

1
2
3
4
5
6
7
8
9

Statistical context dictates the relationship between feedback-related EEG signals and learning

10 Matthew R. Nassar^{1,2}, Rasmus Bruckner^{4,5,6}, Michael J. Frank^{1,3}

11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27

¹Robert J. & Nancy D. Carney Institute for Brain Science, Brown University, Providence RI 02912-1821, USA

²Department of Neuroscience, Brown University, Providence RI 02912-1821, USA

³Department of Cognitive, Linguistic, and Psychological Sciences, Brown University, Providence RI 02912-1821, USA

⁴Department of Education and Psychology, Freie Universität Berlin, 14195 Berlin, Germany

⁵Center for Lifespan Psychology, Max Planck Institute for Human Development, 14195 Berlin, Germany

⁶International Max Planck Research School on the Life Course (LIFE), Berlin, Germany

28
29
30
31
32
33

Acknowledgements:

34
35
36
37
38
39
40
41
42

We would like to thank Julie Helmers and Andrea Mueller for their help collecting EEG and behavioral data. This work was funded by NIH grants F32MH102009 and K99AG054732 (MRN), NIMH R01 MH080066-01 and NSF Proposal #1460604 (MJF). RB was supported by a Promos travel grant from the German Academic Exchange Service (DAAD). The funders had no role in study design, data collection and analysis, decision to publish or preparation of the manuscript.

43
44

Competing interests:

45
46

The authors have no financial or non-financial conflicts of interest related to this work.

47 **Abstract**

48

49 Successful decision-making requires learning expectations based on
50 experienced outcomes. This learning should be calibrated according to the
51 surprise associated with an outcome, but also to the statistical context dictating
52 the most likely source of surprise. For example, when occasional changepoints
53 are expected, surprising outcomes should be weighted heavily, demanding
54 increased learning. In contrast, when signal corruption is expected to occur
55 occasionally, surprising outcomes can suggest a corrupt signal that should be
56 ignored by learning systems. Here we dissociate surprising outcomes from the
57 degree to which they demand learning using a predictive inference task and
58 computational modeling. We show that the updating P300, a stimulus-locked
59 electrophysiological response previously associated with adjustments in learning
60 behavior, does so conditionally on the source of surprise. Larger P300 signals
61 predicted greater learning in a changing context, but predicted less learning in a
62 context where surprise was indicative of a one-off outlier (oddball). The
63 conditional predictive relationship between the P300 and learning behavior was
64 persistent even after adjusting for known sources of learning rate variability. Our
65 results suggest that the P300 provides a surprise signal that is interpreted by
66 downstream learning processes differentially according to statistical context in
67 order to appropriately calibrate learning across complex environments.

68

69

70

71 **Introduction**

72

73 People are capable of rationally adjusting the degree to which they incorporate
74 new information into their beliefs about the world (1-5). In environments that
75 include discontinuous changes (changepoints) normative learning requires
76 increasing learning when beliefs are uncertain or when observations are most
77 surprising (2,6). Human participants display both of these tendencies, albeit to
78 varying degrees (2,6,7).

79

80 A major open question in the learning domain is how the brain achieves such
81 apparent adjustments in learning rate. This question has fueled a number of
82 recent studies that have identified neural correlates of surprise in functional
83 magnetic resonance imaging (fMRI) (8), electroencephalography (EEG) (9,10),
84 and pupil signals (6) that predict subsequent learning behavior. These signals
85 might reflect candidate mechanisms for a general system to adjust learning rate
86 (1,11,12), yet the generality has yet to be established outside of discontinuously
87 changing environments, where surprise and learning are tightly coupled.

88

89 The relationship between surprise and learning is complex and depends critically
90 on the overarching statistical context. We refer to learning as the degree to which

91 an observed prediction error promotes measurable behavioral updating. While
92 changing environments require increased learning in the face of surprising
93 information, stable environments with outliers (“oddballs”), dictate less learning
94 from surprising information (4). People are capable of this type of robust learning
95 rate adjustment that deemphasizes surprising information (3,4,13), yet the
96 learning signals measured under such conditions do not correspond directly to
97 those observed in changing environments. Most notably, a number of candidate
98 learning signals measured through fMRI do not reflect learning rate when
99 considering a broader set of statistical contexts (4).

100

101 However, prior studies on EEG correlates of learning seem to favor the idea that
102 a late, stimulus-locked positivity referred to as the P300, tracks learning in a
103 broader range of statistical contexts. The central parietal component of the P300
104 (P3b) reflects surprise (14) and relates to learning (15) even after controlling for
105 the degree of surprise in changing environments (9,10). In a stationary
106 environment where integration of sequential samples is required to make a
107 subsequent decision, a late posterior positivity, reminiscent of the P300, predicts
108 the degree to which a particular sample influences the subsequent decision (16).
109 Interestingly, within this particular task, more surprising outcomes tended to exert
110 less influence on decisions (3,13), suggesting that this late positivity might
111 provide a general learning or updating signal, irrespective of statistical context.
112 This idea would be in line with a prominent theory of P3b function, which
113 emphasizes its role in updating context representations – sometimes defined in
114 terms of items stored in working memory (17-20).

115

116 Here we tested the idea that the P3b provides a general learning signal that is
117 independent of the statistical context. In particular, we measured learning
118 behavior using a modified predictive inference task and normative learning model
119 and examined how learning behavior and surprise related to evoked potentials
120 measured through EEG. We found that people are capable of contextually
121 adjusting learning in response to surprise: they tended to learn more from
122 surprising outcomes when those outcomes were indicative of changepoints, but
123 learned less from surprising outcomes when those outcomes were indicative of
124 an oddball. Outcome evoked potentials reminiscent of a parietal P300 were
125 related to surprising events irrespective of context. The magnitude of this P300
126 response on a given trial positively predicted learning in the presence of
127 changepoints, but negatively predicted learning in the presence of oddballs.
128 These relationships persisted even when controlling for variability in learning
129 behavior that could be explained by the best behavioral model. Taken together
130 these findings suggest that the P300 does not naively reflect increased
131 behavioral updating, but may play a role in adaptively increasing or decreasing
132 learning in response to surprising information, depending on the statistical
133 context.

134

135

136 **Results**

137

138 We used EEG to measure electrophysiological signatures of feedback
139 processing while participants performed a modified predictive inference task (2)
140 designed to dissociate surprise from learning. Predictions were made in the
141 context of a video game that required participants to place a shield at a location
142 on a circle in order to block cannonballs that would be fired from a cannon
143 located at the center of the circle (Fig 1A). Surprise and learning were
144 manipulated independently using two different task conditions. In the *oddball*
145 condition, the aim of the cannon drifted slowly from one trial to the next (Fig 1B,
146 dotted line) and cannonball locations were distributed around the point of cannon
147 aim (Fig 1B, green points nearby dotted line) or, occasionally and unpredictably,
148 uniformly distributed around the circle (*oddballs*; see green point on trial 11 of Fig
149 1B for example). In the *change point* condition, the cannon aim remained constant
150 for an unpredictable duration, and was then re-aimed at a new location on the
151 circle at random (*change points*; Fig 1C, dotted line). Cannonball locations were
152 always distributed around the point of cannon aim in this condition (Fig 1C, green
153 points).

154

155 *Behavior of human participants and normative model*

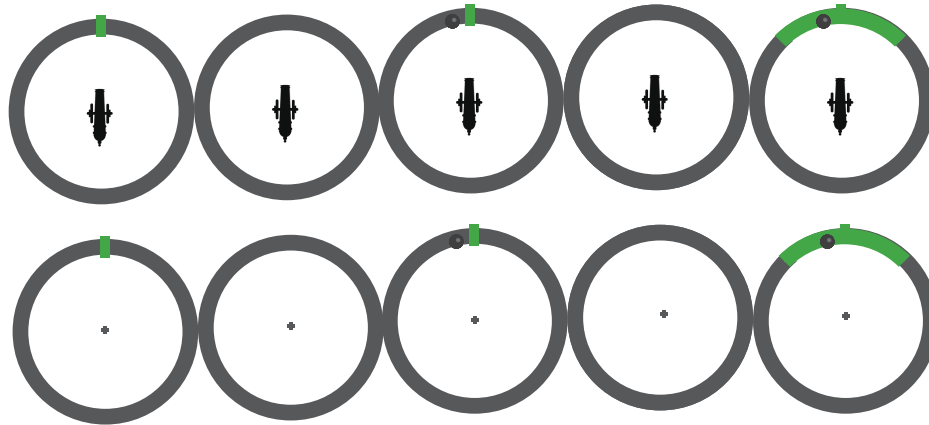
156

157 In both conditions, participants were instructed to place a shield on each trial in
158 order to maximize the chances of blocking the upcoming cannonball (Figure
159 1B&C, orange line). However, behavior differed qualitatively in these two
160 conditions, which can be observed clearly in the example participant data in
161 Figure 1. In particular, shield placements were not updated in response to
162 extreme outcomes in the oddball condition (*oddballs*; Fig 1B) but were updated
163 dramatically in response to extreme outcomes in the change point condition
164 (*change points*; Fig 1C).

165

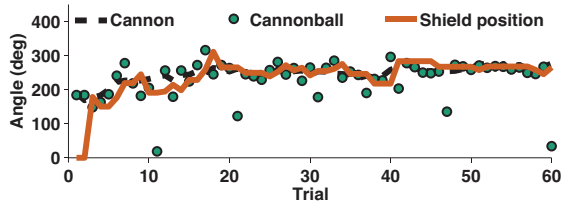
166

A



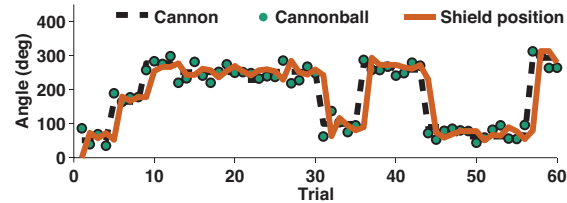
B

Oddball

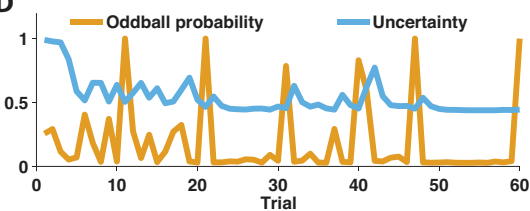


C

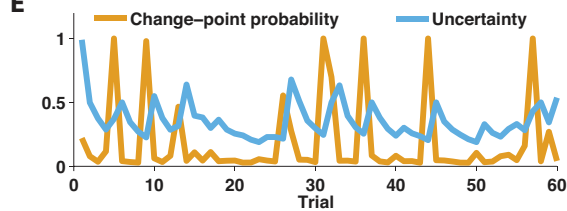
Changepoint



D



E



167

168

169

170

171

172

173

174

175

176

177

178

179

180

181

182

183

184

185

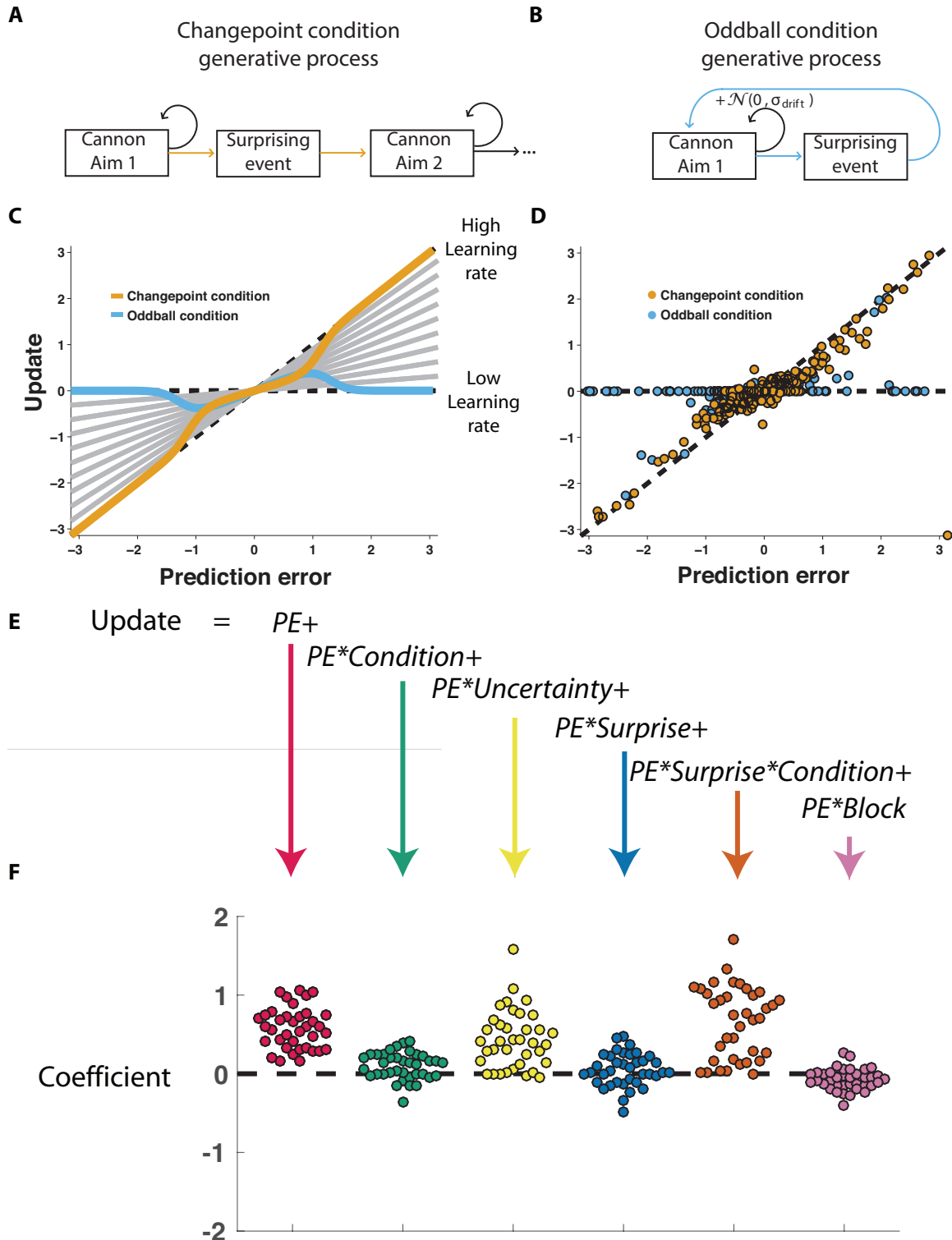
186

Figure 1: Measuring learning in different statistical contexts with a predictive inference task. **A)** Participants were trained to place the center of a shield (green tick; prediction phase [left]) at the aim location of a cannon (training task; top) in order to block a cannonball shot from it (outcome phase [top middle]) with a shield that varied in size from trial to trial and was revealed at the end of the trial (shield phase [top right]). After training, participants were asked to complete the same task, but without a visual depiction of the cannon, which required them to infer the aim of the cannon based on the sequence of previously observed cannonballs (test task; [bottom]). **B)** In oddball blocks, cannon aim (dotted black line) followed a random walk and cannonball locations were typically drawn from a Von Mises distribution centered on the true cannon aim (green points), but occasionally drawn from a uniform distribution across the entire circle (oddball trials). Participants placed their shield on each trial (brown line) providing information about their inference about the cannon aim. **C)** In changepoint blocks, cannon aim was stationary for most trials but was occasionally resampled uniformly from possible angles (changepoint) and cannonball locations were always drawn from a Von Mises distribution centered on the true cannon aim (green points). **D&E)** Optimal inference could be approximated in both generative environments by tracking and adjusting learning according to relative uncertainty and the probability of an unlikely event (oddball or changepoint).

187 To quantitatively analyze the differences between the two task conditions, we
188 extended a previously developed normative learning model (2,7). The model
189 approximates optimal inference using an error-driven learning rule by adjusting
190 learning from trial to trial according to two latent variables. The first latent variable
191 tracks the probability with which the most recent outcome was generated from an
192 unexpected generative process (oddball probability in Fig 1D; changepoint
193 probability in Fig 1E), whereas the second latent variable tracks the model's
194 uncertainty about the true cannon aim (Fig 1D&E; uncertainty). Critically, the
195 model stipulates that surprising events in the oddball condition, which is tracked
196 through the model's estimate of oddball probability, should reduce learning, as
197 oddballs are unrelated to future cannonball locations (4). In contrast, the model
198 stipulates that surprising events in the changepoint condition, which are tracked
199 through the model's estimate of changepoint probability, should amplify learning,
200 as changepoints render prior cannonballs (and thus prior beliefs) irrelevant to the
201 problem of predicting future ones (21,22). Qualitatively, behavior from the
202 example participant seems to follow these prescriptions, with adjustments in
203 shield position fairly minimal on trials that include a spike in oddball probability
204 (Fig 1 B,D), but fairly large on trials that include a spike in changepoint probability
205 (Fig 1 C,E).

206
207 The normative model also makes quantitative prescriptions for how learning
208 should be adjusted according to surprise differentially in the changepoint and
209 oddball conditions. The surprise of a given outcome can be measured crudely
210 through the degree to which a cannonball location differed from that which was
211 predicted (e.g., the shield position). Larger absolute prediction errors indicate a
212 higher degree of surprise, and higher oddball or changepoint probabilities
213 depending on the task condition. Learning in this task can be measured through
214 the degree to which a participant adjusts the shield position in response to a
215 given prediction error (2), and a fixed rate of learning would correspond to a
216 straight line mapping each prediction error onto a corresponding shield update,
217 where the slope of the line can be thought of as the *learning rate* (Fig 2C, gray
218 lines). The normative learning model does not prescribe a fixed learning rate
219 across all levels of surprise; instead it prescribes higher learning rates for more
220 surprising outcomes in the changepoint condition (Fig 2C, orange) and lower
221 learning rates for more surprising outcomes in the oddball condition (Fig 2C,
222 blue).

223
224
225
226
227



228
229
230
231
232
233
234

Figure 2: Participants scale learning according to surprise differently in changepoint and oddball contexts as would be expected for normative learning rate adjustment. **A)** In the changepoint condition, surprising events (change points) signaled a transition in the aim of the cannon whereas **B)** in the oddball condition, surprising events (oddballs) were unrelated to the process through which the aim of the cannon transitioned. **C)** Learning rate in the cannon task can be described by the slope of the relationship between prediction error (signed distance between

235 cannonball and shield; abscissa) and update (signed change in shield position after observing
236 new cannonball location; ordinate). Fixed learning rate updating corresponds to a line in this
237 space whose slope is uniform across prediction errors and reflects the learning rate (gray lines).
238 In contrast, normative learning dictates that the slope should decrease for extreme prediction
239 errors in the oddball condition (blue) but increase for extreme prediction errors in the changepoint
240 condition (orange). **D**) Prediction error (abscissa) and update (ordinate) for each trial (points) in
241 each condition (designated by color) completed by a single example participant. **E**) Trial updates
242 for each subject were fit with a regression model that included prediction errors (to measure fixed
243 learning rate) as well as several interaction terms to assess how learning depended on various
244 factors. **F**) Coefficients from regression model fit to individual subjects (points) revealed an overall
245 tendency to update toward recent cannonball locations (red, $t = 3.5$, $dof = 36$, $p = 10^{-15}$), and a
246 tendency to do so more in the changepoint condition (green, $t = 13.5$, $dof = 36$, $p = 0.001$), when
247 uncertain (yellow, $t = 7.2$, $dof = 36$, $p = 2 \times 10^{-8}$), and on trials where the cannonball was not
248 blocked by the shield (pink, $t = -3.3$, $dof = 36$, $p = 0.002$). The model revealed that there was no
249 consistent effect of surprise on learning across both conditions (blue, $t = 1.5$, $dof = 36$, $p = 0.15$),
250 but that there was a strong interaction between surprise and condition (orange, $t = 8.8$, $dof = 36$,
251 $p = 2 \times 10^{-10}$) whereby surprise tended to increase learning in the changepoint condition but
252 decrease learning in the oddball condition.

253

254

255 Participants adjusted learning behavior in accordance with normative predictions,
256 albeit with considerable heterogeneity across trials and participants. Shield
257 updating behavior and corresponding prediction errors for an example participant
258 reveal the basic trend predicted by the normative model, although exact updates
259 were variable from one trial to the next (Fig 2D). To summarize the degree to
260 which updating behavior of individual subjects was contingent on key task
261 variables, we constructed a linear regression model that described trial-by-trial
262 updates in terms of prediction errors as well as key task variables thought to
263 modulate the degree to which prediction errors are translated into updates (Fig
264 2E) including condition (changepoint versus oddball block), surprise (as
265 measured by changepoint or oddball probability estimates from normative
266 model), and their multiplicative interaction (capturing the degree to which learning
267 is increased for surprising outcomes in the changepoint context, but decreased
268 for surprising outcomes in the oddball context). As expected, prediction error
269 coefficients were positive, capturing a tendency for participants to update shield
270 position toward the most recent cannonball position (Fig 1F, red; mean/SEM beta
271 = $0.58/0.04$, $t = 13.5392$, $dof = 36$, $p = 10^{-15}$). Furthermore, participants
272 systematically adjusted the degree to which they did so according to condition
273 (Fig 1F, green; mean/SEM beta = $0.1/0.03$, $t = 3.5$, $dof = 36$, $p = 0.001$), but not
274 significantly according to surprise (Fig 1F, blue; mean/SEM beta = $0.05/0.03$, $t =$
275 1.5 , $dof = 36$, $p = 0.15$). Critically, surprise robustly impacted learning in opposite
276 directions for the two conditions, as indicated by the interaction between surprise
277 and condition (Fig 1F, orange; mean/SEM beta = $0.66/0.07$, $t = 8.8$, $dof = 36$, $p =$
278 1.5×10^{-10}). Specifically, positive coefficients indicate that sensitivity to prediction
279 errors was increased for surprising outcomes in the changepoint condition and
280 decreased for surprising outcomes in the oddball condition, as predicted by the
281 normative model.

282

283 *Electrophysiological signatures of feedback processing*

284

285 We took a data driven approach to identify electrophysiological signatures of
286 feedback processing. First we regressed feedback-locked EEG data collected
287 simultaneously with task performance onto an explanatory matrix that included
288 separate binary variables reflecting changepoint and oddball trials, amongst other
289 terms (Fig 3A, left). Spatiotemporal maps for changepoint and oddball
290 coefficients were combined to create a *surprise* contrast (changepoint + oddball)
291 and a *learning* contrast (changepoint – oddball) for each subject. Contrasts were
292 aggregated across subjects to create a map of t-statistics (Fig 3A, right), and
293 spatiotemporal clusters of electrode/timepoints exceeding a cluster-forming
294 threshold were tested against a permutation distribution of cluster mass to
295 spatially and temporally organized fluctuations in voltage that related to task
296 variables.

297

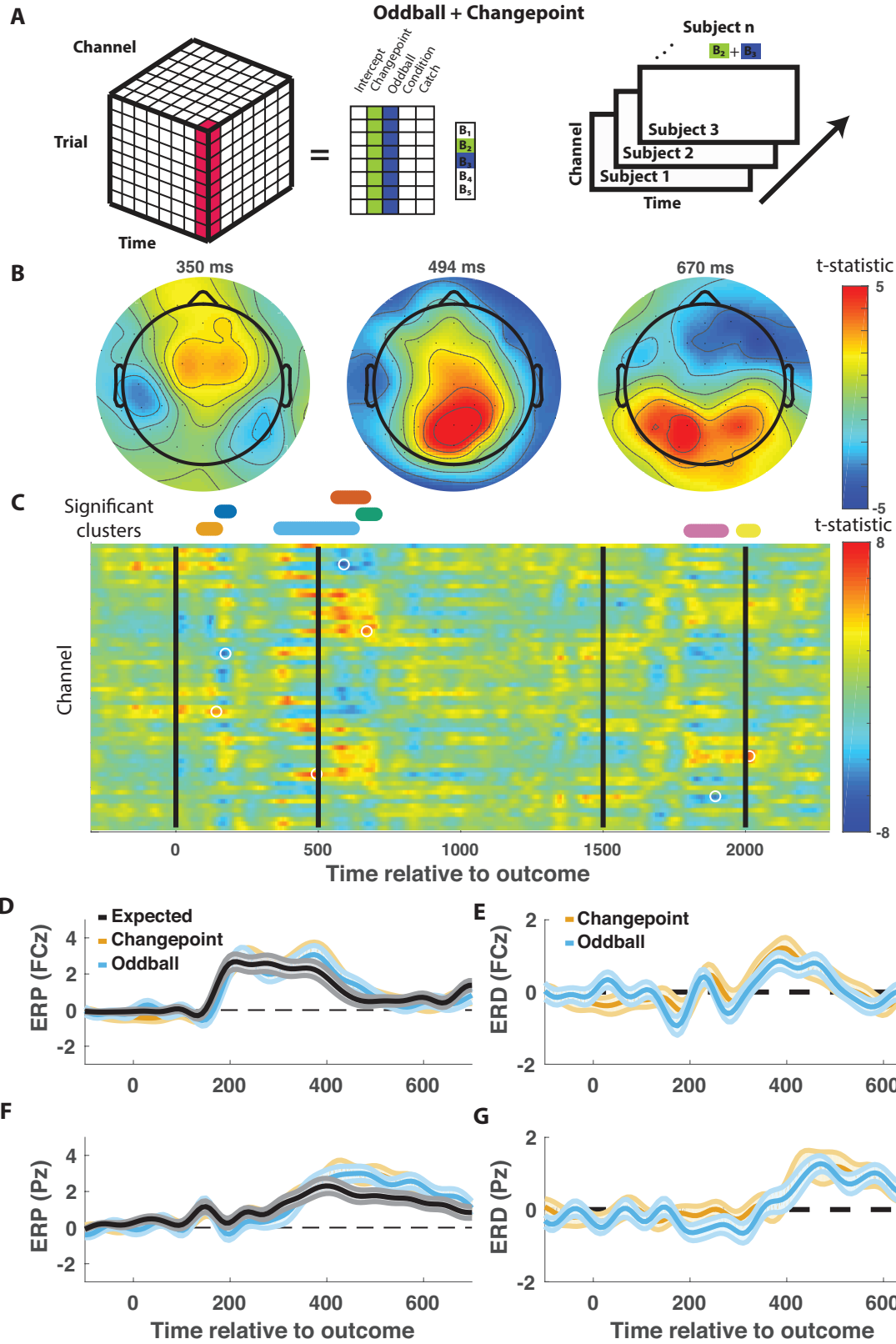
298 When applied to the *surprise* contrast, this procedure yielded a large number of
299 significant clusters distributed across electrodes and timepoints (Fig 3C). Two
300 clusters of positive coefficients occurring 350 to 700 ms after onset of the
301 cannonball location were of particular interest, given the consistency of their
302 timing and direction with the canonical p300 response. Examining the spatial
303 distribution of coefficients during this period reveals an early frontocentral locus
304 of positive coefficients (350 ms; Fig 3B, left) that moves posterior and hits a peak
305 t-statistic at 494 ms (Fig 3b, middle). Later on, the positive coefficients spread
306 laterally and reach a second peak at 670ms (Fig 3B, right). The clustering
307 procedure divides the positive coefficients observed from 350ms to 700ms into
308 two clusters peaking at 494 and 670 ms.

309

310 The time course of positive coefficients within anterior and posterior central
311 electrodes suggests that these clusters are picking up on P3a and P3b
312 components of a P300 response to the outcome delivery. Average outcome-
313 locked event related potentials in a frontocentral electrode (FCz) reveal a positive
314 deflection from 300-500 ms (Fig 3D, black). This deflection is enhanced on both
315 changepoint and oddball trials (Fig 3D,E, orange and blue), reminiscent of the
316 P3a component, also referred to as the novelty P300. Posterior electrode (Pz)
317 event-related potentials (ERPs) reveal a later and longer lasting positive
318 deflection in response to a new outcome (Fig 3F, black). This positive deflection
319 is enhanced on both changepoint and oddball trials (Fig 3F,G, orange and blue),
320 reminiscent of the P3b, or updating component of the P300. Since the spatial and
321 temporal profiles of our clusters were consistent with what has been referred to in
322 previous literature as the P300, we will refer to the clusters peaking at 494 and
323 670 ms as early and late components of the P300, respectively.

324

325 In contrast to the EEG signature of *surprise*, which included a robust and
326 extended P300 response, the only signals identified by the *learning* contrast
327 (change-point-oddball) were early (peak at 158 ms) and transient (Fig S3-1).
328



329
330
331
332

Figure 3: Outcome-locked central positivity reflects surprise irrespective of context. **A)** Trial-series of EEG data for a given electrode and timepoint was regressed onto an explanatory matrix that contained separate binary regressors for changepoint and oddball trials (left). A t-statistic

333 map was created for each electrode and time point on the surprise coefficient contrast (right).
334 **B&C**) T-statistic map for surprise contrast across time (abscissa; **C**) and channel (ordinate; **C**)
335 along with corresponding topoplots **B**. Separate spatiotemporal clusters that survived multiple
336 comparisons correction via permutation testing are depicted in different colors (**C**; above heat
337 plot). The time and channel corresponding to the maximum absolute t-statistic for each such
338 cluster are depicted with a white circle (**C**). **D&F**) Mean/SEM (line/shading) event related
339 potentials (microvolts) sorted by trial type (orange=changepoint, blue=oddball, black=other trials)
340 for frontocentral (**C**; FZc) and central posterior (**F**; Pz) electrodes. **E&G**) Mean/SEM (line/shading)
341 event related difference waveforms computed by subtracting the ERP for typical trials from the
342 average ERP for change-point and oddball trials at frontocentral (**E**; FZc) and central posterior (**G**;
343 Pz) electrodes.

344

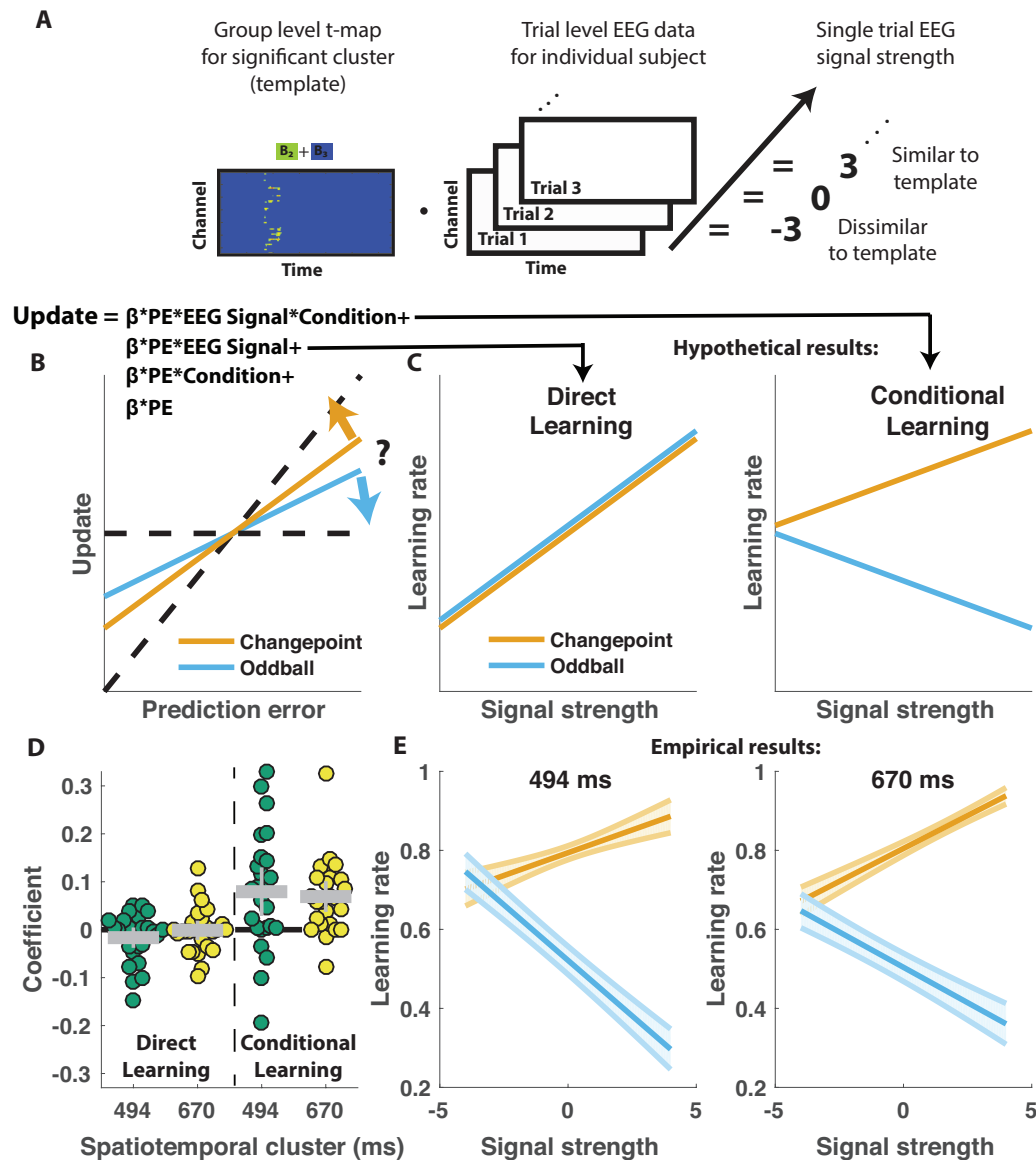
345 *Behavioral relevance of the P300*

346

347 Competing theories posit different functional roles for the signal underlying the
348 P300. In particular, some theories suggest that the P300 reflects a general
349 surprise signal, whereas others attribute a more specific role in accumulating
350 information, for example about the current state of the world. To test how early
351 and late P300 components may relate to learning behavior in our task we
352 extracted trial-to-trial measures of these components by taking the dot product of
353 the cluster t-map and each single trial ERP (Fig 4A, (23)). The dot product
354 indexes the degree to which a single trial ERP displays the profile of a given
355 spatiotemporal cluster, thereby allowing us to test the degree to which the
356 measured signal on any given trial might relate to behavior. We then examined
357 how trial-to-trial behavioral updates in shield position related to these single trial
358 EEG signal strengths using a regression model similar to that employed in the
359 behavioral analysis (Fig 4B). The regression model included two key terms to
360 characterize the influence of 1) the multiplicative interaction of prediction error
361 with the EEG signal strength, and 2) the interaction between prediction error,
362 EEG signal strength and condition. The first EEG-based term provided a
363 measure of the relationship between learning and the P300 that was independent
364 of condition, and thus allowed us to test the prediction that the P300 reflects a
365 *direct learning* signal (Fig 4C). The second EEG-based term provided a measure
366 of the relationship between learning and the P300 that depended on condition
367 (*conditional learning*), and thus allowed us to test the prediction that any learning
368 impact of the P300 is bidirectionally sensitive to the source of surprise (Fig 4D).

369

370



371
 372 **Figure 4:** Central positivity predicts learning in opposite directions for changepoint and oddball
 373 contexts. **A)** T-maps corresponding to significant spatiotemporal clusters were used as templates
 374 to estimate trial-by-trial signal strength. **B)** Single trial updates for each participant were fit with a
 375 regression model that included additional terms to describe 1) the degree to which learning was
 376 increased on trials in which the EEG signal was stronger (PE times EEG signal) as would be
 377 expected for a canonical learning signal and 2) the degree to which learning was conditionally
 378 modulated by the EEG signal (PE times condition times EEG signal) as would be expected for a
 379 surprise signal that influenced downstream learning computations. **C)** Hypothetically, the learning
 380 rate (slope of the relationship between updates and prediction errors) might increase for stronger
 381 EEG signals (left) which would be captured by the PE times EEG direct learning regressor.
 382 Alternatively, the learning rate may increase for stronger EEG signals in the changepoint
 383 condition and decrease for stronger EEG signals in the oddball condition, as measured by the
 384 *conditional learning* regressor. **D)** Individual subject coefficients revealed no significant main
 385 effect of either early (494 ms; green) or late (674 ms; yellow) P300 signals on direct learning
 386 (left), but a strong positive interaction (*conditional learning*) effect at both time points (right),
 387 indicating that the signals were differentially predictive of learning in the changepoint and oddball
 388 conditions. **E)** Learning rates predicted by the regression model (ordinate) increased as a function

389 of signal strength (abscissa) for each P300 cluster (left=494ms; right=674ms) in the changepoint
390 condition (orange) but decreased as a function of signal strength in the oddball condition (blue).

391

392

393 Indeed, participant learning behavior systematically related to trial-by-trial
394 measures of the P300, but only in a manner that depended critically on task
395 condition. *Direct learning* coefficients from the model revealed that neither early
396 (mean/SEM = -0.012/0.01, dof = 24, t = 1.6, p = 0.12) nor late (mean/SEM = -
397 0.001/0.009, dof = 24, t = -0.12, p = 0.90) components of the P300 were
398 systematically related to learning in the same manner across both conditions (Fig
399 4D, left). In contrast, *conditional learning* coefficients from both clusters tended to
400 be positive across subjects (mean/SEM for 494,670ms cluster = 0.07/0.02,
401 0.069/0.02, dof = 24,24, t = 3.2, 4.5, p = 0.004, 0.0002). Learning rate predictions
402 derived from the regression model show that higher P300 signal strength predicts
403 more learning in the changepoint condition (Fig 4E, orange), but less learning in
404 the oddball condition (Fig 4E, blue). Thus, there was a systematic relationship
405 between P300 and learning, but that relationship was oppositely modulated by
406 the task condition and hence the inferred source of surprise.

407

408 The relationship between the P300 and participant learning behavior persisted
409 even after controlling for all known sources of variability in learning behavior. In a
410 model of shield updating behavior that included predictions from the behavioral
411 model described previously (Fig 2E) *conditional learning* coefficients for the late
412 P300 component were reduced relative to the previous regression model, but still
413 greater than zero (Fig 5B, yellow; mean/SEM for 670ms cluster = 0.03/.01 dof =
414 24, t = 2.5, p = 0.02). *Conditional learning* coefficients for the earlier P300 cluster
415 were inconclusive (Fig 5B, green; mean/SEM for 494 ms cluster = 0.02/0.01, dof
416 = 24, t = 1.6, p = 0.11). However, the participants who showed the greatest
417 behavioral modulation of learning according to surprise and condition tended to
418 also have the highest *conditional learning* coefficients indicating the degree to
419 which P300 conditionally predicted learning beyond what could be achieved with
420 our best behavioral model (Fig 5C; r = 0.42, p = 0.04). Thus, the magnitude of the
421 P300 signal predicted learning increases in changepoint contexts and learning
422 decreases in oddball contexts and did so beyond what could be predicted with
423 behavioral modeling alone.

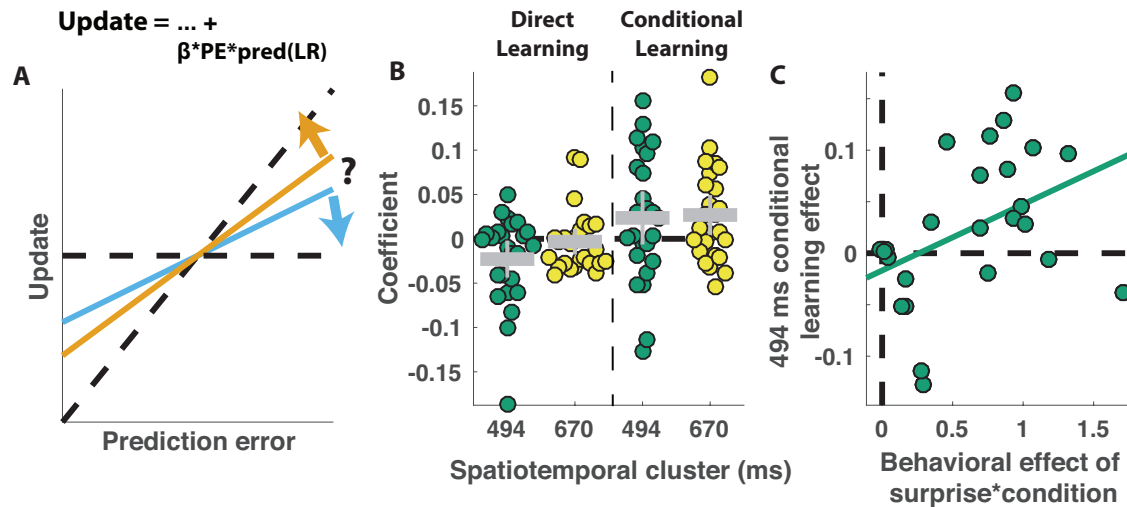
424

425 We applied the same trial-by-trial behavioral analysis to the spatiotemporal
426 clusters identified in our learning (changepoint-oddball) contrast and did not find
427 systematic relationships between EEG signals and learning behavior (ps for all
428 coefficients and spatiotemporal clusters > 0.07; Fig S4-1) even when predictions
429 from our behavioral model (Fig 2E) were not included in the analysis.

430

431

432



433
 434 **Figure 5:** Central positivity explains trial-to-trial learning behavior that could not be otherwise
 435 captured through behavioral modeling. **A)** Single trial updates for each participant were fit with a
 436 regression model that included the best estimates of learning rate provided by our behavioral
 437 regression model (β times PE times pred(LR)) as well as additional terms to describe the degree to
 438 which learning was increased on trials in which an EEG signal was present or the degree to
 439 which learning was contextually modulated by the EEG signal. **B)** Direct learning (left) and
 440 *conditional learning* (right) coefficients for EEG terms in the regression model are plotted for early
 441 (494 ms) and late (670 ms) components of the P300 response for each subject (colored points).
 442 Gray rectangle/lines indicate group mean/SEM. *Conditional learning* coefficients were significantly
 443 greater than zero for the late P300 component (mean/SEM for 670ms cluster = 0.03/.01, dof = 24,
 444 $t = 2.5$, $p = 0.02$) indicating that this signal predicted learning in a contextual manner even after
 445 accounting for behavioral variability that could be captured by our computational model (Figure 2).
 446 **C)** *Conditional learning* coefficients for early P300 component (494 ms; ordinate) were positively
 447 related to the behavioral index of conditional surprise sensitivity (abscissa) as measured through
 448 our behavioral regression ($r = 0.42$, $p = 0.04$).

449
 450
 451
 452
 453
 454
 455
 456
 457
 458
 459
 460
 461
 462
 463
 464
 465
 466
 467

468 Discussion

469

470

471 The brain receives a steady stream of sensory inputs, but these inputs differ
472 dramatically from moment to moment in the degree to which they should affect
473 ongoing inferences about the world. People and animals do not treat each datum
474 in this stream the same, and instead tend to rely more heavily on some pieces of
475 information than others. Identifying the mechanisms through which these
476 adjustments occur could be an important step toward understanding why learning
477 occurs more rapidly in some domains or for some people, yet our understanding
478 of these mechanisms has been heavily conditioned on specific statistical
479 contexts, namely changing environments in which the degree to which one
480 should learn from information is closely coupled to the surprise associated with it.
481 Here we examined how relationships between learning and a specific brain
482 signal, the P300 evoked EEG potential, depend on the statistical context that
483 they are measured in.

484

485 We show that the P300 relates systematically to learning, but that the direction of
486 this relationship depends critically on the statistical context. In a context where
487 surprising events indicated changepoints (Fig 1C,E) and participants learned
488 more from surprising information (Fig 2), larger P300 responses predicted
489 increased learning (Fig 4). In contrast, in a context where surprising events
490 indicated oddballs (Fig 1B,D) and participants deemphasized surprising
491 information (Fig 2), larger P300 responses predicted reduced learning (Fig 4).
492 These context-dependent predictive relationships explained variance in learning
493 beyond what could be captured through computational modeling of behavior
494 alone (Fig 5), suggesting that the P300 signal may be involved in adjustments of
495 learning rate, but does so by mediating the subjective response to surprise,
496 rather than translating surprise into a conditionally appropriate learning signal.

497

498 *Implications for theories of P300 function*

499

500 Our findings are consistent with a number of studies that have suggested the
501 P300 is related to surprise (9,14,17,24), but extend them by demonstrating the
502 role of the signal in controlling the degree to which new information affects
503 updated beliefs. In contrast, our results are inconsistent with standard
504 interpretations of the context updating interpretation of the P300 (17-20). If the
505 P300 signal controlled the degree to which new information was loaded into
506 working memory one would expect a consistent positive relationship between the
507 P300 and learning across conditions (Fig 4C), but our results reveal that this
508 relationship differed markedly depending on the statistical context (Fig 4F,G).

509

510 However, as is the case with many verbal theories, predictions offered by the
511 context updating theory depend critically on the how specific concepts are linked

512 to actual mechanistic processes. If the definition of context were changed to
513 reflect the process that gave rise to the outcome (e.g., normal, changepoint, or
514 oddball), for example, and we assume that participants expected each trial to be
515 normal, then a context updating signal could account for our data (as recognizing
516 more confident recognition of changepoints should lead to more learning, but
517 more confident recognition of oddballs should lead to less learning). Thus, our
518 results constrain potential interpretations of the context updating theory, although
519 they do not falsify the theory altogether.

520

521 Similarly our results could also be viewed as constraining more recent theories
522 about P300 signaling. One more recent theory posits that the event locked
523 central parietal positivity reflects accumulated evidence for a particular decision
524 or course of action (25,26). When accumulated evidence is framed in terms of
525 the action ultimately executed (e.g., shield placement) one might extrapolate to
526 predict that P300 would predict higher learning in both contexts, which is not
527 what we observed (Fig 4F&G). Nonetheless, it is difficult to extrapolate decision
528 variables to our continuous task, and there are other mechanistic schemes in
529 which an evidence accumulation signal over a binary decision categorizing
530 outcome type (normal versus oddball or changepoint) might give rise to our
531 observed results. Such an explanation would also call for response inhibition to
532 prevent premature responding before the default category (e.g., non-oddball trial)
533 was overturned, offering a potential link to another prominent theory of P300
534 function (24,27). Nonetheless, our data do not arbitrate between these theories,
535 and instead highlight their implications for learning when mechanistic
536 interpretations are refined and applied to our task and data.

537

538 *Neural representations of surprise and updating*

539

540 A key question that has motivated a number of recent studies is how does the
541 brain represent surprise differently than the belief updating it sometimes
542 prescribes. Under most conditions, the degree of surprise is tightly linked to the
543 update that is required. However, recent fMRI studies have exploited cued
544 updating paradigms (11), irrelevant stimulus dimensions (28,29), and
545 complementary statistical contexts (4) in order to tease apart neural
546 representations of surprise and updating. While there are trends that seem to
547 generalize across task boundaries (for example, dorsal anterior cingulate cortex
548 (dACC) reflecting updating in cued updating and irrelevant stimulus dimension
549 paradigms (11,29)) there is also a good deal of inconsistency across different
550 tasks in terms of the roles of specific signals. For example, even though BOLD
551 responses in dACC were identified as reflecting updating in two studies, they
552 were shown to represent surprise in another (4) and manipulations of statistical
553 context failed to reveal *any* brain regions that provide a pure updating signal (4).

554

555 One possible explanation for this discrepancy is that the component processes of
556 updating and non-updating might overlap in some specific paradigms. For
557 example, the oddball outcomes that led to reduced learning in our paradigm and
558 that of d'Acremont & Bossaerts were dissimilar to all previous outcomes and
559 indistinguishable on other feature dimensions (in contrast to (11)). Thus, while
560 these outcomes do not contain information pertinent to ongoing beliefs about
561 future outcomes, they did contain information critical for perception, namely that
562 prior expectations should not be used to bias their perceptual representations
563 (30). Interestingly, recent work has suggested that people dynamically adjust the
564 degree to which percepts are biased using systems, including the pupil linked
565 arousal system, that are closely linked to the systems implicated in adjusting
566 learning rate (6,30-34). Thus, one possible explanation for the inconsistency in
567 previous studies attempting to dissociate surprise from updating is that these
568 studies have differed in the degree to which they inadvertently manipulated
569 systems for controlling perceptual biases.

570
571 Like in the previous fMRI study relying on statistical context to dissociate learning
572 from surprise (4), our EEG results revealed a large number of signals related to
573 surprise and no signals that convincingly reflected learning rate in a context
574 independent manner. This comes as somewhat of a surprise given previous work
575 identifying EEG signals analogous to a late P300 component reflecting surprise,
576 predicting learning and influence on choice even in paradigms where this
577 influence was unrelated to surprise (3,9,10,15,16). In line with previous work from
578 fMRI studies, we interpret the differences in our results from what might have
579 been predicted based on previous work as pertaining to unique strategy we
580 employed for dissociating learning from surprise through the use of different
581 statistical contexts.

582
583

584 *Mechanisms of learning rate adjustment*

585

586 Our results, particularly when taken in the context of previous studies examining
587 how the brain adjusts learning in accordance with surprise, constrain possible
588 models of learning rate adjustment in the brain. We show that that the updating
589 P300 signal, which positively predicts learning in changing environments (Fig
590 4E), also negatively predicts learning in a context with infrequent statistical
591 outliers (Fig. 4E). Thus, in a most basic sense, our results suggest that the P300
592 signals reflects an early contribution to learning rate adjustment, and that this
593 signal is untangled according to statistical context at some downstream stage of
594 processing. The lack of robust ERP correlates of direct learning signals (Fig S3-1
595 & S4-1) suggests that this downstream process does not have a task-locked
596 electrophysiological signature.

597

598 One potential mechanism for learning rate adjustment that fits well with these
599 constraints is the notion that adjustments in learning might be implemented
600 through flexible replacement of state representations (35-37). Learning rate
601 adjustment is adaptive in changing environments because it can effectively
602 partition data relevant to the current predictive context from data that are no
603 longer relevant to prediction (21,22). One possible implementation of this
604 partitioning would be to change the active state representations that serve as the
605 substrate for contextual associations. Recent work has identified signals in OFC,
606 a region implicated in representations of latent states (38), that change more
607 rapidly during periods of rapid learning (39). If this is indeed the implementation
608 through which learning rate adjustments occur, observed learning rate signals
609 might actually signal the need to adjust the representation of the latent state.

610
611 Interestingly, replacement of the active latent state, or partitioning of data more
612 generally, might also be an effective way to implement the decreased learning
613 observed in response to surprising observations in the oddball condition of our
614 task. In the case of an oddball, one strategy would be to recognize the oddball as
615 having been generated by an alternative causal process (e.g., oddball
616 distribution) and to attribute learning to a latent representation of this process
617 (40). Under such conditions, implementation would require a surprise signal that
618 reflects the relevance of this oddball latent state. After the new observation is
619 attributed to the oddball context, the system would require a transition back into
620 the original “non-oddball” state in order to make a prediction that is unaffected by
621 the most recent oddball outcome. The more effectively surprise is recognized and
622 responded to through state changes (e.g., the stronger the surprise signal) the
623 more effectively this implementation would partition an oddball observation from
624 ongoing beliefs about the standard generative process, and therefore the smaller
625 learning rates would be. Thus, one mechanistic interpretation of the P300 results
626 might be that it is providing a partitioning signal that results in transitions in the
627 internal state representation, which can either increase or decrease learning
628 depending on the statistical context.

629
630 Confirming our proposed mechanistic interpretation of these results would require
631 future studies more closely relating P300 signals to purported state
632 representations (39). Furthermore, given that our study relied completely on
633 computational modeling and correlations with behavior, our results
634 raise important questions as to whether the observed associations could be
635 manipulated directly pharmacologically or through biofeedback paradigms. Thus,
636 our work provides new insight into the underlying mechanisms of learning rate
637 adjustment and the role of the P300 in this process, but leaves many
638 unanswered questions to be addressed in future research.

639

640

641 **Methods**

642

643 *Participants*

644

645 Participants were recruited from the Brown University community: $n = 37$, 21
646 female, mean age = 20.2 (SD = 3.1, range = 18-36). Behavioral data from all
647 participants was included in behavioral analyses. Data from 12 participants were
648 excluded from EEG analysis due to low data quality (> 25% of epochs rejected
649 during preprocessing). Thus, 37 participants were included in the behavioral
650 analyses and 25 participants were included in the EEG analyses. All human
651 subject procedures were approved by the Brown University Institutional Review
652 Board and conducted in agreement with the Declaration of Helsinki.

653

654 *Cannon Task*

655

656 Participants performed a modified predictive inference task programmed in
657 Matlab (The Mathworks, Natick, MA), using the Psychtoolbox-2
658 (<http://psychtoolbox.org/>) package. The task was based on predictive inference
659 tasks in which participants are asked to predict the next in a series of outcomes
660 (2,6,7), but differed from previous such tasks the following ways: 1) the outcomes
661 were generated from both changepoint and oddball processes to dissociate
662 learning from surprise, 2) information necessary for performance evaluation was
663 not available at time of outcome so that signals related to belief updating could
664 be dissociated from valenced performance evaluation signals, 3) the task space
665 was circular, and 4) the generative process was cast in terms of a cannon
666 shooting cannonballs.

667

668 Participants were instructed to place a shield at some position along a circle
669 subtending 5 degrees of visual angle in order to maximize the chances of
670 catching a cannonball that would be shot on that trial (Fig 1a). During an
671 instructional training period, the generative process that gave rise to cannonball
672 locations was made explicit to participants. During this phase, participants were
673 shown a cannon in the center of the screen. On each trial, a cannonball would be
674 “shot” from that cannon with some angular variability (Von Mises distributed
675 “Noise”, concentration = 10 degrees). A key manipulation in our design was how
676 the aim of the cannon evolved from one trial to the next. The cannon would either
677 1) remain stationary on the majority of trials and re-aim to a random angle with an
678 average hazard rate of 0.14 (changepoint condition) or 2) change position slightly
679 from one trial to the next according to a Von Mises distributed random walk with
680 mean zero and concentration 30 degrees (oddball condition). In the changepoint
681 condition, all cannonballs were displayed as originating at the cannon in the
682 center of the circle, whereas in the oddball condition a small fraction (0.14) of
683 trials were oddballs, in which the cannonball location was sampled uniformly
684 across the entire circle and the cannonball appeared without a trajectory.

685

686 After completing the instructional training, in which the generative process was
687 fully observable, participants were asked to perform the same basic task without
688 being able to see the cannon. In this experimental phase participants were forced
689 to use knowledge of the generative structure gained during training, along with
690 the sequence of prior cannonball locations, in order to infer the aim of the cannon
691 and to inform shield placement. Participants completed four blocks of 60 trials for
692 each task condition (changepoint and oddball) in order randomized across
693 participants. The 240 experimental trials for each condition always followed the
694 instructional training period for that condition in order to minimize ambiguity over
695 which generative structure was giving rise to the experimental outcomes.

696
697 On each trial of the experimental task, participants would adjust the position of
698 the shield through key presses (starting at the shield position from the previous
699 trial) until they were satisfied with its location (Fig 1a; prediction phase). After
700 participants locked in their prediction (through a key press) there was a 500 ms
701 delay and then the cannonball location was revealed for 500 ms (Fig 1A;
702 outcome phase). The cannonball then disappeared for 1000 ms before it
703 reappeared, along with a full depiction of the participants shield (Fig 1A; shield
704 phase). The shield was always centered on the position indicated by the
705 participant during the prediction phase, but differed in size from one trial to the
706 next in a random and unpredictable fashion that ensured subjects could not
707 predict whether they would successfully “catch” the cannonball during the
708 outcome phase. Thus, information provided during the outcome phase provided
709 all necessary information to update beliefs about the cannon aim, but did not
710 contain sufficient information to determine whether the cannonball would be
711 successfully caught on the trial. In addition to trial feedback provided during the
712 shield phase, participants were also provided information about their
713 performance at the end of each block that included the fraction of cannonballs
714 that were caught. Participants were paid an incentive bonus at task completion
715 that was based on the number of cannonballs that were caught.

716

717

718 *Computational Model*

719

720 Optimal inference in the changepoint condition would require considering all
721 possible durations of stable cannon position (21,22) but can be approximated by
722 collapsing the mixture of predictive distributions expected to arise from this
723 optimal solution into a single Gaussian distribution, which approximates the
724 posterior probability distribution over cannon locations, achieves near optimal
725 inference, reduces to an error driven learning rule in which learning rate is
726 adjusted from moment to moment according to environmental statistics, and
727 provides a detailed account of human behavior (2,7). Similarly, the ideal observer
728 for the oddball generative process would require tracking the predictive
729 distributions and posterior probabilities associated with each possible sequence

730 of oddball/non oddball trials that could have preceded the time step of interest.
731 Like in the changepoint condition, this algorithm can be simplified by
732 approximating the set of all possible predictive distributions with a single
733 Gaussian distribution, leading to an error driven learning rule in which learning
734 rate is adjusted dynamically from trial to trial, allowing us to derive normative
735 prescriptions for learning for both conditions.

736
737 While the normative model for the changepoint condition has been described
738 elsewhere (7) the analogous model for the oddball condition is not, and thus we
739 describe the normative account of oddball learning in full detail. In order to
740 minimize the differences between experienced and modeled latent variables, we
741 formulate our model in terms of the prediction errors made by participants on
742 each trial (rather than those that would have been made by the model) (7). On
743 each trial of the oddball condition, the normative model: 1) updated its
744 representation of uncertainty, 2) observed a prediction error and computed the
745 probability that the prediction error reflects an oddball, 3) computed the normative
746 learning rate by combining uncertainty (step 1) and oddball probability (step 2), 4)
747 adjusted prediction about cannon position according learning rate and prediction
748 error.

749
750 Relative uncertainty, which reflects the fraction of uncertainty about an upcoming
751 cannonball location that is due to imperfect knowledge of the cannon aim and is
752 analogous to the Kalman gain, was updated on each trial according to the most
753 recent observation (which should decrease uncertainty about cannon position)
754 and the expected drift in the aim of the cannon occurring between trials (which
755 should increase uncertainty about cannon position). Given that relative
756 uncertainty is expressed as a fraction of total uncertainty, it is useful to think of
757 the numerator of the fraction, or the estimation uncertainty over possible cannon
758 aims, which is the variance on a gaussian mixture distribution and is updated as
759 follows:
760

$$\sigma_{\mu}^2 = \Omega_t \frac{\sigma_N^2 \tau_t}{1 - \tau_t} + (1 - \Omega_t) \sigma_N^2 \tau_t + \Omega_t (1 - \Omega_t) (\delta_t \tau_t)^2 + \sigma_{drift}^2$$

761
762
763 where Ω_t is the probability that an oddball occurred on trial t , σ_N^2 reflects the
764 variance on the distribution of cannonball locations around the true cannon aim
765 (noise), τ_t reflects the relative uncertainty on trial t , δ_t is the prediction error made
766 in predicting the outcome on trial t , and σ_{drift}^2 reflects the degree to which the
767 cannon position drifts from one trial to the next. The first two terms in the model
768 reflect the oddball and non-oddball contributions to the updated uncertainty, the
769 third term reflects uncertainty resulting from the difference between predictions
770 for trial $t+1$ conditioned on an oddball or non-oddball having occurred on trial t ,
771 and the last term reflects uncertainty resulting from the expected drift of the

772 cannon position between trials. Relative uncertainty for trial t+1 is then updated
773 as the updated fraction of uncertainty about the upcoming outcome that is
774 attributable to imprecise knowledge of the true cannon position, rather than to
775 noise in the distribution of exact cannonballs around that position:
776

$$\tau_{t+1} = \frac{\sigma_{\mu}^2}{\sigma_{\mu}^2 + \sigma_N^2}$$

777
778
779
780
781
782

The updated relative uncertainty, along with assumed knowledge of the overall noise and hazard rate, were used to calibrate the oddball probability associated with each new prediction error:

$$\Omega_{t+1} = \frac{\frac{H}{2\pi}}{\frac{H}{2\pi} + (1 - H) \mathcal{N}\left(\delta_{t+1}; 0, \frac{\sigma_N^2}{1 - \tau_{t+1}}\right)}$$

783
784
785
786
787
788
789
790
791
792

Where H is the average hazard of an oddball (0.14) and δ_{t+1} is the new prediction error, and the second term in the denominator reflects the probability density on a normal distribution centered on the predicted location and with variance derived from relative uncertainty. The model's prediction about cannon aim was then updated according to a fraction of the prediction error δ_{t+1} with the exact fraction, or learning rate, determined according to the updated uncertainty and oddball probability:

$$\alpha_{t+1} = \tau_{t+1} - \Omega_{t+1}\tau_{t+1}$$

793
794
795
796
797
798

Note that relative uncertainty (τ_{t+1}) contributes positively to the learning rate, whereas oddball probability (Ω_{t+1}) reduces the learning that would otherwise be dictated by the current level of uncertainty.

799 *Behavioral analysis*

800
801
802
803
804
805
806
807
808
809

Two key behavioral measures were extracted from each trial. First, the *prediction error* on a trial was defined as the circular distance between the cannonball location and the shield position for that trial. Second, the *update* on a given trial was defined as the circular distance between the shield position on that trial and the shield position on the subsequent trial (e.g., the updated shield position). In order to better understand the computational factors governing adjustments in shield position, we fit *updates* with a linear model that included an intercept term to model overall biases in learning along with a prediction error term to capture general tendencies to adjust the shield towards the most recent cannonball

810 location. The model also included additional terms to model how the influence of
811 recent cannonball locations changed dynamically according to task context.
812 These terms included: 1) prediction error times uncertainty interaction (to model
813 how much more participants updated shield position under conditions of
814 uncertainty – as assessed by the computational model), 2) prediction error times
815 surprise (where surprise was indexed by changepoint probability or oddball
816 probability from computational model depending on the context), 3) prediction
817 error times surprise times condition (where condition was +1 for changepoint
818 blocks and -1 for oddball conditions), 4) prediction error times block (a categorical
819 variable indicating whether the shield “blocked” the most recent cannonball. The
820 regression model was fit to each participant and t-tests were performed on the
821 regression coefficients across participants to test for significant contributions of
822 each term to update behavior.

823

824

825 *EEG Acquisition*

826

827 EEG was recorded from a 64-channel Synamps2 system (0.1–100 Hz bandpass;
828 500 Hz sampling rate). Continuous EEG data was epoched with respect to the
829 outcome presentation for each trial. Preprocessing was done manually in Matlab
830 (Mathworks, Natick MA) using the EEGLAB toolbox
831 (<https://sccn.ucsd.edu/eeglab/index.php>) as described previously (23) and
832 included the following steps: 1) epoching and alignment to outcome onset, 2)
833 epoch rejection by inspection, 3) channel removal and interpolation by inspection,
834 4) bandpass filtering [.5-50 hz], 5) removal of blink and eye movement
835 components using ICA. Participants for whom more than 25 percent of epochs
836 were rejected were not included in analyses of EEG data.

837

838

839 *EEG Analysis*

840

841 EEG Data for individual participants were analyzed using a mass univariate
842 approach. Specifically, the trial series EEG data for a given participant, channel,
843 and time relative to outcome onset was regressed onto an explanatory matrix
844 that included the following explanatory variables: 1) intercept, 2) changepoint, 3)
845 oddball, 4) condition, 5) catch . Explanatory variables 2 & 3 were binary variables
846 marking trials in which a surprising event occurred (i.e. changepoint or oddball)
847 whereas 4 reflected the overall task context (i.e. whether oddballs or
848 changepoints were present in the current statistical context), and 5 conveyed
849 whether the participant successfully “caught” the cannonball on each trial.
850 Surprise and learning contrasts were created as the sum and difference of the
851 changepoint and oddball coefficients, respectively. T-statistics were computed
852 across subjects to assess the consistency of contrasts at each electrode and
853 timepoint.

854

855 T-statistic maps were thresholded (cluster forming threshold of $p=0.001$, 2 tailed)
856 and spatiotemporal clusters were identified as temporally and/or spatially
857 contiguous signals that shared a common sign of effect and exceeded the
858 cluster-forming threshold. Cluster mass was computed as the average absolute t-
859 statistic within a cluster times its size (number of electrode timepoints contained
860 within it). Cluster mass for each spatiotemporal cluster was compared to a
861 permutation distribution for cluster mass generated using sign flipping to correct
862 for multiple comparisons (41).

863

864 Trial-to-trial EEG analyses were conducted by computing the dot product of the t-
865 statistic map for a given spatiotemporal cluster and the ERP measured on a
866 given trial. The resulting measure of EEG signal strength was then z-scored
867 across all trials and included in a behavioral regression model to explain trial-to-
868 trial updating behavior. Like for the behavioral analyses, trial-to-trial updates were
869 regressed onto an explanatory matrix that included intercept and prediction error
870 terms to capture updating biases and static tendencies to update toward recent
871 cannonball locations. In addition, EEG informed regression models included 1)
872 the interaction between the EEG signal strength computed above and prediction
873 error (direct learning), and 2) the three-way interaction between EEG signal
874 strength, prediction error, and condition (conditional learning). Positive *direct*
875 *learning* coefficients indicated an unconditional increase in learning for trials in
876 which EEG signal strength was greater, whereas positive conditional learning
877 coefficients indicated a positive relationship between EEG signal strength and
878 learning in the changepoint condition but a negative relationship between EEG
879 signal strength in the oddball condition. In order to test the degree to which EEG-
880 updating relationships persisted after accounting for variability in behavior that
881 could be captured by our computational model, we also used a version of the
882 EEG informed regression that additionally included the predicted update from the
883 behavioral model (\hat{y}) as an explanatory variable (Fig 5).

884

885

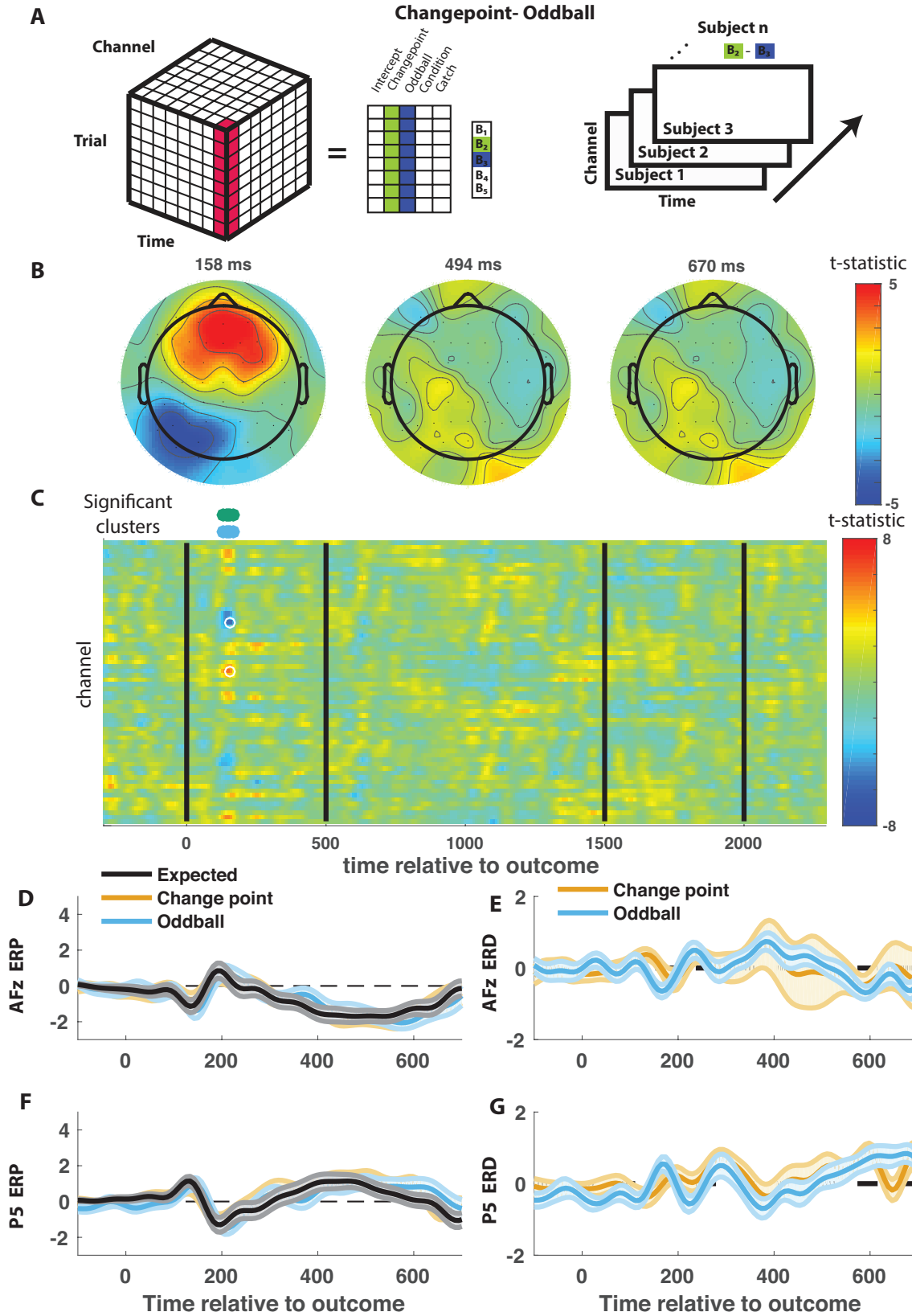
886

887

888

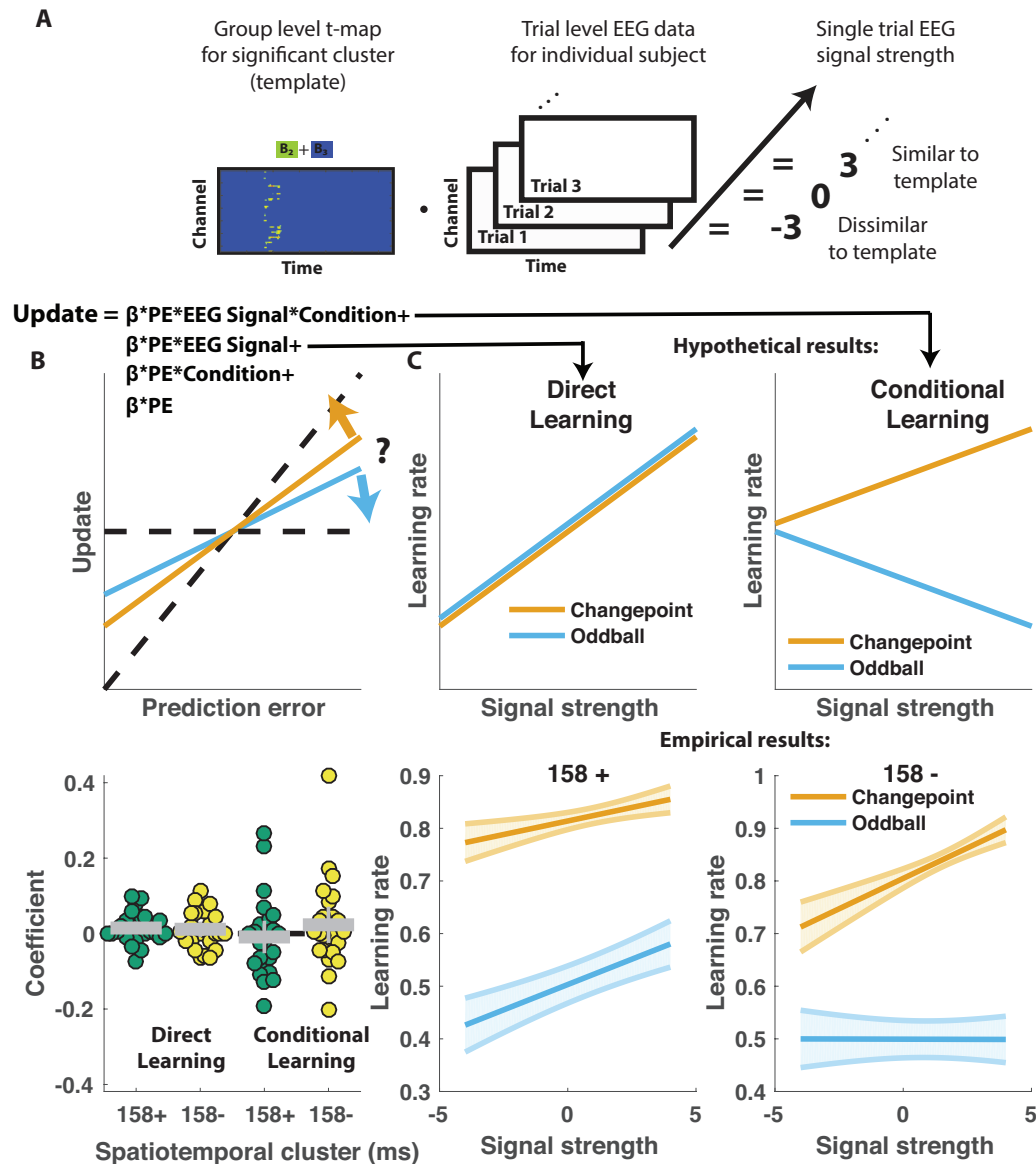
889

890



891
892

893 Figure S3-1: Early frontal, but not late central, positivity is greater in changepoint than oddball trials.
 894 **A)** Trial-series of EEG data for a given electrode and timepoint was regressed onto an
 895 explanatory matrix that contained separate binary regressors for changepoint and oddball trials
 896 (left). A t-statistic map was created for each electrode and time point on the learning contrast
 897 (right). **B&C)** T-statistic map for learning contrast across time (abscissa; **C**) and channel
 898 (ordinate; **C**) along with corresponding topoplots (**B**). Separate spatiotemporal clusters that
 899 survived multiple comparisons correction via permutation testing are depicted in different colors
 900 (**C**; above heat plot). **D&F)** Mean/SEM (line/shading) event related potentials (microvolts) sorted
 901 by trial type (orange=changepoint, blue=oddball, black=other trials) for frontal (**C**; AFz) and
 902 posterior (**F**; P5) electrodes that distinguish between changepoints and oddballs. **E&G)**
 903 Mean/SEM (line/shading) event related difference waveforms computed by subtracting the ERP
 904 for typical trials from the average ERP for change-point and oddball trials at frontocentral (**E**; AFz)
 905 and central posterior (**G**; P5) electrodes.
 906
 907
 908



909

910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955

Figure S4-1: Weak direct relationships between EEG signals and learning. **A)** T-maps corresponding to significant spatiotemporal clusters were used as templates to estimate trial-by-trial signal strength (as the cosine product of the template with the outcome locked EEG data recorded on each trial). **B)** Single trial updates for each participant were fit with a regression model that included additional terms to describe 1) the degree to which learning was increased on trials in which the EEG signal was stronger (PE times EEG signal) as would be expected for a canonical learning signal and 2) the degree to which learning was contextually modulated by the EEG signal (PE times condition times EEG signal) as would be expected for a surprise signal that influenced downstream learning computations. **C)** Hypothetically, the learning rate (slope of the relationship between updates and prediction errors) might increase for stronger EEG signals (left) which would be captured by the (PE times EEG signal) regressor (*direct learning*). Alternatively, the learning rate may increase for stronger EEG signals in the changepoint condition and decrease for stronger EEG signals in the oddball condition, consistent with *conditional learning* signal. **D)** Individual subject coefficients from the changepoint-oddball spatiotemporal clusters (figure S3-1) revealed no systematic *direct learning* effect of either positive (158 +; green) or negative (158 -; yellow) EEG signals (left) (mean[SEM] coefficient = 0.01[0.008], 0.01[0.01], $p = 0.08, 0.25$). Nor did coefficients reveal *conditional learning* effect (right) (mean[SEM] coefficient = -0.01[0.02], 0.02[0.02], $p = 0.67, 0.32$). **E)** Learning rates predicted by the regression model (ordinate) tended to increase as a function of signal strength (abscissa) for both clusters (left=158+; right=158-) in the changepoint condition (orange) but were less consistent in the oddball condition (blue).

1. Behrens TEJ, Woolrich MW, Walton ME, Rushworth MFS. Learning the value of information in an uncertain world. *Nature Neuroscience*. 2007 Sep;10(9):1214–21.
2. Nassar MR, Wilson RC, Heasly B, Gold JI. An approximately Bayesian delta-rule model explains the dynamics of belief updating in a changing environment. *Journal of Neuroscience*. 2010 Sep 15;30(37):12366–78.
3. Cheadle S, Wyart V, Tsetsos K, Myers N, de Gardelle V, Castañón SH, et al. Adaptive Gain Control during Human Perceptual Choice. *Neuron*. Elsevier Inc; 2014 Mar 19;81(6):1429–41.
4. d'Acremont M, Bossaerts P. Neural Mechanisms Behind Identification of Leptokurtic Noise and Adaptive Behavioral Response. *Cerebral Cortex*. 2016 Feb 4;:bhw013.
5. Diederer KMJ, Spencer T, Vestergaard MD, Fletcher PC, Schultz W. Adaptive Prediction Error Coding in the Human Midbrain and Striatum Facilitates Behavioral Adaptation and Learning Efficiency. *Neuron*. Elsevier Inc; 2016 Jun 1;90(5):1127–38.

- 956 6. Nassar MR, Rumsey KM, Wilson RC, Parikh K, Heasley B, Gold JI. Rational
957 regulation of learning dynamics by pupil-linked arousal systems. *Nature*
958 *Neuroscience*. 2012 Jul;15(7):1040–6.
- 959 7. Nassar MR, Bruckner R, Gold JI, Li S-C, Heekeren HR, Eppinger B. Age
960 differences in learning emerge from an insufficient representation of
961 uncertainty in older adults. *Nature Communications*. 2016;7:11609.
- 962 8. McGuire JT, Nassar MR, Gold JI, Kable JW. Functionally dissociable
963 influences on learning rate in a dynamic environment. *Neuron*. 2014 Nov
964 19;84(4):870–81.
- 965 9. Jepma M, Murphy PR, Nassar MR, Rangel-Gomez M, Meeter M,
966 Nieuwenhuis S. Catecholaminergic Regulation of Learning Rate in a
967 Dynamic Environment. O'Reilly JX, editor. *PLoS Comput Biol*. 2016 Oct
968 28;12(10):e1005171.
- 969 10. Jepma M, Brown SBRE, Murphy PR, Koelewijn SC, de Vries B, van den
970 Maagdenberg AM, et al. Noradrenergic and Cholinergic Modulation of
971 Belief Updating. *J Cogn Neurosci*. 2018 Jul 31;:1–18.
- 972 11. O'Reilly JX, Schüffelgen U, Cuell SF, Behrens TEJ, Mars RB, Rushworth
973 MFS. Dissociable effects of surprise and model update in parietal and
974 anterior cingulate cortex. *Proceedings of the ...* 2013.
- 975 12. Iglesias S, Mathys C, Brodersen KH, Kasper L, Piccirelli M, Ouden den
976 HEM, et al. Hierarchical Prediction Errors in Midbrain and Basal Forebrain
977 during Sensory Learning. *Neuron*. Elsevier Inc; 2013 Oct 16;80(2):519–30.
- 978 13. Summerfield C, Tsetsos K. Do humans make good decisions? *Trends in*
979 *Cognitive Sciences*. Elsevier Ltd; 2015 Jan 1;19(1):27–34.
- 980 14. Garrido MI, Teng CLJ, Taylor JA, Rowe EG, Mattingley JB. Surprise
981 responses in the human brain demonstrate statistical learning under high
982 concurrent cognitive demand. *npj Science Learn*. 2016 Jun
983 8;1(1):e1002999.
- 984 15. Fischer AG, Ullsperger M. Real and Fictive Outcomes Are Processed
985 Differently but Converge on a Common Adaptive Mechanism. *Neuron*.
986 2013 Sep;79(6):1243–55.
- 987 16. Wyart V, de Gardelle V, Scholl J, Summerfield C. Rhythmic Fluctuations in
988 Evidence Accumulation during Decision Making in the Human Brain.
989 *Neuron*. Elsevier Inc; 2012 Nov 21;76(4):847–58.
- 990 17. Donchin E. Presidential address, 1980. Surprise!...Surprise?

- 991 Psychophysiology. 1981 Sep;18(5):493–513.
- 992 18. Donchin E, Coles MGH. Is the P300 component a manifestation of context
993 updating? Behav Brain Sci. 2010 Feb 4;11(03):357.
- 994 19. Polich J. THEORETICAL OVERVIEW OF P3a AND P3b. 2003 Jan 24;:1–
995 2.
- 996 20. Polich J. Updating P300: An integrative theory of P3a and P3b. Clinical
997 Neurophysiology. 2007 Oct;118(10):2128–48.
- 998 21. Prescott Adams R, MacKay DJC. Bayesian Online Changepoint Detection.
999 eprint arXiv:07103742. 2007 Oct 1;:–.
- 1000 22. Wilson RC, Nassar MR, Gold JI. Bayesian online learning of the hazard
1001 rate in change-point problems. Neural Comput. 2010 Sep 1;22(9):2452–76.
- 1002 23. Collins AGE, Frank MJ. Within- and across-trial dynamics of human EEG
1003 reveal cooperative interplay between reinforcement learning and working
1004 memory. Proceedings of the National Academy of Sciences. 2018 Mar
1005 6;115(10):2502–7.
- 1006 24. Wessel JR. A Neural Mechanism for Surprise-related Interruptions of
1007 Visuospatial Working Memory. Cerebral Cortex. 2016 Dec 1;28(1):199–
1008 212.
- 1009 25. Kelly SP, O'Connell RG. Internal and external influences on the rate of
1010 sensory evidence accumulation in the human brain. Journal of
1011 Neuroscience. 2013 Dec 11;33(50):19434–41.
- 1012 26. O'Connell RG, Dockree PM, Kelly SP. A supramodal accumulation-to-
1013 bound signal that determines perceptual decisions in humans. Nature
1014 Neuroscience. 2012 Oct 28;15(12):1729–35.
- 1015 27. Wessel JR, Aron AR. Perspective. Neuron. Elsevier Inc; 2017 Jan
1016 18;93(2):259–80.
- 1017 28. Schwartenbeck P, FitzGerald THB, Dolan R. Neural signals encoding shifts
1018 in beliefs. NeuroImage. The Authors; 2016 Jan 15;125(C):578–86.
- 1019 29. Nour MM, Dahoun T, Schwartenbeck P, Adams RA, FitzGerald THB,
1020 Coello C, et al. Dopaminergic basis for signaling belief updates, but not
1021 surprise, and the link to paranoia. Proceedings of the National Academy of
1022 Sciences. 2018 Oct 23;115(43):E10167–76.
- 1023 30. Krishnamurthy K, Nassar MR, Sarode S, Gold JI. Arousal-related
1024 adjustments of perceptual biases optimize perception in dynamic

- 1025 environments. *Nat hum behav.* 2017 May 8;1:0107.
- 1026 31. Nieuwenhuis S, De Geus EJ, Aston-Jones G. The anatomical and
1027 functional relationship between the P3 and autonomic components of the
1028 orienting response. *Psychophysiology.* 2011 Jan 6;48(2):162–75.
- 1029 32. Vazey EM, Moorman DE, Aston-Jones G. Phasic locus coeruleus activity
1030 regulates cortical encoding of salience information. *Proceedings of the*
1031 *National Academy of Sciences.* 2018 Oct 2;115(40):E9439–48.
- 1032 33. Urai AE, Braun A, Donner TH. Pupil-linked arousal is driven by decision
1033 uncertainty and alters serial choice bias. *Nature Communications.* 2017
1034 Mar 3;8:14637.
- 1035 34. de Gee JW, Colizoli O, Kloosterman NA, Knapen T, Nieuwenhuis S,
1036 Donner TH. Dynamic modulation of decision biases by brainstem arousal
1037 systems. *eLife.* 2017 Apr 11;6.
- 1038 35. Collins A, Koechlin E. Reasoning, Learning, and Creativity: Frontal Lobe
1039 Function and Human Decision-Making. O'Doherty JP, editor. *PLoS Biol.*
1040 2012 Mar 27;10(3):e1001293.
- 1041 36. Collins AGE, Frank MJ. Cognitive control over learning: creating,
1042 clustering, and generalizing task-set structure. *Psychological Review.* 2013
1043 Jan;120(1):190–229.
- 1044 37. Wilson RC, Takahashi YK, Schoenbaum G, Niv Y. Orbitofrontal cortex as a
1045 cognitive map of task space. *Neuron.* 2014 Jan 22;81(2):267–79.
- 1046 38. Schuck NW, Cai MB, Wilson RC, Niv Y. Human Orbitofrontal Cortex
1047 Represents a Cognitive Map of State Space. *Neuron.* Elsevier Inc; 2016
1048 Sep 21;91(6):1402–12.
- 1049 39. Nassar MR, McGuire JT, Ritz H, Kable JW. Dissociable Forms of
1050 Uncertainty-Driven Representational Change Across the Human Brain.
1051 *Journal of Neuroscience.* 2019 Feb 27;39(9):1688–98.
- 1052 40. Gershman SJ, Niv Y. Learning latent structure: carving nature at its joints.
1053 *Current Opinion in Neurobiology.* 2010 Apr;20(2):251–6.
- 1054 41. Nichols TE, Holmes AP. Nonparametric permutation tests for functional
1055 neuroimaging: a primer with examples. *Hum Brain Mapp.* 2002
1056 Jan;15(1):1–25.
- 1057