they made an eye

161   movement; Choice decoder), or c) predict the activity of the population of SC neurons we

162   recorded (SC decoder).

163

164   The information coding hypothesis posits that attention improves the stimulus information that

165   could be gleaned by an optimal stimulus decoder. However, attention did not significantly affect

166   the performance of an optimal decoder in our data set, even when we used a decoder optimized

167   separately for each attention condition (Figure 4A, left bars). Furthermore, theoretical work

168   suggests that the effects of attention on the stimulus information that can be decoded from small

169   neuronal populations are likely to be even more minimal for larger populations [22-24].

170

171   The communication hypothesis posits that attention improves communication between visual and

172   decision areas, meaning that it should change the weights relating MT responses to either

173   behavior or SC responses. We found that attention had larger effects on the stimulus information

174   that is related to the animals' choices on individual trials (Figure 4A, middle bars) or that is

175   shared with the SC (Figure 4A, right bars). However, this could arise from either a weight

176   change (Figure 1A) or a change within MT, to better align MT responses with static readout

177   weights (Figure 1B).

178

179

180   *Figure 4. Effects of attention on the stimulus*
181   *information that can be decoded from small*
182   *populations of MT neurons. (A) Ability of a cross-*
183   *validated linear decoder to distinguish the original*
184   *from changed stimuli for each decoder. Error bars*
185   *represent SEM. The effect of attention was significant*
186   *for the Choice and SC decoders (paired t-tests, p<.05)*
187   *but not for the Stimulus decoder (p=0.28). The effects*
188   *of attention on the Choice and SC decoders were*
189   *greater than for the stimulus decoder (paired t-tests,*
190   *p<.05), but not significantly different from each other*
191   *(p=0.21). (B) Weight swapping analysis demonstrates*
192   *that decoding performance was typically better using*
193   *the MT responses from the cued condition and the*
194   *choice decoder weights from the uncued condition (y-*
195   *axis) than using the MT responses from the uncued*
196   *condition and the Choice decoder weights from the*
197   *cued condition (x-axis; paired t-test, p<.05). (C)*
198   *Same, using the weights from the SC decoder (paired*
199   *t-test, p<.05).*
200

201   We found no evidence that attention changes weights

202   relating MT responses to SC responses or behavior.

203   Both the Choice and SC decoders gleaned more

204   stimulus information from MT responses in the

205   attended than unattended condition, even when we

206   used the weights computed in the opposite attention

207   condition for which they were calculated (Figures 4B

208   and 4C). Together, these neuronal population analyses

209   that use the animals' behavior and the activity of

210   downstream neurons to assess the hypothesized

211   attention mechanisms reveal that none of the three

212   previously hypothesized mechanisms provides a

Figure 4
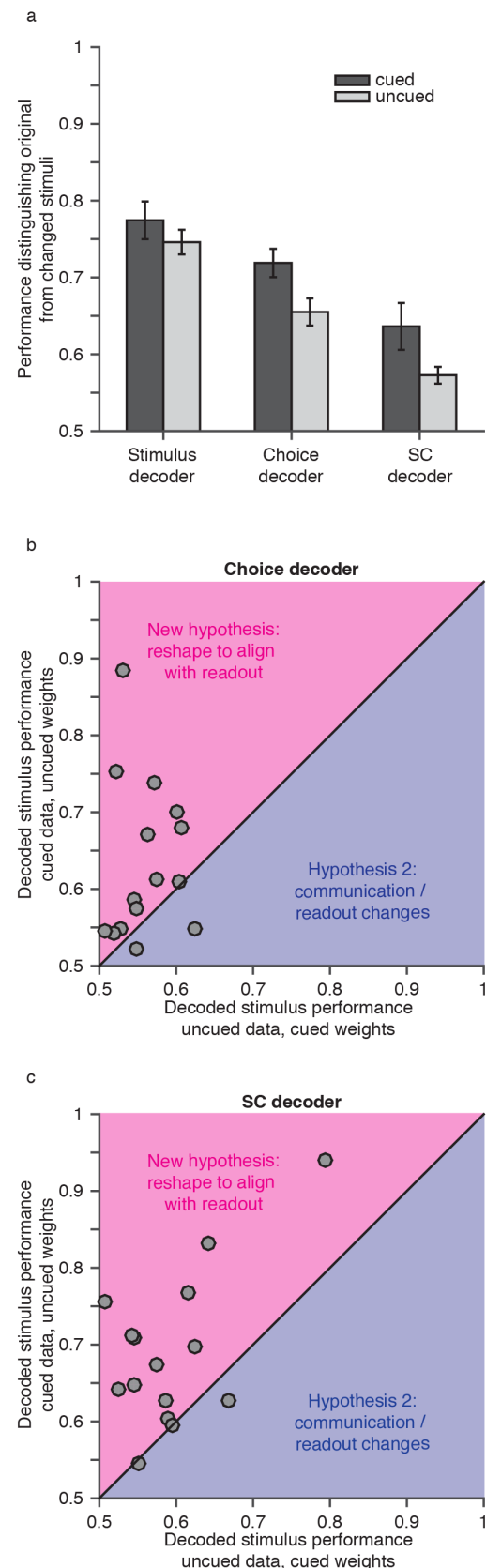
213   satisfactory account of the observed attention-related behavioral improvement.

214

215   Our data support the hypothesis that attention reshapes the representation of attended stimuli to

216   more effectively guide behavior (Figure 1B). In this scenario, the critical changes are in visual

217   cortex. However, this does not result in a large improvement in the stimulus information that can

218   be gleaned by an optimal stimulus decoder. Instead, the modulated neuronal activity in MT better

219   aligns with the readout dimensions using relatively static weights. Modifying stimulus

220   representations without changing readout dimensions may be simple mechanistically[25], allowing

221   animals to discriminate stimulus features (e.g. motion direction) of any stimuli they encounter in

222   the natural world, without optimizing readout for the reduced stimulus sets used in laboratory

223   tasks like ours [18].

224

225   It is important to note that not only is this reshaping of the representation of the attended

226   stimulus consistent with the simple signatures of attention that have been observed for many

227   years (e.g. changes in firing rate gain or noise correlations [2]), it may be a direct result of those

228   changes. For example, we showed that in this task, monkeys' choices are well-aligned with the

229   axis in population space that explains the most correlated noise [13], perhaps because that axis is

230   well-positioned to decode the motion direction of any stimulus [18]. If so, reducing noise

231   correlations and increasing firing rate gains would improve the stimulus information projected

232   along that readout axis (following the intuitions in [26]). Attention-related changes in gains and

233   correlations have been observed for many years. The key insight that they may improve

234   performance by aligning the responses of populations of visual neurons to fixed readout

235   dimensions was only possible using population analyses constrained by the animals' behavior

236   rather than focusing on one or two neurons.

237

238   By combining psychophysics and multi-neuron, multi-area recordings, our study provides a

239   framework for understanding the neuronal mechanisms underlying a wide variety of sensory,

240   cognitive, and motor processes. We showed that the effects of covert spatial attention are best

241   described as a change in the visual information used to guide behavior. In the future, it will be

242   interesting to use the same approach to determine whether a similar mechanism can account for

243   behavioral changes associated with other cognitive processes (e.g. feature attention) that might

244   seem more likely to change the weights relating stimulus information to downstream neurons or

245   behavior. Further, many neuropsychiatric disorders (including disorders of attention, Autism, and

246   schizophrenia) are thought to involve changes in the same computations thought to underlie

247   attention [27]. An exciting possibility is that these changes might be identified and potential

248   therapies evaluated in animal models using the combination of behavioral evaluation and multi-

249   neuron, multi-area recordings that we described here.

250

251   Methods Summary

252

253   The subjects were two adult male rhesus monkeys (*Macaca mulatta*, 8 and 9 kg). All animal

254   procedures were approved by the Institutional Animal Care and Use Committees of the

255   University of Pittsburgh and Carnegie Mellon University. Using linear 24 channel moveable

256   probes (Plexon), we simultaneously recorded extracellular activity from direction-selective

257   neurons in area MT and neurons in the superior colliculus that responded either visually, prior to

258   a saccade, or both. Before beginning the experiment, we searched for neurons in both areas that

259   had overlapping spatial receptive fields (Figure 1E) as determined by mapping with both drifting

260   gratings and a delayed saccade task. The monkeys performed a direction change-detection task

261   that commenced upon fixation of a central spot (Figure 1C). Two drifting Gabor stimuli, whose

262   direction was selected to drive the recorded population of MT neurons well, flashed on and off

263   until the direction of one stimulus changed at a random, unsignaled time. The monkeys signaled

264   detection of the change by making a saccade to the changed stimulus within 450 ms of its onset.

265   The location of the stimulus change was cued using instruction trials prior to each block of trials

266   and the cue was valid 80% of the time. On the other 20% of trials, the change happened at the

267   uncued location. In order to encourage fixation on longer trials, catch trials, where no stimulus

268   changed direction and monkeys were rewarded for maintaining fixation, were intermixed. We

269   analyzed spiking activity in response to all visual stimuli except the first stimulus in each trial.

270

271   Materials and Methods

272   The subjects were two adult male rhesus monkeys (*Macaca mulatta*, 8 and 9 kg). All animal

273   procedures were approved by the Institutional Animal Care and Use Committees of the

274   University of Pittsburgh and Carnegie Mellon University.

275   We presented visual stimuli using custom software (written in MATLAB using the

276   Psychophysics Toolbox [28,29] on a CRT monitor (calibrated to

277   linearize intensity; 1024 768 pixels; 120 Hz refresh rate) placed 54 cm from the animal. We

278   monitored eye position using an infrared eye tracker (Eyelink 1000; SR Research) and recorded

279   eye position and pupil diameter (1000 samples/s), neuronal responses (30,000 samples/s), and

280   the signal from a photodiode to align neuronal responses to stimulus presentation times (30,000

281   samples/s) using hardware from Ripple.

282

283   Behavioral Task

284   As previously described [11], a trial began when the monkey fixated a small, central spot within a

285   1.25° per side, square fixation window in the center of a video display while two peripheral

286   Gabor stimuli (one overlapping the receptive fields of the recorded neurons, the other in the

287   opposite visual hemifield) synchronously flashed on (for 200 ms) and off (for a randomized

288   period between 200-400 ms) until, at a random, unsignaled time, the direction of one of the

289   stimuli changed from that of the preceding stimuli (Figure 1C). The monkey received a liquid

290   reward for making a saccade to the stimulus that changed. Attention was cued in blocks of 50-

291   100 trials, and alternated between blocks where attention was cued to either the left or the right

292   stimulus. In each block, the direction change occurred at the cued stimulus on 80% of trials, and

293   at the uncued stimulus in 20% of trials (all uncued changes used either the middle or largest

294   orientation change, Figure 1D). Catch trials, where no stimulus changed direction and the

295   monkey was rewarded for maintaining fixation, were randomly intermixed throughout each

296   block. Psychometric data were fit with Weibull functions. Before recording commenced, the

297   monkeys were extensively trained to have stable thresholds across a range of spatial locations (3-

298   6 months). During recording, the size, location, spatial frequency and direction of the Gabor

299   stimuli were selected to drive the recorded units well and the direction of all of the stimuli prior

300   to the orientation change was constant throughout a recording session.

301

302   Electrophysiological Recordings

303   We simultaneously recorded extracellularly from single units and sorted multiunit clusters (the

304   term "unit" refers to either). The dataset consisted of a total of 306 responsive MT units and 345

305   responsive SC units total (36-58 units per session, mean 20 in MT, 24 in the SC for Monkey HO;

306   36-53 units per session, mean 21 in MT, 22 in SC for Monkey ST) in both MT and the SC in the

307   right hemisphere using moveable, linear 24-channel V-probes (Plexon; inter-electrode spacing in

308   MT = 50μm, SC = 100μm). We presented visual stimuli and tracked eye position as previously

309   described [12]. The data presented are from 6 days of recording for Monkey HO and 9 days of

310   recording for Monkey ST. Each day consisted of multiple blocks of the attention task (Figure 1C;

311   mean 1015 of trials for Monkey HO, 745 for Monkey ST) preceded by receptive field mapping

312   using a delayed saccade task and direction tuning during passive fixation.

313   Data Analysis

314   All spike sorting was done offline manually using Offline Sorter (version 3.3.5; Plexon). We

315   based our analyses on both single units and multiunit clusters and use the term "unit" to refer to

316   either. All neuronal analyses prior to Figure 3 used spike count responses between 50-250 ms

317   after stimulus onset to account for visual latencies in the two areas. To remove response

318   contamination from eye movements during change stimuli, data presented in the decoding

319   analyses in Figure 3 and 4 used shorter response windows. Responses to both unchanged and

320   changed stimuli were measured from 50-185 ms after stimulus onset for monkey HO and 50-220

321   ms for monkey ST. These times were selected based on the distribution of each animal's reaction

322   times, with the goal of maximizing the number of trials that could be included in the analyses.

323   Trials with reaction times that began during those windows were excluded. To minimize the

324   impact of adaptation on our results, we did not analyze the first stimulus presentation in each

325   trial. We only analyzed a recorded MT unit if its stimulus-driven firing rate was 10% higher than

326   its firing rate as measured in the 100 ms prior to the onset of the first stimulus. We only analyzed

327   a recorded SC unit if its stimulus-driven firing rate was 10% higher than its firing rate as

328   measured in the 100 ms prior to the onset of the first stimulus or if its response during a 100 ms

329   epoch prior to a saccade on correct trials to the contralateral side was 10% larger than that same

330   baseline. Stimulus presentations during which a microsaccade was detected were excluded from

331   analyses [12,30]).

332

333   For firing rate analyses in Supplemental Figure 1A and B, attention indices were calculated using

334   average spike counts on the stimulus presentation prior to correct detections of the intermediate

335   change amount depending on whether attention was directed into or out of the receptive fields of

336   the recorded neurons using the formula $(attend_{in} - attend_{out})/(attend_{in} + attend_{out})$. Significance of

337   individual units was determined by a paired t-test ($p < 0.05$).

338

339   Noise correlations

340

341   We defined the correlated variability of each pair of simultaneously recorded units (quantified as

342   spike count correlation or r_{SC} [19]) as the Pearson correlation coefficient between the responses of

343   the two units to repeated presentations of the same stimulus. This measure of r_{SC} represents noise

344   correlations rather than signal correlations because the responses used in this analysis were

345   always to an identical visual stimulus. For Supplemental Figure 1C, we included responses from

346   stimulus presentations 2 though 10 from trials that ended with either a hit, miss or correct catch

347   trial and that were immediately followed by the maintenance of fixation and continuation of the

348   trial. We z-scored responses as a function of the stimulus presentation number in each trial and

349   then pooled data across stimulus presentations before calculating noise correlations. Results did

350   not qualitatively change if we did not perform this z-score procedure. For Supplemental Figure

351   1D, we included data from all stimulus presentations prior to the change stimulus (except the

352   first) and sorted them depending on what the behavioral outcome was on the subsequent stimulus

353   presentation. Pairs of units that were recorded on the same electrode were not included in

354   correlation analyses. The data presented in Supplementary Figures 1C consisted of 3,285 MT

355   pairs, 3,948 SC pairs and 6,934 between area pairs.

356

357   Slow timescale auto- and cross-correlations

358

359   To test the hypothesis that attention affects uncontrolled fluctuations in internal states, we

360   created a procedure to identify slow fluctuations in population responses. We plotted the

361   responses of the populations of simultaneously recorded MT or SC neurons to the stimuli before

362   the motion direction change (excluding the first stimulus) in a high dimensional space in which

363   each neuron's response comprises one axis. We performed principal components analysis (PCA)

364   on this cloud of points. Because the visual stimuli in this subset of the data were all identical, the

365   first principal component (PC) represents the axis that captures the most variance in non-

366   stimulus related population responses. We searched for slow fluctuations in internal states by

367   measuring the auto- and cross-correlations in projections onto this first PC in each area.

368

369   Decoding

370   We focused our decoding analyses on trials in which the third largest (middle) direction change

371   occurred, because changes of that magnitude occurred in both attention conditions. We used the

372   decoding strategy schematized in Figure 3. We began by constructing four matrices in each

373   attention condition: 'MT responses' (a # MT neurons x 2*# trials matrix of MT responses to the

374   stimuli before the direction change and the changed stimulus on the relevant trials), 'motion

375   direction' (a 1 x 2*# trials vector of zeros for the stimulus before the change, referred to as

376   'original', and ones for the changed stimulus, referred to as 'change'), 'choice' (a 1 x 2*# trials

377   vector of zeros for stimulus presentations on which the animal did not make an eye movement,

378   referred to as 'no saccade', and ones when the animal made an eye movement, referred to as

379   'saccade'), and 'SC responses' (a # SC neurons x 2*# trials matrix of SC responses to the stimuli

380   before the direction change and the changed stimulus on the relevant trials).

381

382   We cross validated by holding out the two stimulus presentations (for the original and changed

383   stimuli) from one trial at a time to perform the rest of our analyses. To reduce the number of

384   weights we needed to fit and therefore improve our confidence in the weights we did fit, we

385   performed PCA on the MT and SC responses to find the first 10 PCs in each area. We then

386   performed linear regression to find the weight vectors (for the Stimulus and Choice decoders) or

387   weight matrices (for the SC decoder) that related projections along the first ten MT PCs plus a

388   vector of ones to 'motion direction', 'choice', or projections along the first 10 SC PCs in each

389   attention condition.

390

391   We assessed the stimulus information in each decoder (Figure 4) by multiplying projections of

392   MT responses to the original and changed stimuli from the held-out trial by the fitted weights

393   and either determining whether those weighted sums correctly classified the stimuli as original or

394   changed (Stimulus and Choice decoders) or whether a linear classifier correctly classified those

395   stimulus presentations on the basis of the predicted SC responses (SC decoder).

396

397   Data Availability

398

399   The data that support the findings of this study are available from the corresponding author upon

400   reasonable request.

401

402   Code Availability

403

404   The MATLAB code that was written to perform the analyses in this paper are available from the

405   corresponding author upon reasonable request.

406

407   References:

408   1   Carrasco, M. Visual attention: the past 25 years. *Vision Res* **51**, 1484-1525,
409   doi:10.1016/j.visres.2011.04.012 (2011).
410   2   Maunsell, J. H. R. Neuronal Mechanisms of Visual Attention. *Annu Rev Vis Sci* **1**, 373-391,
411   doi:10.1146/annurev-vision-082114-035431 (2015).
412   3   Reynolds, J. H. & Chelazzi, L. Attentional modulation of visual processing. *Annual Review*
413   *of Neuroscience* **27**, 611-647 (2004).
414   4   Fries, P. Rhythms for Cognition: Communication through Coherence. *Neuron* **88**, 220-
415   235, doi:10.1016/j.neuron.2015.09.034 (2015).
416   5   Ecker, A. S., Denfield, G. H., Bethge, M. & Tolias, A. S. On the Structure of Neuronal
417   Population Activity under Fluctuations in Attentional State. *J Neurosci* **36**, 1775-1789,
418   doi:10.1523/JNEUROSCI.2044-15.2016 (2016).
419   6   Born, R. T. & Bradley, D. C. Structure and function of visual area MT. *Annual review of*
420   *neuroscience* **28**, 157-189, doi:10.1146/annurev.neuro.26.041002.131052 (2005).
421   7   Wurtz, R. H. & Albano, J. E. Visual-motor function of the primate superior colliculus.
422   *Annual review of neuroscience* **3**, 189-226, doi:10.1146/annurev.ne.03.030180.001201
423   (1980).
424   8   Sparks, D. L. Conceptual issues related to the role of the superior colliculus in the control
425   of gaze. *Current opinion in neurobiology* **9**, 698-707, doi:10.1016/S0959-4388(99)00039-
426   2 (1999).
427   9   Gandhi, N. J. & Katnani, H. A. motor functions of the superior colliculus. *Annual Review*
428   *of Neuroscience* **34**, 205-231, doi:10.1016/j.micinf.2011.07.011.Innate (2011).
429   10   Horwitz, G. D. & Newsome, W. T. Separate Signals for Target Selection and Movement
430   Specification in the Superior Colliculus. *Science* **284**, 1158-1161,
431   doi:10.1126/science.284.5417.1158 (1999).
432   11   Cohen, M. R. & Maunsell, J. H. R. Attention improves performance primarily by reducing
433   interneuronal correlations. *Nature Neuroscience* **12**, 1594-1600 (2009).
434   12   Ruff, D. A. & Cohen, M. R. Attention Increases Spike Count Correlations between Visual
435   Cortical Areas. *Journal of Neuroscience* **36**, 7523-7534, doi:10.1523/JNEUROSCI.0610-
436   16.2016 (2016).
437   13   Ni, A. M., Ruff, D. A., Alberts, J. J., Symmonds, J. & Cohen, M. R. Learning and attention
438   reveal a general relationship between population activity and behavior. *Science* **359**,
439   463-465 (2018).
440   14   Parker, A. J. & Newsome, W. T. Sense and the single neuron: probing the physiology of
441   perception. *Annual review of neuroscience* **21**, 227-277,
442   doi:10.1146/annurev.neuro.21.1.227 (1998).
443   15   Horwitz, G. D. & Newsome, W. T. Target selection for saccadic eye movements: prelude
444   activity in the superior colliculus during a direction-discrimination task. *Journal of*
445   *neurophysiology* **86**, 2543-2558 (2001).
446   16   Krauzlis, R. J., Lovejoy, L. P. & Zénon, A. Superior Colliculus and Visual Spatial Attention.
447   *Annual Review of Neuroscience* **36**, 165-182, doi:10.1146/annurev-neuro-062012-
448   170249.Superior (2013).
449   17   Mitchell, J. F., Sundberg, K. A. & Reynolds, J. H. Spatial attention decorrelates intrinsic
450   activity fluctuations in macaque area V4. *Neuron* **63**, 879-888 (2009).

451 18   Ruff, D. A., Ni, A. M. & Cohen, M. R. Cognition as a Window into Neuronal Population
452   Space. *Annu Rev Neurosci* **41**, 77-97 (2018).
453 19   Cohen, M. R. & Kohn, A. Measuring and interpreting neuronal correlations. *Nature*
454   *Neuroscience* **14**, 811-819 (2011).
455 20   Ruff, D. A. & Cohen, M. R. A normalization model suggests that attention changes the
456   weighting of inputs between visual areas. *Proc Natl Acad Sci U S A* **114**, E4085-E4094,
457   doi:10.1073/pnas.1619857114 (2017).
458 21   Treue, S. & Martinez-Trujillo, J. C. Feature-based attention influences motion processing
459   gain in macaque visual cortex. *Nature* **399**, 575-579 (1999).
460 22   Moreno-Bote, R. *et al.* Information-limiting correlations. *Nature Neuroscience* **17**, 1410-
461   1417 (2014).
462 23   Kanitscheider, I., Coen-Cagli, R. & Pouget, A. Origin of information-limiting noise
463   correlations. *Proceedings of the National Academy of Sciences of the United States of*
464   *America* **112**, E6973-6982, doi:10.1073/pnas.1508738112 (2015).
465 24   Kohn, A., Coen-cagli, R., Kanitscheider, I. & Pouget, A. Correlations and Neuronal
466   Population Information. *Annual Review of Neuroscience* **39**, 237-256,
467   doi:10.1146/annurev-neuro-070815-013851 (2016).
468 25   Huang, C. *et al.* Circuit models of low dimensional shared variability in cortical networks.
469   *bioRxiv* **217976**, doi:https://doi.org/10.1101/217976 (2017).
470 26   Averbeck, B. B., Latham, P. E. & Pouget, A. Neural correlations, population coding and
471   computation. *Nature Reviews Neuroscience* **7**, 358-366 (2006).
472 27   Carandini, M. & Heeger, D. J. Normalization as a canonical neural computation. *Nature*
473   *Reviews Neuroscience* **13**, 51-62, doi:10.1038/nrn3136 (2012).
474 28   Brainard, D. H. The Psychophysics Toolbox. *Spatial vision* **10**, 433-436 (1997).
475 29   Pelli, D. The VideoToolbox software for visual psychophysics: Transforming numbers into
476   movies. *Spatial vision* **10**, 437-442 (1997).
477 30   Engbert, R. & Kliegl, R. Microsaccades uncover the orientation of covert attention. *Vision*
478   *Research* **43**, 1035-1045, doi:10.1016/S0042-6989(03)00084-1 (2003).
479
480

Supplemental Figure 1

495
496 *Supplemental Figure 1. Effects of attention on common analyses of individual units and pairs of*
497 *units (A) Attention increases firing rates in MT, quantified as the difference in firing rates in the*
498 *different attention conditions divided by the sum. Units with significant differences in average*
499 *responses for the two conditions are specified by black bars (t-test, p<.05). This distribution*
500 *(mean = 0.04, median = 0.04) is significantly different from zero (Wilcoxon signed rank test, p<*
501 *10^{-21}). (B) Same as A, for SC data. This distribution (mean = 0.073, median = 0.05) is*
502 *significantly different from zero (Wilcoxon signed rank test, p< 10^{-43}). (C) Within and between*
503 *area noise correlations calculated from spike counts during stimulus presentations that preceded*
504 *successful maintenance of fixation from trials that ended with either a hit or miss or were a*
505 *successful catch trial. Attention decreases average correlations within MT (Wilcoxon signed*
506 *rank test, p< 10^{-12}), not in the SC (Wilcoxon signed rank test, p=0.8) and increases them between*
507 *the two areas (Wilcoxon signed rank test, p< 10^{-41}). Error bars are standard error of the mean.*
508 *(D) Within and between area noise correlations calculated from spike counts that immediately*
509 *preceded different behavioral outcomes during cued trials. Misses and false alarms are*
510 *associated with higher correlations within MT (t-test, p<10^{-3}) and SC (t-test, p<10^{-3}) but not*
511 *between the two areas (t-test, p=0.23). Error bars are standard error of the mean.*
512

Supplemental Figure 2
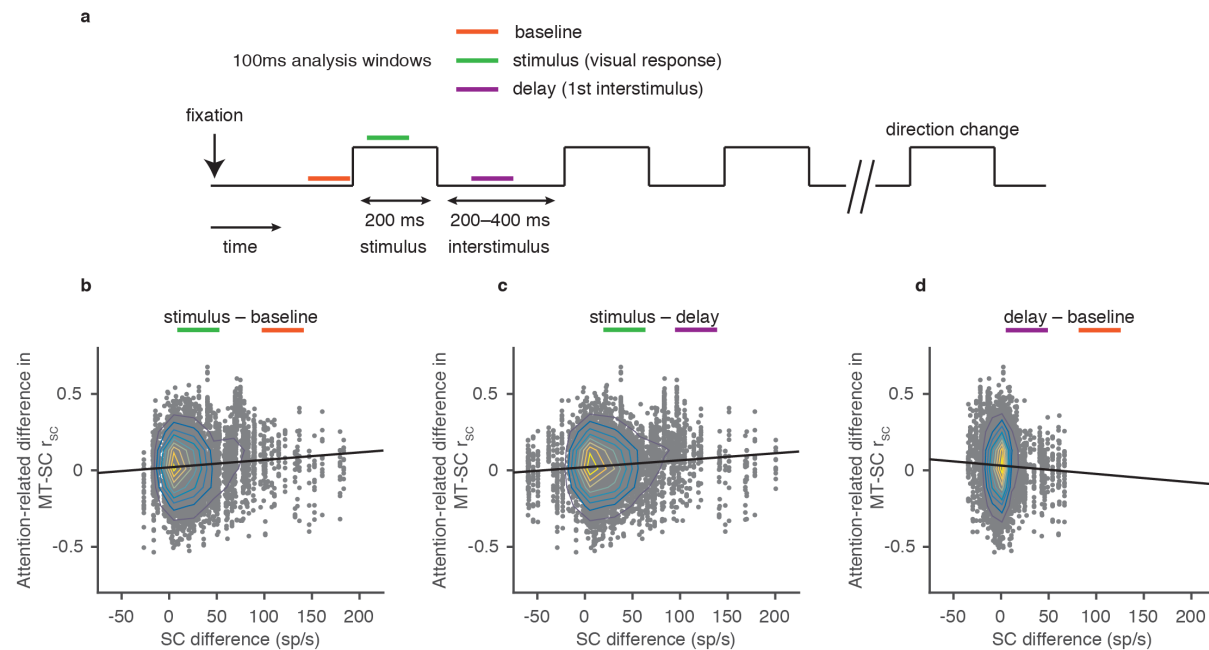
513
514 *Supplemental Figure 2. Relationship between SC responses during different task epochs and*
515 *attention-related correlation changes with MT. (A) Schematic of task timing depicts the three*
516 *100ms epochs used to count spikes in SC units. The baseline period began 100ms before the first*
517 *stimulus appeared, which is after stable fixation had been acquired. The stimulus period was*
518 *shifted 30 ms after the appearance of the visual stimulus, to account for the earliest visual*
519 *latencies observed in the SC. The delay period began 100ms after the first stimulus turned off*
520 *and always ended prior to the onset of the second stimulus. (B) Attention-related changes in MT-*
521 *SC r_{SC} plotted against the difference between each SC unit's response during the stimulus and*
522 *baseline periods. There are multiple MT-SC correlation differences measured for each SC unit.*
523 *Correlations between MT and SC were calculated using the same data and methods as*
524 *Supplementary Figure 1C (Pearson correlation, rho=0.087, p< 10^{-12}). Isolines depicting the*
525 *decile boundaries are overlaid over the individual data points. (C) Similar to B, but data are*
526 *now sorted by the difference between each SC unit's response during the stimulus and delay*
527 *periods (Pearson correlation, rho=0.092, p< 10^{-14}). (D) Similar to B, but data are now sorted by*
528 *the difference between each SC unit's response during the delay and baseline periods (Pearson*
529 *correlation, rho=-0.042, p< 10^{-4}).*
530