1 Title:

2 Surprise Signals in the Supplementary Eye Field: Rectified Prediction Errors Drive

3 Exploitation-Exploitation Transitions

4

5 Authors:

6 Norihiko Kawaguchi,[1,2] Kazuhiro Sakamoto,[3] Naohiro Saito,[1] Yoshito Furusawa,[1] Jun Tanji,[1]

7 Masashi Aoki,[2] and Hajime Mushiake[1,4]

8

9 Affiliations:

10 [1]Department of Physiology, Tohoku University School of Medicine, 2-1 Seiryo-machi, Aoba-ku,

11 Sendai, 980-8575, Japan

12 [2]Department of Neurology, Tohoku University School of Medicine, 1-1 Seiryo-machi, Aoba-ku,

13 Sendai, 980-8574, Japan

14 [3]Research Institute of Electrical Communication, Tohoku University, 2-1-1 Katahira, Aoba-ku,

15 Sendai, 980-8577, Japan

16 [4]Core Research for Evolutional Science and Technology (CREST), JST, Tokyo

17

18 Running Head: SEF Surprise Signals Drive Behavioral Transitions

19

20 Corresponding author: Hajime Mushiake

21 Department of Physiology, Tohoku University School of Medicine,

22 2-1 Seiryo-machi, Aoba-ku, Sendai, 980-8575, Japan

23 e-mail: hmushiak@med.tohoku.ac.jp, phone: +81-22-717-8073

24

25 Number of pages: 51

26 Number of figures: 9

27 Number of tables: 2

28 Numbers of words: abstract 230, Introduction 468, Discussion 2545.

29

30 Conflicts of interest: The authors declare no competing financial interests.

31

Abstract

Visual search is coordinated adaptively by monitoring and predicting the environment. The

supplementary eye field (SEF) plays a role in oculomotor control and outcome evaluation. However,

it is not clear whether the SEF is involved in adjusting behavioral modes based on the preceding

feedback. We hypothesized that the SEF drives exploration–exploitation transitions by generating

"surprise signals" or rectified prediction errors, which reflect differences between predicted and

actual outcomes. To test this hypothesis, we introduced an oculomotor two-target search task in

which monkeys were required to find two valid targets among four identical stimuli. After they

detected the valid targets, they exploited their knowledge of target locations to obtain a reward by

choosing the two valid targets alternately. Behavioral analysis revealed two distinct types of

oculomotor search patterns: exploration and exploitation. We found that two types of SEF neurons

represented the surprise signals. The error-surprise neurons showed enhanced activity when the

monkey received the first error feedback after the target-pair change, and this activity was followed

by an exploratory oculomotor search pattern. The correct-surprise neurons showed enhanced

activity when the monkey received the first correct feedback after an error trial, and this increased

activity was followed by an exploitative, fixed-type search pattern. Our findings suggest that

error-surprise neurons are involved in the transition from exploitation to exploration, and that

correct-surprise neurons are involved in the transition from exploration to exploitation.

Introduction

Dynamic control of exploration and exploitation is essential for searching optimally in an uncertain environment (Gittins 1979; Krebs et al. 1978). During a search for a new object, we explore the visual field based on predictions about the location of the object and the behavioral outcomes. After we find the object, we memorize the location and exploit the knowledge in the next search. This process involves the use of eye movements to search for an object. Thus, the higher-order oculomotor center is expected to be involved in adaptive searches using eye movements.

The supplementary eye field (SEF) plays a crucial role in the control of higher-order eye movements. The SEF, first identified as an oculomotor area in the dorsomedial frontal cortex (Schlag and Schlag-Rey 1987), is involved in the control of anti-saccades (Schlag-Rey et al. 1997) and the temporal organization of multiple saccades (Berdyyeva and Olson 2010; Isoda and Tanji 2003; Lu et al. 2002). The SEF has recently been found to exert non-motor supervisory functions in oculomotor control (Schall et al. 2002). Specifically, distinct groups of SEF neurons were activated after errors, after the successful withholding of a prepared eye movement, or in association with reinforcement in an oculomotor countermanding task (Stuphorn and Schall 2002; Stuphorn et al. 2000). In addition, the SEF is related to prediction and detection of reward (Amador et al. 2000) and to the evaluation of the amount of reward (Roesch and Olson 2003; Uchida et al. 2007). Furthermore, the SEF is suggested to have executive control over saccade generation by monitoring behavioral consequences (Schall and Boucher 2007; Stuphorn and Schall 2006). These studies show that the SEF is involved in both the motor and cognitive aspects of eye

71   movement control. However, it remains unclear how the SEF contributes to adaptive oculomotor

72   search based on performance monitoring. Thus, we attempted to investigate how the error- and

73   correct-outcome signals represented in the SEF are used to adjust subsequent searches. We

74   hypothesized that the SEF controls transitions between exploration and exploitation by producing

75   two types of surprise signals in response to unexpected erroneous and unexpected correct

76   outcomes (Fig. 1). To test this hypothesis, we introduced an oculomotor two-target search task in

77   which monkeys were required to explore valid targets with eye movements and to exploit

78   memorized target locations to select the next valid target. To classify SEF neurons, we constructed a

79   reinforcement learning model in which the predictions of correct and error feedback were updated

80   via surprise signals, defined as the differences between prediction and outcomes. We found that two

81   types of surprise neurons responded to unexpected correct and error feedback, respectively. Their

82   post-feedback activities were related to the shift between oculomotor search patterns in the

83   following trials, implying that the SEF plays a supervisory role in the exploration–exploitation

84   transitions based on performance monitoring.

85 **Methods**

86 *Subjects and apparatus*

87 Two Japanese monkeys (*Macaca fuscata*: monkey K, 6.5 kg; and monkey M, 6.1 kg) were trained to

88 perform the task in this study. They were cared for in accordance with the National Institutes of

89 Health guidelines and the guidelines of Tohoku University. The monkeys were kept in individual

90 primate cages in an air-conditioned room with food available *ad libitum*. During the experiments, the

91 monkeys sat in a primate chair with their heads restrained and faced a screen on which visual stimuli

92 were presented. Eye position was monitored with an infrared corneal reflection system sampling at

93 250 Hz.

94

95 *Behavioral task*

96 We introduced an oculomotor search task in which the monkeys were required to find two valid

97 targets among four identical targets and to make saccades alternately to the two targets (Fig. 2*A*).

98 Because the positions of the valid targets were not shown explicitly, the monkeys had to search for

99 the targets by trial and error. After a predetermined number of correct trials, the valid-target pair was

100 changed without instruction; thus, the monkeys had to change their search strategy (i.e., exploration

101 and exploitation) after receiving error feedback. The events in a trial are shown in Figure 2*A*.

102 Monkeys were allowed to choose one target in each trial. First, the monkeys were required to fixate

103 on a central fixation point for 500–700 ms. The onset of fixation was signaled when the color of the

104 central fixation point changed from white to red. When the monkeys fixated on the point for 500–700

105 ms successfully, the color was changed from red to green; this was followed by the appearance of

106 four identical visual targets: 2 × 2° white points presented 8.5° from the central fixation point. If the

107 monkeys maintained fixation on the central fixation point for an additional 500–700 ms, the fixation

108 point disappeared (Go signal), and the monkeys were allowed to make a saccade to one of the four

109 peripheral targets. When the monkeys fixated on a target, the color of the selected target turned

110 from white to red, signaling the start of fixation. The monkeys had to maintain fixation for 700 ms to

111 receive feedback, which indicated whether the selected target was valid. There were two types of

112 visual feedback. When the monkey fixated on an invalid target, the four dots disappeared (error

113 feedback). When the monkeys fixated on the valid target, the color of the selected stimulus turned

114 from red to green (correct feedback), and the four stimuli turned off 500 ms after the color change.

115 The reward was then provided 500 ms after the disappearance of the four stimuli. After an interval of

116 2,000–2,500 ms, the next trial began. After the monkeys selected the valid target, the currently valid

117 target was shifted to the other one within the currently valid pair (Fig. 2*A*, see the changes between

118 "Bottom-right" and "Bottom-left"). If the fixation was interrupted before the feedback, the four

119 peripheral targets were turned off, and the next trial began without reward (fixation-break trial). The

120 fixation-break trials were excluded from our analyses.

121   We defined four types of valid pairs, each of which consisted of two adjacent targets among the

122 four presented (Fig. 2*B*). The valid pair was changed in a pseudorandom manner after a

123 predetermined number of consecutive correct choices (6–10 trials) or total number of correct

124 choices (10–20 correct trials) after each pair change. Because the current valid pair was shifted

125 without instruction, the monkeys had to search the valid paired targets based on behavioral

126   outcomes.

127   We classified the oculomotor search pattern into three types: "Adjacent," "Diagonal," and

128   "Repeat" (Fig. 2*C*). These patterns were defined by the relative location of the target chosen in the

129   current trial (n) compared with that of the target chosen in the previous trial (n–1). For example, if the

130   subject selected a target that was adjacent to the previously selected one, the search was classified

131   as an "Adjacent" search pattern. The adjacent search patterns were further divided into two types

132   based on whether the target selected in the current trial was included in the current valid pair or was

133   not. "Diagonal" patterns were defined as searches in which the subject selected the target diagonal

134   to the previously selected one. "Repeat" patterns were defined as searches in which the same target

135   selected in the previous trial was selected again.

136   We distinguished two phases of search behavior: exploration and exploitation. We defined the

137   exploration phase as the trials from the first error trial after a valid-pair change to the first correct trial

138   in which the monkey found the first valid target. The exploitation phase was defined as the

139   consecutive correct trials after the trial in which the monkey found both targets of the current valid

140   pair. If the subject made an error during the consecutive correct trials, the trials after the error were

141   not categorized into the exploitation phase. Based on these definitions, we inferred the internal

142   behavioral mode of the monkeys from their oculomotor search patterns and determined the timing of

143   the behavioral transitions. Examples of behaviors and feedback are presented in Figure 2*D*. After

144   the valid targets were changed without instruction, the monkeys made errors and began to search

145   for new valid targets.

146

147   *Computational model*

148   *General characteristics.* To categorize the task-related neurons, we developed a model for an

149   evaluation system to monitor behavioral outcomes, predict forthcoming feedback, and generate

150   "surprise signals". The following two general assumptions underpinned this model.

151   (1) We assumed that the evaluation system detects changes in the environment by calculating the

152   differences between predictions and actual outcomes. The product of such comparisons were

153   so-called "prediction errors," which can be represented as either signed or unsigned values. A

154   signed prediction error was calculated by subtracting the prediction value from the actual outcome.

155   Thus, a prediction error was either positive or negative, depending on the two values. Dopamine

156   neurons are well-known examples of signed prediction errors (Schultz 1997; 2006). In our analysis,

157   signed prediction errors were rectified, and only the positive components were used as surprise

158   signals that encoded the occurrence of an unexpected event or the non-occurrence of expected one

159   (Alexander and Brown 2011). Specifically, when an event occurred unexpectedly, the surprise signal

160   was calculated as a rectified prediction error by subtracting the prediction value from the actual

161   outcome value. Based on Alexander's original model, we hypothesized two types of rectified

162   prediction error signals: correct and error. The correct type responded to correct feedback with an

163   increasing firing rate, and the error type responded to error feedback with an increasing firing rate.

164   Although the two types of rectified prediction errors are two aspects of the signed prediction error,

165   this issue of whether the surprise signals are encoded in the brain as single or multiple prediction

166 errors is not trivial, because the rectified prediction errors could lead to different behaviors.

167   (2) We assumed that the evaluation system would be able to predict future events such as the

168 occurrences of correct and error feedback. Furthermore, the two signals were followed by different

169 behavioral adjustments as shown in the results of the behavioral analyses. It is plausible that two

170 separate representations of correct and error outcomes enable flexible behavioral adjustments by

171 facilitating the prompt switch to the next search.

172   Based on these assumptions, we defined two types of prediction errors: surprise signals for correct

173 and error feedback. The mathematical details are presented in the Appendix.

174

175 *Surgical and recording methods*

176 After the monkeys were trained to select targets at a rate of 90% correct during the exploitation

177 phase, acrylic recording chambers (40 × 30 mm) were implanted in their skulls under aseptic

178 conditions. Using Horsley–Clarke's stereotactic coordinates, markers indicating reference points

179 were placed on the skull. Anesthesia was induced using ketamine hydrochloride (8 mg/kg i.m.) plus

180 atropine sulfate, followed by pentobarbital sodium (30 mg/kg i.m.). Antibiotics and analgesics were

181 used to prevent postsurgical infection and pain, respectively.

182   Standard electrophysiological techniques for single-cell recording were used to record the activity

183 from the SEF (Mushiake et al. 2006). We located the SEF using previously established criteria to

184 differentiate the SEF from adjacent areas, such as the supplementary motor area, based on

185 responses to intracortical microstimulation (22–42 cathodal pulses of 0.2-ms duration at 333 Hz)

186 (Isoda and Tanji 2003; Wang et al. 2005). When eye movements were elicited by the stimuli at

187 currents of <50 µA in response to more than half the stimuli, the neuron was judged to be in the SEF.

188 The cells were also assessed by location (2–8 mm rostral to the genu of the arcuate sulcus (AS

189 genu) and within a few millimeters of the interhemispheric cleft) (Moorman and Olson 2007).

190 Neuronal activity was recorded using glass-insulated Elgiloy electrodes (impedance of 0.8–1.0 MΩ)

191 and parylene-insulated tungsten microelectrodes (impedance of 1.0–1.8 MΩ). An electrode was

192 inserted through the dura using a hydraulic microdrive (Narishige, Tokyo, Japan). Neuronal activity

193 was amplified and band-pass filtered (1–6 kHz), and single-unit potentials were discriminated using

194 a template-matching algorithm (MCP-Plus and MSD; Alpha Omega Engineering, Nazareth, Israel).

195 Data were stored in a microcomputer for off-line analysis.

196

197 *Analysis of single-unit activity*

198 *Task-related neurons.* If neuronal activity changed during the feedback or saccade period, the

199 neurons were judged to be task-related. Specifically, if neuronal activity during a pre-feedback

200 (0–300 ms before feedback) or post-feedback (0–300 ms after feedback) period differed significantly

201 from that during the control period (200–500 ms after onset of fixation to the central fixation point,

202 Mann–Whitney, $p < 0.05$), the neuron was judged to be feedback-related (Fujii et al. 2002). If the

203 activity during a saccade period (0–300 ms after the Go signal) differed significantly from that during

204 the control period (Mann–Whitney, $p < 0.05$), the neuron was judged to be "eye movement-related".

205

206 *Regression analysis.* We hypothesized that SEF neurons would play evaluative roles in the

207 oculomotor search task. Therefore, we tested whether the neurons represented the values in the

208 evaluation system described above. To classify the neurons, we performed regression analyses

209 containing a variable representing each value in the aforementioned model. We determined the

210 value that fits best for each neuron by comparing the coefficients of determinants. We performed

211 regression analyses on a trial-by-trial basis as

212 $$Firing\ rate = \alpha_i Value_i + \varepsilon_i\ (i = 1..10),$$

213 where $Value_i$ denotes the explanatory variables that were the values described above (O_c, V_c, SS_c,

214 *signed* PE_c, O_e, V_e, SS_e, *signed* PE_e) and by the feedback in the current and previous trials. α_i are

215 regression coefficients, and ε_i are intercepts for the explanatory variables. We calculated the

216 coefficient of determination (R^2) for each model for each neuron. We defined a neuron as

217 feedback-related when the R^2 was higher than 0.1. Based on this analysis, we selected the model

218 with the higher R^2 as the best one to characterize neuronal activity.

219   To eliminate the simple visual-stimuli-responsive neurons from the feedback-related neurons, we

220 compared the neuronal activity following the visual stimuli that signaled a correct or an error

221 outcome from those that signaled only task phases. If the neuron did not show a significant change

222 in activity after the outcome signals ($p < 0.05$), we excluded that neuron from further analyses.

223 Specifically, to eliminate a neuron that responded to the change from red to green, we compared the

224 activities during the periods after fixation onset with those after correct feedback (both stimuli

225 consisted of a change in the color of the fixated target from red to green). We excluded the neuron if

226   it did not exhibit significantly different activities during the two periods. Additionally, to exclude

227   neurons related to the absence of stimuli and those related to eye movements, we compared the

228   activities during the periods in the error and correct trials in which stimuli were disappeared. We

229   excluded the neuron if it did not show a significant difference between the activities in error and

230   correct trials.

231

232   *Directional selectivity.* To examine whether task-related neurons were modulated by saccade

233   direction, we performed regression analyses using the following equation:

234                              $Firing\ rate = \alpha_1 H + \alpha_2 V + \varepsilon,$

235   where *H* denotes the horizontal location of the selected target, left (0) or right (1), *V* denotes the

236   vertical location, bottom (0) or top (1). α_1 and α_2 are coefficients, and ε is an intercept. If the effect of

237   either *H* or *V* were significant ($p < 0.05$) during the saccade period (0–300 ms after Go signal), the

238   neuron was classified as having directional selectivity.

239

240   *Normalized firing rate.* The firing rate for each neuron was normalized in a 20-ms window by

241   constructing a *z*-score of activity for each neuron in each bin. The *z*-score was computed across all

242   correct and error trials by normalizing by the standard deviation in the mean activity of each neuron

243   (Buschman and Miller 2009).

244

245

246 *Statistical analyses of the neuronal activity and behavior:* To investigate the activity pattern of the

247 surprise neurons, we examined the changes in the activity during the post-feedback period during

248 trials that were crucial to oculomotor target selection. Specifically, we calculated the firing rates for

249 the following four types of trials: the first error trial after the change in the valid-target pair (1st E), the

250 first correct trial after the pair change (1st C), the second correct trial after the pair change (2nd C),

251 and the correct trials after more than two consecutive correct trials (CC). We calculated the

252 averaged firing rates of the surprise neurons in these index trials. We used ANOVA with Tukey's

253 HSD correction (factor: trial types, $p < 0.05$) to assess the differences in neuronal activity between

254 the activities in these trial types. For other statistical tests, we used paired or simple *t*-tests for

255 comparisons of two values. To investigate the relationship between the activity and oculomotor

256 search patterns, we categorized trials by the oculomotor search pattern in the following trials

257 (diagonal or adjacent) and calculated the mean post-feedback activity in the trials. In this analysis,

258 the exploration phase was analyzed for the error-surprise and error-outcome neurons, and the

259 exploitation phase was analyzed for the correct-surprise and correct-outcome neurons.

260

261 *Choice probability:* To determine the extent to which the activity of a neuron predicted the

262 oculomotor search pattern (adjacent or diagonal), we calculated the choice probabilities of

263 task-related neurons (Britten et al. 1996). In this situation, the choice probability was an estimate of

264 the probability that an observer correctly predicted whether the monkey would engage in the

265 adjacent or diagonal search pattern in the subsequent trial. To compute the choice probability, a

266   receiver operating characteristic curve was computed from the distributions of neuronal activities for

267   each of the two search patterns of the animal. The area under the curve was integrated to produce

268   the choice probability (Yang and Heinen 2014).

269

270

271 **Results**

272 *Behavioral results*

273 After training, the accuracy of the two monkeys reached ~90% in the exploitation phase (monkey K:

274 $91 \pm 3\%$, monkey M: $88 \pm 4\%$, mean $\pm$ SD). After a valid-pair change, they began to explore new

275 targets and found two targets in a new pair within an average of 4.2 trials (K: 4.2 ± 1.2, M: 4.3 ± 1.3).

276 We first examined the monkeys' oculomotor search patterns in the exploration and exploitation

277 phases in terms of the three search patterns described in the Methods section: diagonal, adjacent,

278 and repeat. During the exploitation phase, the adjacent search pattern was dominant (K: $97 \pm 2\%$,

279 M: $96 \pm 2\%$), and the diagonal search was observed only rarely (K: $1 \pm 1\%$, M: $1 \pm 1\%$) (Fig. 4*A*). In

280 contrast, during the exploration phase, the probability of the diagonal search increased (K: $54 \pm 5\%$,

281 M: $48 \pm 5\%$), and the adjacent search decreased (K: $39 \pm 5\%$, M: $47 \pm 4\%$). This increase in the

282 probability of the diagonal search during the exploration phase compared with the exploitation

283 phase was statistically significant (paired *t*-test; $p < 0.001$). In the exploration phase, 4.2 trials were

284 required to find the current pair of targets after a pair change. However, this raised the question of

285 whether the 4.2 trials were optimal. To answer this, we performed the following analyses.

286 To investigate whether the subsequent oculomotor search pattern was dependent on behavioral

287 history, we analyzed the search patterns in the next trials (n+1) based on the outcomes in the current

288 (n) and previous (n–1) trials. The results showed that the frequency of diagonal search increased

289 after the monkey made errors, especially when the error occurred after correct trials (Fig. 4B, error

290 →correct). However, adjacent searches were dominant when correct feedback occurred in the

291 current trial (n), irrespective of the outcome in the previous trial (n-1). We observed that the monkeys

292 choose the diagonal target in the specific situation in which the subject made an error after selecting

293 the target that was adjacent to the previous correct target. This pattern was included in the correct

294 →error in Figure 4B.

295   Next, we analyzed the dependency of the monkeys' behavior on the history of the outcomes of the

296 two preceding trials. Specifically, we compared the probability of a diagonal search occurring

297 between the trials in which the two preceding trials (n-2) were correct and erroneous, respectively.

298 The outcome of the two-preceding trials (n-2) did not have a significant effect on the search patterns

299 in the next trials (n+1) (paired t-test, $p > 0.001$). Importantly, this suggests that the monkeys chose

300 the search pattern based on the outcomes of the current (n) and previous (n-1) trials and did not use

301 the information provided before the previous trial. This result implied that the monkeys showed

302 suboptimal behavior during our task. Based on the finding that the monkeys used a combination of

303 information from the current and previous trials, we categorized the neurons by the outcomes of the

304 current and previous trials.

305   In addition, to examine whether the search pattern was affected by the direction of the chosen

306 target, we analyzed the relationship between the position of the target selected in the current trial

307 and the search patterns in the subsequent trials. The results revealed no significant relationship

308 between search patterns and saccade directions (ANOVA, factor: direction of stimuli, $p = 0.17$).

309 These results suggest that distinct search strategies operated during the exploration and

310 exploitation phases.

311

Neuronal database

313 We recorded 295 neurons in two monkeys (monkey K: 268, monkey M: 27). Of these SEF neurons,

314 200 cells (K:184, M: 16) were found to be feedback-related, and 212 cells (K: 197, M: 15) were

315 eye-movement-related. Among the 212 eye-movement-related neurons, 105 (K: 100, M: 5) showed

316 significant directional selectivity (p < 0.05) during the saccade period. We present the various types

317 of neurons in Table 1.

318   In the following section, we initially discuss outcome neurons that responded to correct or error

319 feedback. We next address the surprise neurons that showed enhanced activity when the feedback

320 was unexpected. Finally, we investigate the relationship between neuronal activities and behavioral

321 patterns and present data showing that the enhanced activity of the surprise neurons was related to

322 shifts between oculomotor search patterns.

323

Outcome neurons

325 We identified SEF neurons showing elevated activity in response to error or correct feedback. The

326 error-outcome neurons responded to error feedback in the current (n) trial, irrespectively of the

327 outcomes of previous trials (Fig. 5*A*, *B*). The error-outcome neurons were categorized in terms of

328 encoding the outcome values of error feedback (OV_e). We identified 47 error-outcome neurons. On

329 the other hand, the correct-outcome neurons were divided into two types. The first type of neuron

330 showed elevated activity in response to correct feedback, irrespectively of the outcomes of previous

331 trials (Fig. 5*C*). The second type of neuron showed distinctly decreasing activity in response to error

332 feedback (Fig. 5*D*). We categorized 29 cells in the correct-outcome neurons, 21 of which were

333 categorized as the first type of neuron, and the remaining eight of which were categorized as the

334 second type. These correct-outcome neurons were viewed as encoding the outcome values of

335 correct feedback (OV_c). Population-level activity was observed to increase slightly in the correct

336 trials and to decrease slightly in the error trials (Fig. 5*E*). These neurons were not classified as

337 prediction-error neurons but as outcome neurons, because they responded differentially depending

338 on outcome history.

339

340 ***Surprise neurons***

341 The surprise neurons showed enhanced activity when the unpredicted-outcome signals were fed

342 back. There were two types of surprise neurons. First, error-surprise neurons increased in activity in

343 response to the error feedback after a correct trial. Second, the correct-surprise neurons increased

344 in activity in response to the correct feedback after an error trial. Afterwards, we investigated how

345 these neurons were related to the behavioral transitions between exploration and exploitation.

346   The first example was an error-surprise neuron, which showed elevated activity in response to the

347 error feedback in the current (n) trial when correct feedback was received in the previous (n-1) trial

348 (Fig. 6*A,* correct→error). However, the activity was moderate when the error feedback occurred

349 repeatedly and was predicted (Fig. 6*A,* error→error). Because the outcome of the current trial

350 differed from that of the previous trial, the current-error outcome was surprising, and we refer to

351 these neurons as "error-surprise neurons". These neurons were viewed as encoding the surprise

352 signal as error feedback (SS_e). A total of 15 SEF neurons were placed in this group. Population-level

353 analysis showed the highest activity when the error feedback was received in the current (n) trial

354 after correct feedback had been received in the previous (n-1) trial (Fig. 6*B,* correct→error). There

355 was no response when the correct feedback was received in the current (n) trial (correct→correct

356 and error→correct). The second example was a correct-surprise neuron, which exhibited elevated

357 activity when correct feedback was given in the current (n) trial after error feedback was received in

358 the previous (n-1) trial (Fig. 6*C,* error→correct). The post-feedback activity decreased when correct

359 trials occurred successively and were predicted (correct→correct). Because the outcome of the

360 current trial differed from that of the previous trial, the current correct outcome was surprising, and

361 we refer to these neurons as "correct-surprise neurons." These neurons were viewed as encoding

362 the surprise signal for correct feedback (SS_c). A total of 14 neurons were placed in this group.

363 Population-level analysis showed similar activity when correct feedback was received in the current

364 (n) trial following error trials (Fig. 6*D,* error→correct). Notably, these error- and correct-surprise

365 neurons differed from neurons reflecting conventional prediction errors (*signed PE_e* and *signed PE_c*),

366 because the surprise neurons encoded *"rectified"* positive parts of the prediction errors.

367   To confirm the validity of the categorization of the surprise neurons, we compared the R^2 values of

368 the surprise neurons in terms of the different values in the model. Of the 15 error-surprise neurons,

369 11 had R^2 values higher than 0.1 for signed prediction errors (*signed PE_e*), and 10 had R^2 values

370 higher than 0.1 for error outcomes (OV_e) (Table 2). However, the R^2 values for error-surprise signals

371 were significantly higher than those of the other two groups (paired *t*-test, SS_e vs. *signed PE_e*, p =

372 0.009; SS_e vs. OV_e, p = 0.002). On the other hand, among the 14 correct-surprise neurons, nine had

373 R^2 values higher than 0.1 for signed prediction errors (*signed PE_c*), and seven had R^2 values higher

374 than 0.1 for correct outcomes (OV_c) (Table 2). However, the R^2 values for correct-surprise signals

375 were significantly higher than those in other groups (paired *t*-test, SS_c vs. *signed PE_c*, p = 0.001; SS_c

376 vs. OV_c, p = 0.0007).

377   Finally, to investigate whether the surprise neurons were related to specific target directions, we

378 used regression analyses to examine the directional selectivity of the neurons during saccade

379 periods. We found that 10 error-surprise neurons (10/15) and seven correct-surprise neurons (7/14)

380 showed directional selectivity during the 0–300-ms period after the feedback signals (ANOVA;

381 factor: directions of stimuli; p < 0.01). On the other hand, 38 error-outcome neurons (38/47) and 15

382 correct-outcome neurons (15/29) showed directional selectivity. This result is consistent with a

383 previous finding that half the SEF neurons showed directional selectivity (Roesch and Olson 2003).

384

385 ***Surprise neurons and history of behavioral outcomes***

386 We investigated the relationship between the neuronal activity of surprise neurons and the outcome

387 history. We identified two types of surprise neurons in the SEF that exhibited enhanced activity

388 when an unpredicted outcome occurred. First, we investigated how the surprise neurons reflected

389 changes in outcomes after pair shifts. Specifically, we analyzed the averaged firing rates of the

390 surprise neurons in the following trials: the first error trial after a pair change (1st E), the first correct

391 trial after a pair change (1st C), the second correct trial after a pair change (2nd C), and consecutive

392 correct trials (CC). These trials were crucial because they provided clues about the valid targets by

393 tracking the history of behavioral outcomes. As a result, the error-surprise neurons showed

394 significantly enhanced responses in the first error trial after the pair shifts compared with those in the

395 consecutive correct trials (Fig. 7A: 1st E vs. CC, p = 0.012). The activity in the first error trial was

396 significantly higher compared with those in the first and second correct trials (Fig. 7A: 1st E vs. 1st C,

397 p = 0.029; 1st E vs. 2nd C, p = 0.025). We also found that the neuronal activity level of the

398 error-surprise neurons was higher in the first error trial than in the second error trial (1st vs. 2nd, 7.0 $\pm$

399 1.2 vs. 4.8 $\pm$ 0.90, p = 0.013). We did not find a significant change after the third try due to the low

400 number of trials under this condition (< 3 trials per neuron). The correct-surprise neurons showed

401 significantly higher activity in the first correct trials compared with the preceding first error trial (Fig.

402 7B, 1st E vs. 1st C, p = 0.040). Furthermore, the activity in the first correct trial of correct-surprise

403 neurons was larger in the first try than in the second try after a pair change, but this difference was

404 not statistically significant (1st vs. 2nd try, 6.0 $\pm$ 0.8 vs. 4.4 $\pm$ 0.5, p = 0.16). We did not find a

405 significant change in the third try, because of the scarcity of trials performed under this condition (< 3

406 trials per neuron). These findings suggest that surprise neurons play a role in detecting the

407 unexpected trials after the valid pair changes. This result is consistent with a previous study that

408 found that neurons in the frontomedial region showed enhanced activity during the first correct trials

409 after a change in the task condition (Quilodran et al. 2008).

410

411 ***Surprise neurons and exploration–exploitation transitions***

412   To investigate whether the surprise neurons in the SEF were related to the exploitative or

413 exploratory oculomotor search pattern, we analyzed the relationship between the post-feedback

414 activities and search patterns in the subsequent trials. According to our behavioral analysis, there

415 were two main oculomotor search patterns: diagonal and adjacent. The use of the diagonal pattern

416 increased in the exploration phase, but the adjacent pattern was dominant in the exploitation phase

417 (Fig. 4*A*). Thus, we investigated the relationship between the post-feedback activities and the

418 oculomotor search patterns in the following trials. Specifically, we compared the neuronal activities

419 in the trial followed by a diagonal search with those in the trials followed by an adjacent search. The

420 results showed that error-surprise neurons had significantly higher activity levels in trials followed by

421 a diagonal search than those followed by an adjacent search (Fig. 7*C*; p = 0.040). Furthermore,

422 correct-surprise neurons showed significantly higher activity levels in trials followed by an adjacent

423 search than in those followed by a diagonal search (Fig. 7*D*; p = 0.010). In contrast, the activity

424 levels of error- and correct-outcome neurons were not significantly correlated with their activities or

425 search patterns in subsequent trials (Fig. 7*E* and *F*; error-outcome neuron: p > 0.05,

426 correct-outcome neuron: p > 0.05). These results suggest that the increased activity of the surprise

427 neurons are related to the shifts between the oculomotor search patterns, but that the activity of the

428 simple outcome neurons is not.

429   To investigate the influence of post-feedback activity on subsequent behavioral adjustments

430 further, we calculated the choice probability for surprise and outcome neurons (Fig. 7*G*). In this

431 analysis, we computed the extent to which we could predict whether the subsequent search pattern

432 was diagonal or adjacent. According to the results, the choice probability of the error-surprise

433 neurons was higher than that of the error-outcome neurons (p = 0.0028). The choice probability of

434 the correct-surprise neurons was slightly higher than that of the correct-outcome neurons, but this

435 difference was not significant (p = 0.33).

436   In summary, the error-surprise neurons exhibited increased responses in the first error trials, and

437 this activity was related to the behavioral shift from exploitation to exploration in the following trials.

438 The correct-surprise neurons showed significant responses in the first correct trials, and this activity

439 was related to the shift from exploration to exploitation in the following trials. We plotted the

440 distribution of these SEF neurons and found no clear segregation (Fig. 8*B* and *C*).

441

Discussion

442 We introduced a two-target search task in which monkeys had to find two adjacent targets by trial

443

444 and error. Behavioral analysis revealed two distinct oculomotor search patterns during the

445 exploration and exploitation phases and revealed that the monkeys used their memory from only the

446 two preceding trials. We found two types of surprise neurons: error-correct surprise neurons and

447 correct-surprise neurons. These neurons showed elevated responses to unexpected error and

448 correct feedback, respectively. Specifically, the error-surprise neurons selectively showed

449 increasing activity when the monkeys made the first error after a pair change. The correct-surprise

450 neurons showed enhanced activity when the monkeys found the first correct target after a pair

451 change. Furthermore, the surprise neurons were related to the exploration–exploitation transitions.

452 Specifically, the error-surprise neurons showed enhanced activity preceding the transition from

453 exploitative to explorative search pattern. In contrast, the correct-surprise neurons showed

454 enhanced activity preceding the transition from explorative to exploitative search pattern. Our

455 results suggest that the SEF is involved in behavioral transitions between exploration and

456 exploitation by encoding the surprise signals (Fig. 1).

457

458 ***Behavioral shifts during the two-target search task***

459 The unique feature of our search task was the absence of an external cue for valid targets and the

460 attendant requirement that monkeys search for the targets by trial and error. Previous studies used

461 visual search tasks in which monkeys had to find a target among multiple identical stimuli (Amiez et

462 al. 2012; Mirpour and Bisley 2012; Procyk et al. 2000; Quilodran et al. 2008). In contrast, our

463   oculomotor search task required monkeys to make alternate saccades to two valid targets during

464   the exploration phase to motivate them to change the search pattern after a pair change. Thus, the

465   subjects had to maintain their memory of previous choices and feedback to choose subsequent

466   behavioral patterns.

467   Interestingly, the behavioral analyses revealed that the monkeys remembered the outcomes of

468   only the two preceding trials: the current (n) and the previous trial (n-1). In a previous study,

469   Quilodran et al. (2008) demonstrated that monkeys could remember the outcomes of three

470   preceding trials according to their ability to choose the correct target among four possibilities after

471   three consecutive error trials (Quilodran et al. 2008). This discrepancy between Quilodran's and our

472   findings may be explained not by a limit in the monkeys' working memory but rather by their search

473   strategy. In the study by Quilodran et al., the monkeys had to find one correct target among four

474   targets, but our study required the monkeys to select two valid targets alternately. Our task was

475   more demanding, and the monkeys used the two preceding outcomes to choose the subsequent

476   search patterns.

477

478   ***Surprise signals in the SEF***

479   Our results revealed surprise neurons in the SEF. These were defined as neurons that encoded the

480   unsigned, rectified values of the prediction error, the difference between the prediction and the

481   outcome. These surprise neurons engaged in activity that was distinct from that displayed by

482   outcome neurons and reward-prediction-error neurons. Their pattern of activity was similar to that of

483 the surprise neurons in the ACC and prefrontal cortex (PFC) in that they responded to unpredicted

484 outcomes (Asaad and Eskandar 2011; Hayden et al. 2011). Learning-selective activity has been

485 observed in the SEF during an associative learning task, and this type of activity appeared

486 preferentially when a novel stimulus was introduced (Chen and Wise 1995). One recent study

487 revealed a group of neurons that encoded valence-related reward-prediction errors (So and

488 Stuphorn 2012). In this context, we possibly demonstrated that SEF neurons encode a different

489 aspect of the prediction error and that distinct types of surprise neurons represent the rectified

490 prediction errors for correct and error feedback, respectively. Our finding supports the notion that the

491 surprise signal promotes adaptive learning in a changing world (Courville et al. 2006; Pearce and

492 Hall 1980).

493   The SEF encodes two types of prediction errors: one is based on the amount of reward and the

494 other on the subjective probabilities of feedback. The former type of prediction error was identified

495 using a behavioral task that involved changing the amount of the reward systematically (So and

496 Stuphorn 2012). On the other hand, we investigated the latter type of prediction error as a surprise

497 signal. We computed surprise signals based on the subjective probability of feedback, because we

498 assumed that the SEF was able to evaluate outcomes based on the subjects' behavioral history.

499 This is consistent with a recent study in which SEF neurons showed metacognition-related activity

500 that reflected the subjective probability of correct choices (Middlebrooks and Sommer 2012).

501   We found that a portion of the surprise neurons showed directional selectivity during saccade

502 periods, implying that they integrated directional information and outcome signals. We assumed that

503 the directional information in the saccade period and the outcome information during the

504 post-feedback period would be combined to adjust subsequent behavior. On the other hand, a

505 previous report showed that the reward value and directional information were represented

506 separately in the SEF, but these signals were expected to be used in concert to adjust behavior (So

507 and Stuphorn 2012). Future studies should address how information regarding action and outcome

508 is integrated to adjust subsequent behavior.

509

510 ***Surprise signals in the frontal areas***

511   Surprise signals, responses to unpredicted outcomes, were also observed in other frontal areas,

512 such as the ACC, in single-cell recording studies (Amiez et al. 2005; Hayden et al. 2011; Shidara

513 and Richmond 2002). Correct/error-related activity was reported by single-cell recording studies (Ito

514 et al. 2003; Matsumoto et al. 2007) as well as by event-related potential studies (Emeric et al. 2008;

515 Gemba et al. 1986). Neurons in the lateral prefrontal cortex (LPFC) were also reported to reflect

516 positive or negative prediction errors (Asaad and Eskandar 2011; Seo and Lee 2007). The SEF is

517 connected anatomically to the ACC and LPFC (Huerta and Kaas 1990; Wang et al. 2005). These

518 functional and anatomical studies suggest that surprise signals are encoded in the frontal network in

519 the SEF, ACC, and LPFC.

520   Furthermore, human electroencephalogram studies have reported event-related potentials in the

521 medial frontal region in response to unexpected negative, unexpected positive, or novel stimuli

522 (Falkenstein et al. 1991; Gehring et al. 1993; Holroyd et al. 2008; Ridderinkhof et al. 2004; Wessel et

523 al. 2012). Indeed, the relationship between error-related responses and neuronal activity should be

524 investigated in future research.

525

Surprise signals and exploration–exploitation transitions

527 We found that the elevated activities of the surprise neurons in the SEF were followed by specific

528 search patterns. Combined with the behavioral analysis that showed distinctive explorative and

529 exploitative search patterns, our results suggest that the SEF neurons are related to the transitions

530 between exploration and exploitation in visual searches (Fig. 1). As a frontomedial region, the ACC

531 is functionally significant in adaptive, exploratory behavior (Daw et al. 2006; Procyk et al. 2000;

532 Shima and Tanji 1998; Walton et al. 2004; Williams et al. 2004). A recent study showed that the SEF

533 was more active during a matching-pennies task that required exploration than it was during a

534 simple saccade task (Donahue et al. 2013). A previous study demonstrated that ACC neurons

535 encoded the first correct feedback and contributed to the end of exploration in a target-search task

536 (Quilodran et al. 2008). Our study extends their findings by showing that the SEF may contribute not

537 only to the transition from exploration to exploitation but also to that from exploitation to exploration.

538 Thus, the question arises regarding the neuronal mechanisms that underlie exploration–exploitation

539 transitions after detection of surprise signals in the SEF. Cohen et al. proposed that cortical and

540 subcortical areas contribute differently to the management of exploration and exploitation (Cohen et

541 al. 2007). According to their model, cortical areas such as the ACC and the orbitofrontal cortex

542 (OFC) generate evaluation signals and send them to subcortical regions such as the locus

543 coeruleus (LC) or dopaminergic network. The subcortical systems regulate decision-making

544 processes (Doya 2002; Yu and Dayan 2005), and this activity has been related to behavioral

545 transitions (Barnes et al. 2005; Graybiel 2008). Cortical and subcortical contributions to exploration

546 and exploitation have also been supported by studies evaluating genetic factors in humans (Frank et

547 al. 2009). Based on these findings, our data suggest that, along with the ACC and OFC, the SEF

548 plays an evaluative role. Indeed, anatomical studies have shown that the SEF is connected to the

549 LC (Arnsten and Goldman-Rakic 1984; Shook et al. 1990) and the ACC (Huerta and Kaas 1990;

550 Wang et al. 2005). Thus, the SEF is expected to manage the balance between exploration and

551 exploitation by sending surprise signals to the subcortical systems, such as the LC.

552

553 ***Exploration–exploitation transition as set shifting***

554 Our task resembles the Wisconsin Card Sorting Test (WCST) in that it requires cognitive flexibility in

555 switching mental "sets" (Milner 1963; Stuss et al. 2000). Functional MRI and single-cell recordings

556 from monkeys performing an analog of the WCST have showed that the medial and lateral prefrontal

557 cortices are involved in set shifting (Kamigaki et al. 2012; Konishi et al. 2010; Nakahara et al. 2002).

558 Another study showed that ACC neurons responded to correct or error feedback and changed

559 activity depending on the outcome of the previous trial (Mansouri et al. 2006). Compared with the

560 WCST, our task requires understanding the spatial combinations of the targets, but it investigates

561 how the oculomotor system is involved in the task.

562   From a clinical perspective, our finding is consistent with a study of a patient with a focal SEF lesion

563 (Husain et al. 2003). This patient demonstrated difficulty adapting to unexpected rule changes in a

564 rule-reversal task. According to our findings, the patient may have been impaired in the ability to

565 switch between exploration and exploitation strategies appropriately.

566

567 ***Surprise signals from the perspective of predictive coding and timing***

568 The surprise neurons showed increased activity when the outcome was unexpected and decreased

569 activity when the same outcome was repeated. This pattern of responses fits well with predictive

570 coding theory, which holds that the brain infers the cause of sensory inputs based on a generative

571 model and that prediction errors are used to adjust the model until the error is minimized (Friston

572 2005). The framework of the predictive coding explains the neuronal mechanism for repetition

573 suppression, in which stimulus-evoked neural activity is attenuated when the same stimulus is

574 repeatedly presented and activity is enhanced when the stimulus is unpredictable (Summerfield et al.

575 2008; Todorovic and de Lange 2012). Thus, the surprise signal in our task corresponds to the

576 prediction error in predictive coding theory (Friston 2005).

577 We also found that a group of SEF neurons showed increased activity until the predicted event

578 occurred. This type of activity is consistent with predictive timing theory, which posits that the brain

579 estimates the timing of events and holds that prediction errors are generated when expected events

580 do not occur (Arnal and Giraud 2012; Nobre et al. 2007). The predictive timing theory may explain

581 various types of activity in SEF neurons, such as those that responded to predictable target motion

582 (Heinen and Liu 1997) and error-feedback (Stuphorn et al. 2000), reward-predicting (Amador et al.

583   2000), event-predicting, and event-detecting neurons (Campos et al. 2010). Additionally, the

584   pre-supplementary motor area has been reported to encode elapsed time (Mita et al. 2009).

585   Combined with these data, our results provide evidence that the SEF is involved in both predictive

586   coding and predictive timing. Furthermore, our study also suggests that the signals generated

587   according to these theories are important for making subsequent behavioral adjustments.

588

589   ***Effects of feedback signals on outcomes and surprise neurons***

590   We found a substantial difference in the number of neurons and the activity patterns of the error- and

591   correct-outcome neurons. Two possibilities for this difference are discussed below. The first

592   possibility is that these differences were due to variations in the visual stimuli that signal error and

593   correct feedback. Error feedback was signaled by the disappearance of visual stimuli, whereas

594   correct feedback was signaled by a change in the color of a fixation target. The second possibility

595   involves the difference in behavioral context. In our task, error feedback was critical when

596   exploitation was switched to exploration; in contrast, correct feedback was related to the

597   maintenance of current exploitation and the switch from exploration to exploitation. The

598   context-dependent difference in response to error or correct feedback is consistent with the previous

599   finding that error/correct-related activities were influenced by the likelihood of error in the behavioral

600   tasks (Jessup et al. 2010; Wessel et al. 2012).

601   We found 15 error-surprise neurons (10% of all feedback-related neurons) and 14 correct-surprise

602   neurons (10%). These numbers seem small compared with the numbers reported in previous

603 studies showing reward-prediction errors in the ACC (Hayden et al. 2011) and LPFC (Asaad and

604 Eskandar 2011; Seo and Lee 2007). However, our result is comparable with other studies evaluating

605 the SEF (So and Stuphorn 2012) and ACC (Matsumoto et al. 2007). The differences in the numbers

606 of neurons that encode reward-prediction errors may be explained by the type of feedback. Like the

607 present study, the research conducted by So's and Matsumoto's groups investigated neuronal

608 activities during the post-feedback period, and feedback was provided in the form of visual signals.

609 In contrast, the feedback in the former studies was an actual reward. Thus, differences in the type of

610 feedback may be a key contributor to the responsiveness of cortical neurons to unexpected

611 outcomes.

612

613 ***Limitations in the original model and alternatives to the model***

614 We found that the correct-surprise and correct-outcome neurons were suppressed strongly in

615 response to error feedback. These neurons were categorized as correct-outcome or surprise

616 neurons because we fitted each model to the post-feedback activity according to its relative firing

617 rates in correct and error trials. This result raises the possibility that SEF neurons change their

618 response preferentially to behavioral errors rather than to correct feedback. This preference for error

619 responses in the SEF is consistent with findings of the dominance of error-related negativity by

620 event-related potential studies (Falkenstein et al. 1991; Gehring et al. 1993; Holroyd and Coles

621 2002).

622 The correct-surprise neurons showed seemingly "*signed*" responses to the feedback (Fig. 6*C*,*D*),

623 because they exhibited error-related suppression. However, this activity was explained better by the

624 model of "*unsigned*" surprise signals. Indeed, the difference between actual and expected outcomes

625 was reflected only in an increase in the activity and not in a reduction in the activity of the

626 correct-surprise neurons. To resolve this apparent discrepancy, we hypothesized an additional

627 model of correct-surprise signals based on a combination of the two types of outcome activity (Fig.

628 9). In this model, the correct-outcome value was increased in response to correct feedback, and the

629 error-outcome value was reduced in response to error feedback. Previous studies have

630 demonstrated that SEF neurons reflect both types of outcome values, as predicted by this model

631 (Amador et al. 2000; Roesch and Olson 2003; So and Stuphorn 2012; Stuphorn et al. 2000). New

632 types of unsigned surprise signals should be investigated in the future.

633

634 ***Conclusion***

635 We found two types of surprise neurons in the SEF during the two-target oculomotor search task.

636 Error-surprise neurons responded to unexpected error feedback and were involved in the transition

637 from exploitation to exploration. Correct-surprise neurons responded to unexpected correct

638 feedback and were involved in the transition from exploration to exploitation. Our results suggest

639 that the SEF contributes to the control of exploration–exploitation transitions by detecting

640 unexpected behavioral outcomes.

641

642 **_Appendix_**

643 In this section, we present the values used in and theoretical background of our model.

644 *Outcome values.* Outcome values represent whether a specific outcome was fed back. Outcome

645 values at time t were divided into two types: the outcome value of correct feedback ($OV_{c,t}$) and the

646 outcome value of error feedback ($OV_{e,t}$). We set the $OV_{c,t}$ as 1 and the $OV_{e,t}$ as 0 for correct

647 feedback. In contrast, we set the $OV_{c,t}$ as 0 and the $OV_{e,t}$ as 1 for error feedback.

648

649 *Prediction values.* Prediction values represent the probability that a specific outcome occurred at

650 time t. Prediction values were divided into two types: the prediction value of correct feedback ($PV_{c,t}$)

651 and the prediction value of error feedback ($PV_{e,t}$). $PV_{c,t}$ and $PV_{e,t}$ represented the subjective

652 probability of correct and error feedback at time t, respectively. These values were updated at each

653 time t using a temporal difference algorithm (Alexander and Brown 2011; Sutton and Barto 1998).

654 Each iteration (dt) represents 10 ms of real time. The prediction values at time t ($PV_{i,t}$) were

655 computed as

656 $$PV_{i,t} = \Sigma X_t \times W_t,$$

657 where X_t is a set of task-related stimuli modeled as a temporal delay chain. W_t is a learned weight

658 and constrained by $W_t \rightarrow 0$. W_t was computed as follows:

659 $$W_{t+1} = W_t + \alpha \delta_t \varepsilon_t,$$

660 where α is the learning rate ($0.1 \leq \alpha \leq 0.3$) and ε_t is an eligibility trace at time t. $\delta_{c,t}$ and $\delta_{e,t}$ were the

661 temporal difference errors in correct or error feedback, respectively, at time t and were defined as

662 follows:

663
$$\delta_{c,t} = OV_{c,t+1} + \gamma PV_{c,t+1} - PV_{c,t}$$

664
$$\delta_{e,t} = OV_{e,t+1} + \gamma PV_{e,t+1} - PV_{e,t},$$

665 where $OV_{c,t}$ and $OV_{e,t}$ represent occurrences of correct and error feedback, respectively, at time t.

666 $PV_{c,t}$ and $PV_{e,t}$ are the prediction values of correct and error feedback, respectively, at time t. γ is a

667 temporal discounting factor ($0 \le \gamma \le 1$). $PV_{c,t}$ and $PV_{e,t}$ reflect the subjective probability of correct and

668 error feedback based on the histories of outcomes, regardless of actions.

669

670 *Multiple prediction errors.* In this model, signed and unsigned prediction errors were computed

671 based on differences between predicted and actual outcomes. First, we defined signed prediction

672 errors for correct and error feedback (*signed $PE_{c,t}$* and *signed $PE_{e,t}$*) as follows:

673
$$signed: PE_{c,t} = OV_{c,t} - PV_{c,t}$$

674
$$signed: PE_{e,t} = OV_{e,t} - PV_{e,t},$$

675 where $OV_{c,t}$ and $OV_{e,t}$ represent occurrences of correct and error feedback, respectively, at time t.

676 $PV_{c,t}$ and $PV_{e,t}$ are prediction values for correct and error feedback, respectively, at time t.

677 Next, we defined surprise signals (SS) as rectified unsigned prediction errors. The surprise signals

678 were divided into correct- or error-responsive types. An error-surprise signal ($SS_{e,t}$) was generated

679 when an unexpected error outcome was received at time t. A correct-surprise signal ($SS_{c,t}$) was

680 generated when an unexpected correct outcome was received at time t. Thus, these two values

681   were produced to respond to the unpredicted outcome of correct and error feedback at time t. The

682   values were defined as follows:

683
$$SS_{c,t} = [OV_{c,t} - PV_{c,t}]^+$$

684
$$SS_{e,t} = [OV_{e,t} - PV_{e,t}]^+,$$

685   where $[OV_{c,t} - PV_{c,t}]^+$ and $[OV_{e,t} - PV_{e,t}]^+$ are the rectified positive parts of the prediction errors. $OV_{c,t}$

686   and $OV_{e,t}$ represent occurrences of correct and error feedback, respectively, at time t. $PV_{c,t}$ and $PV_{e,t}$

687   are prediction values of correct and error feedback, respectively, at time t. The $SS_{c,t}$ and $SS_{e,t}$ are

688   high when the prediction values are low and unpredicted feedback occurs. These values were

689   shown in Figure 3. The values were emphasized for visualization purpose.

690

691

Figure legends

Figure 1. Hypothetical function of SEF neurons in the exploration–exploitation transitions

during an oculomotor search

We hypothesized that the SEF not only detects the feedback signals but contributes to the

subsequent behavioral adaptation. During exploitative oculomotor search, the subject predicts

correct feedback. When unpredicted-error feedback occurs, the SEF surprise neuron detects an

unpredicted error and produces an error-surprise signal to encourage the subject to shift from

exploitative to exploratory behavior. During exploratory search, the subject predicts error feedback.

When unexpected correct feedback is delivered, the SEF neuron produces a correct-surprise signal,

leading the subject to engage in exploitative behavior.

Figure 2. Oculomotor two-target search task

A: Sequence of events in two consecutive trials. The monkey chose one of four identical targets

(Top-left [TL], Top-right [TR], Bottom-left [BL], Bottom-right [BR]) by making a saccade to and fixating

on it. When a valid target was selected, it turned green and a reward was given. When an invalid

target was selected, the four targets disappeared, and the same trial was repeated until the valid

target was selected. The monkeys were required to choose two targets alternately. In this figure, the

valid-target pair consisted of the two targets, "Bottom-right" and "Bottom-left". The monkey had to

select the target "Bottom-right" (upper row) and then select the target "Bottom-left" in the next trial

(lower row). *B:* Four types of valid-target pairs. The current valid pair (two stimuli enclosed by the

712   dotted line) was switched to the other one in a pseudorandom manner. **C:** Three types of oculomotor

713   search patterns. The patterns were defined by the relative location of the currently chosen stimulus

714   compared with the location of the stimulus that was chosen in the previous trial. **D:** Examples of

715   behaviors before and after a valid-pair change. The monkey made alternate saccades to two valid

716   targets and received rewards in consecutive trials (exploitation phase). When the valid pair changed,

717   the monkey made an error and began to search for new valid targets (exploration phase). Within

718   several trials, the monkey found the valid pair of targets and began to select the two stimuli

719   alternately to gain the reward. Adj: adjacent search pattern. Diag: diagonal search pattern.

720

721   **Figure 3. Model of the error- and correct-surprise signals**

722   The surprise signal (*SS)* was calculated by the two basic values, the prediction value (*PV*) and the

723   outcome value (*OV*). The *OV* represents whether preferred feedback occurred (framed by a gray

724   line). The *PV* represents the subjective probability of the preferred feedback. We categorized the *PV*

725   into two types: higher and lower probability. The *OV* shows transient increases during feedback, and

726   the *PV* shows increases before feedback and decreases after feedback. Dotted lines indicate the

727   timing of the feedback. Both outcome values and surprise signals were divided into two types

728   depending on the preference for correct or error feedback. **A:** Simulated value of the error-surprise

729   signal (SS_e). The value is the rectified (positive) portion of the prediction error produced by

730   subtracting PV_e from OV_e. The SS_e is high when the error feedback was not expected but an error

731   occurred (blue, framed by a black line). The post-feedback response is lower when the error

732   feedback occurred repeatedly (sky blue). ***B:*** Correct-surprise signal (SS_e). The value is the rectified

733   portion of the prediction error produced by subtracting PV_c from OV_c. The SS_c is high when the

734   correct feedback was expected but an error occurred (red, framed by a black line). The response

735   was lower when correct feedback was expected and occurred repeatedly (magenta).

736

737   **Figure 4. Behavioral results**

738   ***A:*** The oculomotor search patterns depended on exploration and exploitation phases. The relative

739   incidences of the specific search patterns of two monkeys were plotted for each phase. The

740   adjacent search pattern was dominant during the exploitation phase. The diagonal search pattern

741   was common during the exploration phase. ***B:*** The search patterns in subsequent trials (n+1)

742   depended on the combination of the outcomes in the current (n) and previous (n-1) trials. The

743   diagonal search pattern was frequent after an error was fed back in the current (n) trial, especially

744   after feedback to a correct response was given in the previous (n-1) trials (correct → error). The

745   adjacent search pattern was dominant after feedback to a correct response was given in the current

746   (n) trial, irrespective of the outcomes in the previous (n-1) trials (correct → correct and error →

747   correct). The adjacent search pattern was divided into groups according to whether the selected

748   target was included in the current pair (other pair).

749

750   **Figure 5. Outcome neurons**

751   ***A:*** An example of an error-outcome neuron in the SEF. This neuron responded to error feedback,

752   regardless of the outcome of the preceding trial. This neuron did not respond to correct feedback. ***B:***

753   Population-level activity of the error-outcome neurons. *C:* An example of a correct-outcome neuron.

754   This neuron showed increasing activity in response to correct feedback, irrespectively of the

755   outcomes in the previous trials. *D:* The second example of correct-outcome neuron. This neuron

756   showed reduced activity in response to error feedback. *E:* Population-level activity of the

757   correct-outcome neurons. All firing rates were normalized, and the average *z*-scores were aligned to

758   error or correct feedback (FB, gray dotted lines), respectively. Colored dotted lines in *B* and *E*

759   indicate the standard errors of the mean *z*-scores.

760

761   **Figure 6. Surprise neurons**

762   *A:* An example of error-surprise neurons in the SEF. This neuron showed elevated activity when an

763   unpredicted error was fed back (correct→error). This post-feedback response decreased when the

764   error was repeated in successive trials (error→error). *B:* Population-level activity of the

765   error-surprise neurons. *C:* An example of a correct-surprise neuron. This neuron showed the

766   highest activity when unpredicted correct feedback was fed back (error→correct). This activity

767   decreased when correct feedback was repeated (correct→correct). *D:* Population-level activity of

768   the correct-surprise neurons. Conventions are the same as in Figure 5.

769

770   **Figure 7. Relationship between neuronal activity and search patterns**

771   *A:* Averaged activity of the error-surprise neurons. The averaged firing rates in the following trials

772   were plotted as follows: the first error trial after a valid-target change (1st E), the first correct trial after

773   a pair change (1st C), the second correct trial after a pair change (2nd C), and the correct trial after

774   consecutive correct trials (CC). Dotted lines show the timings of valid-pair changes. The activity was

775   high at the start of exploration after the valid-pair changes (1st E vs. CC, $*p = 0.012$). The activity in

776   the first error trial was higher than was that in the first correct and the second correct trials (1st E vs.

777   1st C, $*p = 0.029$; 1st E vs. 2nd C, $*p = 0.025$). **B:** Averaged activity of the correct-surprise neurons.

778   The activity level of correct-surprise neurons was high in the first correct trial after the valid-pair

779   changes ($*p = 0.040$). **C–G:** Post-feedback neuronal activities and the subsequent search patterns.

780   **C:** Enhanced activity of the error surprise-neurons was related to the diagonal search pattern in the

781   next trials (paired *t*-test; $*p = 0.040$). **D:** A higher level of activity of the correct-surprise neurons was

782   related to the adjacent search pattern in the subsequent trials (paired *t*-test; $*p = 0.010$). **E,F**: The

783   activities of the error-outcome (**E**) and correct-outcome neurons (**F**) were not related to the search

784   patterns in the next trials (paired *t*-test; $p\rightarrow 0.05$). All error bars indicate the standard error of the

785   mean. **G,H:** Choice probabilities for error-surprise and error-outcome neurons (**G**) and for

786   correct-surprise and correct-outcome neurons (**H**). Bar graphs show the mean choice probability

787   value of each group. The choice probability of the error-surprise neurons was significantly higher

788   than that of the error-outcome neurons ($**p = 0.0028$).

789

790   **Figure 8. Distribution of task-related neurons in the SEF**

791   Data are shown in the dorsal view of the frontal lobe. **A:** Task-related neurons. The positions of the

792   gray dots indicate the locations of the task-related neurons of two monkeys relative to the gross

793 morphological landmarks of the midline and the genu of the arcuate sulcus (AS Genu). The sizes of

794 the dots reflect the number of neurons recorded at each location. The dashed horizontal lines

795 indicate the frontal level of the AS genu . **B:** Error-surprise neurons. The dots indicate the location of

796 error-surprise neurons (orange) and error-outcome neurons (gray). **C:** Correct-surprise neurons.

797 The dots indicate correct-surprise neurons (green) and correct-outcome neurons (grey). The

798 relationship between the size of dots and the number of neurons is the same as in **B**.

799

800 **Figure 9. Hypothetical model of correct-surprise signals**

801 This surprise signal may explain the correct-surprise signals observed in Figure 6**C,D**. In this model,

802 a differential response to correct and error feedback was hypothesized, and the surprise signals

803 were associated with increasing activity in response to correct feedback and decreasing activity in

804 response to error feedback. FB: feedback.

805

806

807

808 **Table 1. Categorized neurons (Monkey K+ monkey M).**

809

	Error	Correct
Outcome neuron	47 (44 + 3)	29 (28 + 1)
Surprise neuron	15 (13 + 2)	14 (13 + 1)
Prediction-error neuron	3 (2 + 1)	8 (8 + 0)
Prediction neuron	20 (20 + 0)	10 (8 + 2)
Subtotal	85 (79 + 6)	61 (57 + 4)
Total	146	

810

811

812 **Table 2. Categorized neurons with overlap**

813

	Outcome value	Surprise signal	Prediction error
Error-outcome neuron	(47)	35	36
Error-surprise neuron	10	(15)	11
Correct-outcome neuron	(29)	7	9
Correct-surprise neuron	12	(14)	1

814

815 The figures in parentheses represent the total numbers of categorized neurons in Table 1.

816

817 **References**

818 **Alexander WH, and Brown JW**. Medial prefrontal cortex as an action-outcome predictor. *Nat Neurosci* 14:

819 1338-1344, 2011.

820 **Amador N, Schlag-Rey M, and Schlag J**. Reward-predicting and reward-detecting neuronal activity in the

821 primate supplementary eye field. *J Neurophysiol* 84: 2166-2170, 2000.

822 **Amiez C, Joseph JP, and Procyk E**. Anterior cingulate error-related activity is modulated by predicted reward.

823 *Eur J Neurosci* 21: 3447-3452, 2005.

824 **Amiez C, Sallet J, Procyk E, and Petrides M**. Modulation of feedback related activity in the rostral anterior

825 cingulate cortex during trial and error exploration. *Neuroimage* 63: 1078-1090, 2012.

826 **Arnal LH, and Giraud AL**. Cortical oscillations and sensory predictions. *Trends Cogn Sci* 16: 390-398, 2012.

827 **Arnsten AF, and Goldman-Rakic PS**. Selective prefrontal cortical projections to the region of the locus

828 coeruleus and raphe nuclei in the rhesus monkey. *Brain Res* 306: 9-18, 1984.

829 **Asaad WF, and Eskandar EN**. Encoding of both positive and negative reward prediction errors by neurons of

830 the primate lateral prefrontal cortex and caudate nucleus. *J Neurosci* 31: 17772-17787, 2011.

831 **Barnes TD, Kubota Y, Hu D, Jin DZ, and Graybiel AM**. Activity of striatal neurons reflects dynamic encoding

832 and recoding of procedural memories. *Nature* 437: 1158-1161, 2005.

833 **Berdyyeva TK, and Olson CR**. Rank signals in four areas of macaque frontal cortex during selection of

834 actions and objects in serial order. *J Neurophysiol* 104: 141-159, 2010.

835 **Britten KH, Newsome WT, Shadlen MN, Celebrini S, and Movshon JA**. A relationship between behavioral

836 choice and the visual responses of neurons in macaque MT. *Vis Neurosci* 13: 87-100, 1996.

837 **Buschman TJ, and Miller EK**. Serial, covert shifts of attention during visual search are reflected by the frontal

838   eye fields and correlated with population oscillations. *Neuron* 63: 386-396, 2009.

839   **Campos M, Breznen B, and Andersen RA**. A neural representation of sequential states within an instructed

840   task. *J Neurophysiol* 104: 2831-2849, 2010.

841   **Chen LL, and Wise SP**. Neuronal activity in the supplementary eye field during acquisition of conditional

842   oculomotor associations. *J Neurophysiol* 73: 1101-1121, 1995.

843   **Cohen JD, McClure SM, and Yu AJ**. Should I stay or should I go? How the human brain manages the

844   trade-off between exploitation and exploration. *Philos Trans R Soc Lond B Biol Sci* 362: 933-942, 2007.

845   **Courville AC, Daw ND, and Touretzky DS**. Bayesian theories of conditioning in a changing world. *Trends*

846   *Cogn Sci* 10: 294-300, 2006.

847   **Daw ND, O'Doherty JP, Dayan P, Seymour B, and Dolan RJ**. Cortical substrates for exploratory decisions in

848   humans. *Nature* 441: 876-879, 2006.

849   **Donahue CH, Seo H, and Lee D**. Cortical signals for rewarded actions and strategic exploration. *Neuron* 80:

850   223-234, 2013.

851   **Doya K**. Metalearning and neuromodulation. *Neural Netw* 15: 495-506, 2002.

852   **Emeric EE, Brown JW, Leslie M, Pouget P, Stuphorn V, and Schall JD**. Performance monitoring local field

853   potentials in the medial frontal cortex of primates: anterior cingulate cortex. *J Neurophysiol* 99: 759-772,

854   2008.

855   **Falkenstein M, Hohnsbein J, Hoormann J, and Blanke L**. Effects of crossmodal divided attention on late

856   ERP components. II. Error processing in choice reaction tasks. *Electroencephalogr Clin Neurophysiol* 78:

857   447-455, 1991.

858 **Frank MJ, Doll BB, Oas-Terpstra J, and Moreno F**. Prefrontal and striatal dopaminergic genes predict

859 individual differences in exploration and exploitation. *Nat Neurosci* 12: 1062-1068, 2009.

860 **Friston K**. A theory of cortical responses. *Philos Trans R Soc Lond B Biol Sci* 360: 815-836, 2005.

861 **Fujii N, Mushiake H, and Tanji J**. Distribution of eye- and arm-movement-related neuronal activity in the SEF

862 and in the SMA and Pre-SMA of monkeys. *J Neurophysiol* 87: 2158-2166, 2002.

863 **Gehring WJ, Goss B, Coles MGH, Meyer DE, and Donchin E**. A neural system for error-detection and

864 compensation. *Psychol Sci* 4: 385-390, 1993.

865 **Gemba H, Sasaki K, and Brooks VB**. 'Error' potentials in limbic cortex (anterior cingulate area 24) of

866 monkeys during motor learning. *Neuroscience Letters* 70: 223-227, 1986.

867 **Gittins JR**. Bandit processes and dynamic allocation indices. *J R Statist Soc B* 41: 148-177, 1979.

868 **Graybiel AM**. Habits, rituals, and the evaluative brain. *Annu Rev Neurosci* 31: 359-387, 2008.

869 **Hayden BY, Heilbronner SR, Pearson JM, and Platt ML**. Surprise signals in anterior cingulate cortex:

870 neuronal encoding of unsigned reward prediction errors driving adjustment in behavior. *J Neurosci* 31:

871 4178-4187, 2011.

872 **Heinen SJ, and Liu M**. Single-neuron activity in the dorsomedial frontal cortex during smooth-pursuit eye

873 movements to predictable target motion. *Vis Neurosci* 14: 853-865, 1997.

874 **Holroyd CB, and Coles MG**. The neural basis of human error processing: reinforcement learning, dopamine,

875 and the error-related negativity. *Psychol Rev* 109: 679-709, 2002.

876 **Holroyd CB, Pakzad-Vaezi KL, and Krigolson OE**. The feedback correct-related positivity: sensitivity of the

877 event-related brain potential to unexpected positive feedback. *Psychophysiology* 45: 688-697, 2008.

878   **Huerta MF, and Kaas JH**. Supplementary eye field as defined by intracortical microstimulation: connections

879   in macaques. *J Comp Neurol* 293: 299-330, 1990.

880   **Husain M, Parton A, Hodgson TL, Mort D, and Rees G**. Self-control during response conflict by human

881   supplementary eye field. *Nat Neurosci* 6: 117-118, 2003.

882   **Isoda M, and Tanji J**. Contrasting neuronal activity in the supplementary and frontal eye fields during

883   temporal organization of multiple saccades. *J Neurophysiol* 90: 3054-3065, 2003.

884   **Ito S, Stuphorn V, Brown JW, and Schall JD**. Performance monitoring by the anterior cingulate cortex during

885   saccade countermanding. *Science* 302: 120-122, 2003.

886   **Jessup RK, Busemeyer JR, and Brown JW**. Error effects in anterior cingulate cortex reverse when error

887   likelihood is high. *J Neurosci* 30: 3467-3472, 2010.

888   **Kamigaki T, Fukushima T, Tamura K, and Miyashita Y**. Neurodynamics of cognitive set shifting in monkey

889   frontal cortex and its causal impact on behavioral flexibility. *J Cogn Neurosci* 24: 2171-2185, 2012.

890   **Konishi S, Hirose S, Jimura K, Chikazoe J, Watanabe T, Kimura HM, and Miyashita Y**. Medial prefrontal

891   activity during shifting under novel situations. *Neuroscience Letters* 484: 182-186, 2010.

892   **Krebs JR, Kacelnik A, and Taylor P**. Test of optimal sampling by foraging great tits. *Nature* 275: 27-31, 1978.

893   **Lu X, Matsuzawa M, and Hikosaka O**. A neural correlate of oculomotor sequences in supplementary eye

894   field. *Neuron* 34: 317-325, 2002.

895   **Mansouri FA, Matsumoto K, and Tanaka K**. Prefrontal cell activities related to monkeys' success and failure

896   in adapting to rule changes in a Wisconsin Card Sorting Test analog. *J Neurosci* 26: 2745-2756, 2006.

897   **Matsumoto M, Matsumoto K, Abe H, and Tanaka K**. Medial prefrontal cell activity signaling prediction errors

898   of action values. *Nat Neurosci* 10: 647-656, 2007.

899   **Middlebrooks PG, and Sommer MA**. Neuronal correlates of metacognition in primate frontal cortex. *Neuron*

900   75: 517-530, 2012.

901   **Milner B**. Effects of Different Brain Lesions on Card Sorting - Role of Frontal Lobes. *Archives of Neurology* 9:

902   90-&, 1963.

903   **Mirpour K, and Bisley JW**. Dissociating activity in the lateral intraparietal area from value using a visual

904   foraging task. *Proc Natl Acad Sci U S A* 109: 10083-10088, 2012.

905   **Mita A, Mushiake H, Shima K, Matsuzaka Y, and Tanji J**. Interval time coding by neurons in the

906   presupplementary and supplementary motor areas. *Nat Neurosci* 12: 502-507, 2009.

907   **Moorman DE, and Olson CR**. Impact of experience on the representation of object-centered space in the

908   macaque supplementary eye field. *J Neurophysiol* 97: 2159-2173, 2007.

909   **Mushiake H, Saito N, Sakamoto K, Itoyama Y, and Tanji J**. Activity in the lateral prefrontal cortex reflects

910   multiple steps of future events in action plans. *Neuron* 50: 631-641, 2006.

911   **Nakahara K, Hayashi T, Konishi S, and Miyashita Y**. Functional MRI of macaque monkeys performing a

912   cognitive set-shifting task. *Science* 295: 1532-1536, 2002.

913   **Nobre A, Correa A, and Coull J**. The hazards of time. *Curr Opin Neurobiol* 17: 465-470, 2007.

914   **Pearce JM, and Hall G**. A model for Pavlovian learning: variations in the effectiveness of conditioned but not

915   of unconditioned stimuli. *Psychol Rev* 87: 532-552, 1980.

916   **Procyk E, Tanaka YL, and Joseph JP**. Anterior cingulate activity during routine and non-routine sequential

917   behaviors in macaques. *Nat Neurosci* 3: 502-508, 2000.

918   **Quilodran R, Rothe M, and Procyk E**. Behavioral shifts and action valuation in the anterior cingulate cortex.

919   *Neuron* 57: 314-325, 2008.

920   **Ridderinkhof KR, Ullsperger M, Crone EA, and Nieuwenhuis S**. The role of the medial frontal cortex in

921   cognitive control. *Science* 306: 443-447, 2004.

922   **Roesch MR, and Olson CR**. Impact of expected reward on neuronal activity in prefrontal cortex, frontal and

923   supplementary eye fields and premotor cortex. *J Neurophysiol* 90: 1766-1789, 2003.

924   **Schall JD, and Boucher L**. Executive control of gaze by the frontal lobes. *Cogn Affect Behav Neurosci* 7:

925   396-412, 2007.

926   **Schall JD, Stuphorn V, and Brown JW**. Monitoring and control of action by the frontal lobes. *Neuron* 36:

927   309-322, 2002.

928   **Schlag-Rey M, Amador N, Sanchez H, and Schlag J**. Antisaccade performance predicted by neuronal

929   activity in the supplementary eye field. *Nature* 390: 398-401, 1997.

930   **Schlag J, and Schlag-Rey M**. Evidence for a supplementary eye field. *J Neurophysiol* 57: 179-200, 1987.

931   **Schultz W**. Behavioral theories and the neurophysiology of reward. *Annu Rev Psychol* 57: 87-115, 2006.

932   **Schultz W**. A neural substrate of prediction and reward. *Science* 275: 1593-1599, 1997.

933   **Seo H, and Lee D**. Temporal filtering of reward signals in the dorsal anterior cingulate cortex during a

934   mixed-strategy game. *J Neurosci* 27: 8366-8377, 2007.

935   **Shidara M, and Richmond BJ**. Anterior cingulate: single neuronal signals related to degree of reward

936   expectancy. *Science* 296: 1709-1711, 2002.

937   **Shima K, and Tanji J**. Role for cingulate motor area cells in voluntary movement selection based on reward.

938   *Science* 282: 1335-1338, 1998.

939   **Shook BL, Schlag-Rey M, and Schlag J**. Primate supplementary eye field: I. Comparative aspects of

940   mesencephalic and pontine connections. *J Comp Neurol* 301: 618-642, 1990.

941   **So N, and Stuphorn V**. Supplementary eye field encodes reward prediction error. *J Neurosci* 32: 2950-2963,

942   2012.

943   **Stuphorn V, and Schall JD**. Executive control of countermanding saccades by the supplementary eye field.

944   *Nat Neurosci* 9: 925-931, 2006.

945   **Stuphorn V, and Schall JD**. Neuronal control and monitoring of initiation of movements. *Muscle & Nerve* 26:

946   326-339, 2002.

947   **Stuphorn V, Taylor TL, and Schall JD**. Performance monitoring by the supplementary eye field. *Nature* 408:

948   857-860, 2000.

949   **Stuss DT, Levine B, Alexander MP, Hong J, Palumbo C, Hamer L, Murphy KJ, and Izukawa D**. Wisconsin

950   Card Sorting Test performance in patients with focal frontal and posterior brain damage: effects of lesion

951   location and test structure on separable cognitive processes. *Neuropsychologia* 38: 388-402, 2000.

952   **Summerfield C, Trittschuh EH, Monti JM, Mesulam MM, and Egner T**. Neural repetition suppression reflects

953   fulfilled perceptual expectations. *Nat Neurosci* 11: 1004-1006, 2008.

954   **Sutton RS, and Barto AG**. *Reinforcement learning*. Cambridge, MA: MIT Press, 1998.

955   **Todorovic A, and de Lange FP**. Repetition suppression and expectation suppression are dissociable in time

956   in early auditory evoked fields. *J Neurosci* 32: 13389-13395, 2012.

957   **Uchida Y, Lu X, Ohmae S, Takahashi T, and Kitazawa S**. Neuronal activity related to reward size and

958   rewarded target position in primate supplementary eye field. *J Neurosci* 27: 13750-13755, 2007.

959   **Walton ME, Devlin JT, and Rushworth MF**. Interactions between decision making and performance

960   monitoring within prefrontal cortex. *Nat Neurosci* 7: 1259-1265, 2004.

961   **Wang Y, Isoda M, Matsuzaka Y, Shima K, and Tanji J**. Prefrontal cortical cells projecting to the

962   supplementary eye field and presupplementary motor area in the monkey. *Neurosci Res* 53: 1-7, 2005.

963   **Wessel JR, Danielmeier C, Morton JB, and Ullsperger M**. Surprise and error: common neuronal architecture

964   for the processing of errors and novelty. *J Neurosci* 32: 7528-7537, 2012.

965   **Williams ZM, Bush G, Rauch SL, Cosgrove GR, and Eskandar EN**. Human anterior cingulate neurons and

966   the integration of monetary reward with motor responses. *Nat Neurosci* 7: 1370-1375, 2004.

967   **Yang SN, and Heinen S**. Contrasting the roles of the supplementary and frontal eye fields in ocular decision

968   making. *J Neurophysiol* 111: 2644-2655, 2014.

969   **Yu AJ, and Dayan P**. Uncertainty, neuromodulation, and attention. *Neuron* 46: 681-692, 2005.
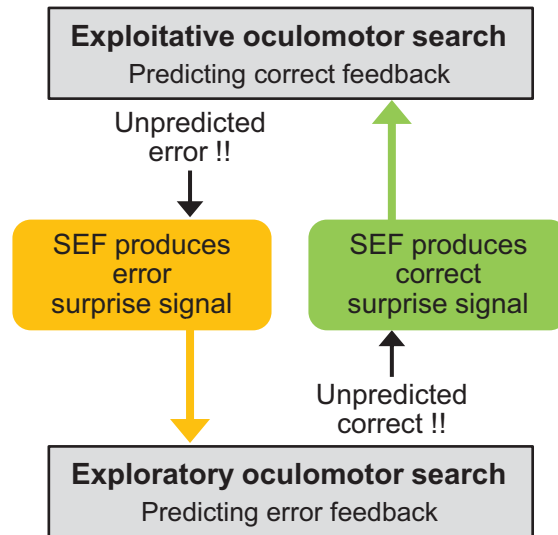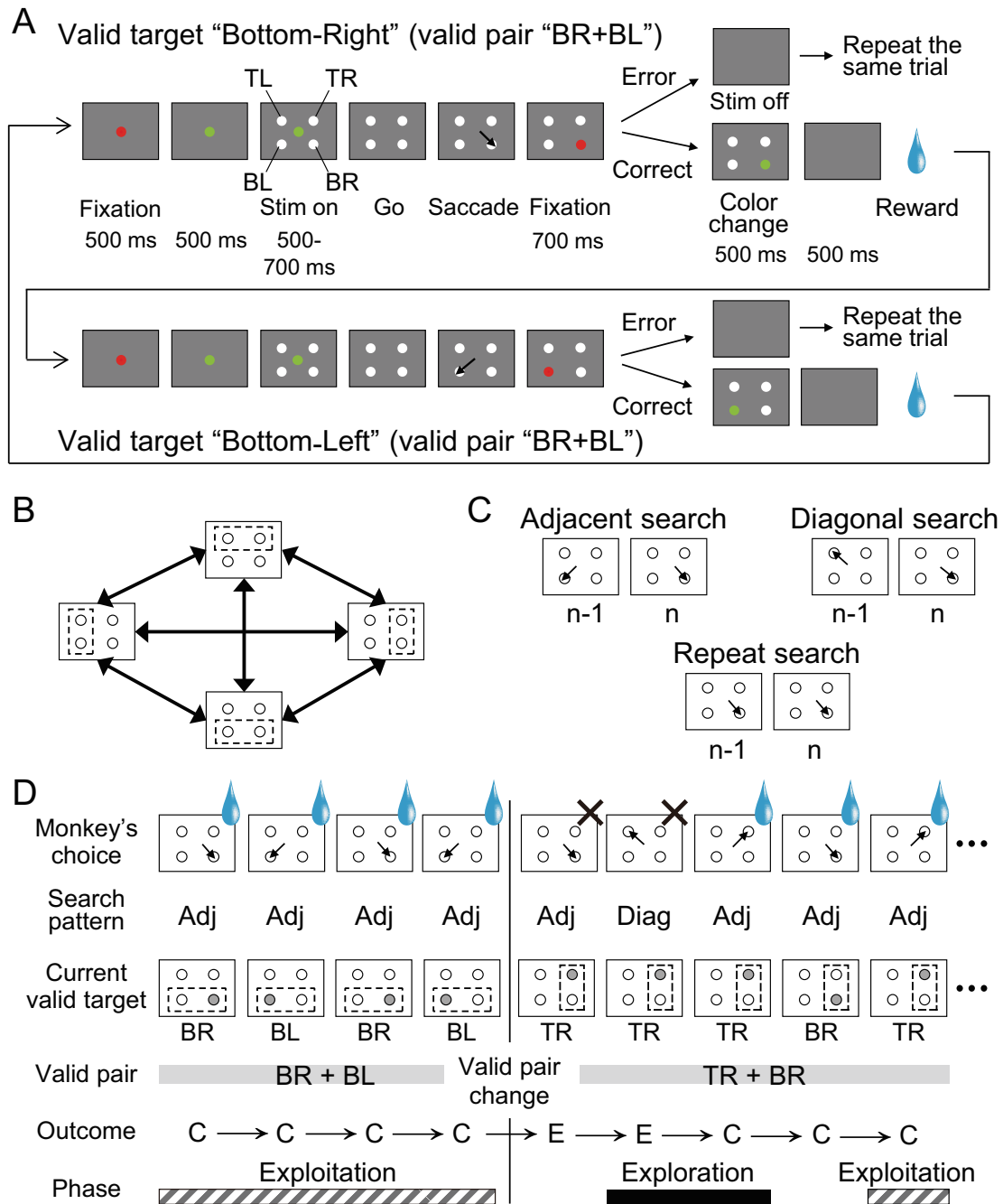
970

971

Figure 1

Figure 2

A Valid target "Bottom-Right" (valid pair "BR+BL")

Fixation 500 ms — 500 ms — Stim on 500-700 ms — Go — Saccade — Fixation 700 ms — Error → Stim off → Repeat the same trial — Correct → Color change 500 ms — 500 ms — Reward

TL TR
BL BR

Valid target "Bottom-Left" (valid pair "BR+BL")

Error → Repeat the same trial
Correct → Reward

B

C Adjacent search   Diagonal search
   n-1   n   n-1   n
   Repeat search
   n-1   n

D
Monkey's choice
Search pattern   Adj Adj Adj Adj | Adj Diag Adj Adj Adj ...
Current valid target   BR BL BR BL | TR TR TR BR TR ...
Valid pair   BR + BL | Valid pair change | TR + BR
Outcome   C → C → C → C | → E → E → C → C → C
Phase   Exploitation | Exploration | Exploitation

Figure 3

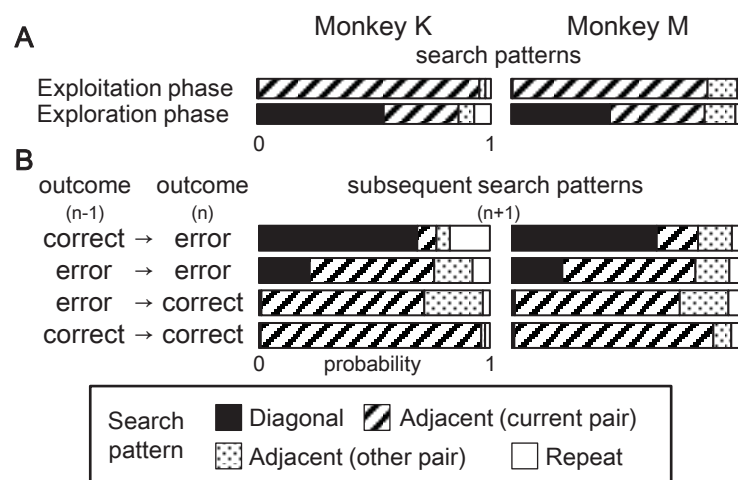**A** Error-surprise signal
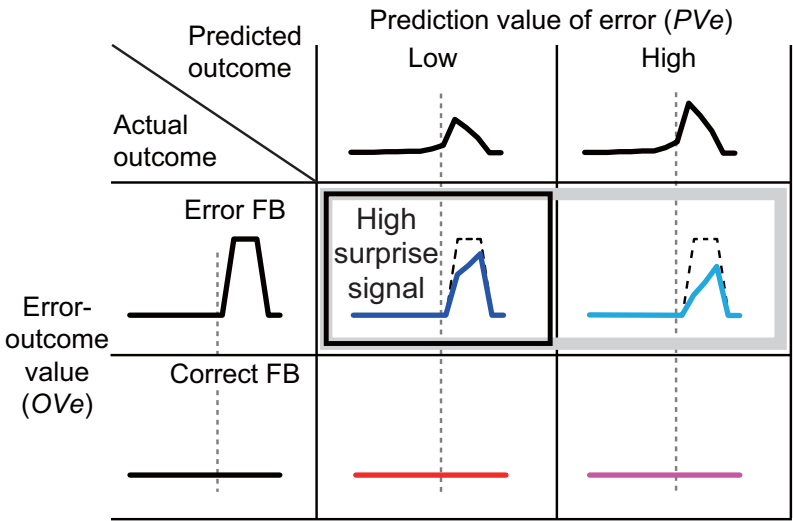
B Correct-surprise signal

Figure 4

A

Monkey K   Monkey M

search patterns

Exploitation phase
Exploration phase

0   1

B

outcome	outcome	subsequent search patterns
(n-1)	(n)	(n+1)
correct →	error	
error →	error	
error →	correct	
correct →	correct	

0   probability   1

Search pattern: ■ Diagonal   ▨ Adjacent (current pair)   ▦ Adjacent (other pair)   □ Repeat

Figure 5

Error-outcome neurons

A Single neuron

B Population

n=47

Correct-outcome neurons

C Single neuron

n-1		n
Correct	→	Error
Error	→	Error
Error	→	Correct
Correct	→	Correct

D Single neuron

E Population

n=29

Figure 6

Error-surprise neuron

A Single neuron

B Population

Correct-surprise neuron

C Single neuron

D Population

n-1		n
Correct	→	Error
Error	→	Error
Error	→	Correct
Correct	→	Correct

Figure 7

Figure 8

A Task-related neurons

> 11 cells
6-10 cells
1-5 cells

← AS Genu

midline 5 mm

B Error-surprise neurons

> 3 cells
1-3 cells

● Error-surprise neurons
● Error-outcome neurons

C Correct-surprise neurons

● Correct-surprise neurons
● Correct-outcome neurons

Figure 9

Correct-surprise signal + baseline