

Supplementary information

Decoding and perturbing decision states in real time

In the format provided by the authors and unedited

1 **Supplementary Information**

2 **Methods**

3 **1. Subjects**

4 Our experiments were performed on two adult male macaque monkeys (*Macaca mulatta*, ages 8-
5 10 and 12-14 years old over the course of the experiments for monkeys H and F, respectively).
6 Monkeys were trained to perform a direction discrimination task with reaching movements of the
7 arm as operant responses. These were the same subjects used in our previous study¹¹, but with new
8 experiments. All training, surgery, and recording procedures conformed to the National Institutes
9 of Health Guide for the Care and Use of Laboratory Animals and were approved by Stanford
10 University Animal Care and Use Committee.

11 **2. Apparatus**

12 Monkeys sat in a custom-made primate chair (Stanford Machine Shop) in front of a video
13 touchscreen, with their heads restrained using a surgical implant. The front plate of the chair could
14 be opened, allowing the subjects to reach the touchscreen with the arm contralateral to the
15 implanted hemisphere. The ipsilateral arm was gently restrained using a Delrin tube and a cloth
16 sling. Stimuli were shown on the video touchscreen (ELO Touchsystems 1939L), which was
17 positioned approximately 35.5 cm away from the monkeys' heads and allowed hand position to be
18 tracked at 75 Hz. Eye position was continuously tracked with an infrared eye tracker at 1 kHz
19 (EyeLink 1000, SR Research, Canada).

20 3. Motion discrimination task

21 The task employed is a variation of the classical random dots motion discrimination task, in which
22 the subject uses an arm movement as the operant response¹¹ (Fig. 1a). We used a variable duration
23 version of this task in which the duration of the stimulus presentation varied from trial to trial.
24 There were two types of trials in our experiments: *open-loop*, in which the stimulus duration was
25 determined by the experimenter at the beginning of the trial and *closed-loop*, in which the duration
26 was contingent on a specific pattern of neural activity detected in real time (see Experiments 1-3).
27 The subject was never cued on what type of trial it was on. For open-loop trials stimulus duration
28 ranged from 500-1200 ms (median 670 ms) and was randomly chosen on each trial by sampling
29 an exponential distribution. For closed-loop trials the possible values for duration ranged between
30 250-1200 ms and were determined on each trial either by the timepoint at which the termination
31 conditions were met or a predetermined random duration sampled from the open-loop distribution,
32 whichever came first. All trials started with the onset of a fixation point (FP; 1.5 degree diameter)
33 on the video touchscreen (Fig. 1a). To initiate the task, the monkey was required to maintain both
34 eye and hand fixation within ± 3 degrees of the FP as long as it remained on the screen. Importantly,
35 throughout the entire trial, the monkey was required to always maintain direct hand contact with
36 the screen, otherwise the trial would be aborted.

37 After 300 ms of fixation, two targets (1.5 degree diameter) appeared on opposite sides of the FP
38 (eccentricities between 10 and 17 degrees). After a 500 ms delay the random dot stimulus was
39 presented for the durations mentioned above, after which it was removed from the screen. The
40 monkey was asked to report the net direction of motion (0 or 180 degrees) by reaching to the target
41 in the corresponding direction. The difficulty of the task was adjusted by changing the fraction of

42 dots moving coherently in one direction (motion strength). After stimulus offset the monkey either
43 entered a delay period during which it was required to withhold his response for 400-900 ms (on
44 30% of the open-loop trials) or was immediately presented the go cue (on 70% of the open-loop
45 trials and all closed-loop trials). The go cue was then signaled by the offset of the FP at which
46 point the monkey was free to gaze anywhere and report his decision with his arm by reaching one
47 of the two targets. Although gaze was monitored, reward acquisition depended solely on reaching
48 to the correct target. Finally, for a response to be considered valid, the monkey was required to
49 hold its hand position within ± 4 degrees of the center of the target for 200 ms. The monkey was
50 then rewarded with a drop of juice for correct choices and given a timeout (2-4 seconds) for
51 incorrect ones. Zero coherence trials were rewarded randomly with a probability of 0.5 since there
52 was no correct response on these trials. The motion discrimination task was run on an Apple Mac
53 Pro running Mac OS.

54 **4. Random dots stimuli**

55 The stimuli used in our psychophysical experiment were random dot kinematograms (RDK)
56 generated using MATLAB and Psychophysics Toolbox. The details for generating the random
57 dots stimuli have been described previously¹¹. However, to allow for closed-loop experiments 1
58 and 2 (see below) we introduced a modification to be able to terminate the dot presentations early
59 if needed. The stimulus code was designed to precompute a sequence of kinematograms that
60 contain both random and moving dots. The sequence was then presented ballistically with no need
61 to continuously compute the content of each frame. Our modification allowed for DV values to be
62 received asynchronously from the real-time decoder and evaluated during the dots presentation. If
63 the DV criteria defined by the particular experiment were met, the dot presentation could then be

64 terminated without the remaining frames being shown. For the experiment in which an additional
65 pulse of motion energy was injected (closed-loop experiment 3, see below), we arranged for two
66 sequences of kinematograms to be precomputed before presentation: one without the pulse, the
67 other for the 200 ms pulse itself. Contingent on the evolution of DV values, the stimulus could
68 then be rapidly switched from the standard sequence to the pulse sequence.

69 For both monkeys, the motion strength could take one of 6 possible values within a set, but the
70 sets were slightly different between subjects: [0%, 1.6%, 3.2%, 6.4%, 12.8%, 25.6%] for monkey
71 H and [0%, 3.2%, 6.4%, 12.8%, 25.6%, 51.2%] for monkey F. The top coherence (51.2%) was
72 dropped and a very low coherence (1.6%) was introduced for monkey H, due to its superior
73 discrimination ability. For pooled analyses across the two subjects, we grouped trials into relative
74 coherence levels as follows: level 0: 0%/0% level 1: 1.6%/3.2%, level 2: 3.2%/6.4%, level 3:
75 6.4%/12.8%, level 4: 12.8%/ 25.6%, and level 5: 25.6%/51.2% for monkey H/F.

76 The direction and coherence of the motion were randomly assigned on each trial by sampling from
77 a uniform distribution with replacement. For zero-coherence stimuli all dots were displaced
78 randomly but, due to the stochasticity of that process, one obtains non-zero net motion toward the
79 targets over a small number of frames.

80 **5. Behavioral training**

81 Both monkeys had been extensively trained on fixed and variable duration versions of the motion
82 discrimination task using an arm reach movement as the operant response prior to the current
83 study¹¹. A few training sessions (all open-loop trials) were used to get the subject accustomed to

84 the new task timing (0.5-1.2 s stimuli and no delay on 70% of the trials). Real time decoding
85 sessions only started when psychophysical performance was stable.

86 **6. Behavioral analysis**

87 Psychophysical performance was assessed in two ways: by describing the percentage of correct
88 choices as a function of (unsigned) stimulus coherence and by describing the percentage of
89 rightward choices as a function of signed stimulus coherence.

90 The percentage of correct choices as a function of motion strength (stimulus coherence) was fit by
91 a cumulative Weibull distribution function:

$$P_{\text{correct}}(c) = 1 - 0.5 \times e^{(-\frac{c}{\alpha})^\beta} \quad \text{Eqn. 1}$$

92 where P_{correct} is probability correct, c is motion strength, α is the psychophysical threshold (the
93 value of c that corresponds to ~82% correct responses), and β is a parameter that controls the shape
94 of the function, especially its steepness. For behavioral analyses shown in Extended Data Fig. 1a,
95 Eqn. 1 was fit separately for trials in each stimulus duration quartile for each subject (monkey H
96 quartiles: Q1: [0.500 , 0.574] s, Q2: [0.574 , 0.680] s, Q3[0.680 , 0.827] s, Q4: [0.827, 1.200] s;
97 monkey F quartiles: Q1: [0.500 , 0.574] s, Q2: [0.574 , 0.667] s, Q3[0.667 , 0.813] s, Q4: [0. 813,
98 1.200] s).

99 The proportion of rightward choices, P_{right} , as a function of motion strength and direction was fit
100 by a logistic regression:

$$P_{\text{right}}(c) = \frac{1}{1 + e^{-\beta_1 \times (\beta_0 + c)}} \quad \text{Eqn. 2}$$

101 where c is motion strength, β_1 is the slope parameter and $-\beta_0$ is the motion strength corresponding
102 to the indifference point. This value was used to assess the monkey's behavioral bias on each
103 session.

104 **7. Electrophysiological recordings**

105 Two multielectrode arrays (Blackrock Microsystems, Utah) with 96 electrodes each (1mm long
106 platinum-iridium electrodes, 0.4 mm spacing, impedance average of approximately 400 KOhm)
107 were implanted in primary motor and dorsal premotor cortex of each monkey (Fig. 1b). The
108 methods for determining the array placement were described in our previous study¹¹. For monkey
109 F, the M1 array became unusable between the end of the previous study and the start of the current
110 study. Due to lack of neural signal from the M1 array, only the PMd array was used for this animal.
111 Continuous neural data were acquired and saved to disk from each channel (sampling rate 30 kHz)
112 and thresholded at -4.5 RMS using the Cerebus recording system (Blackrock Microsystems, Utah)
113 and two separate PCs (one for each array) running Windows 8. Waveforms corresponding to
114 threshold crossings were not sorted and each channel could contain one or more unit(s). Sorting
115 waveforms would require a significant lead-up time before the beginning of the experiment and
116 could negatively affect the ability to combine data and use decoders across days (see below,
117 Decoder training). Since units were not isolated within each channel our resulting units were most
118 likely multi-unit clusters. Any extremely noisy channels were deactivated at the beginning of a
119 session, and all other channels were used in this study.

120 **8. Datasets**

121 Data were collected in two sets of experiments. In the first set of experiments we performed closed-
122 loop experiments 1 and 2 (see below). For this set, for each monkey we analyzed all datasets that
123 met two behavioral inclusion criteria: 1) over 500 trials and 2) a behavioral bias ($|\beta_0|$) under 4%,
124 as determined by a logistic regression fit (see above). These criteria were imposed to ensure that
125 we have a sizeable number of trials per condition (6 coherence x 2 directions = 12 conditions) and
126 that the behavior of the monkey is virtually unbiased, such that both neural and behavioral results
127 are more easily interpretable. These criteria resulted in a selection of 17/17 and 15/20 sessions for
128 a total of 16468 and 15826 trials for monkey H and F, respectively.

129 In the second set of experiments we performed closed-loop experiment 3. This set of experiments
130 was performed later, on separate sessions, but using the same two subjects, arrays and decoding
131 techniques as the first set. In this set of experiments, we analyzed all datasets with over 550 trials.
132 These criteria resulted in a selection of 48/48 and 32/33 sessions for monkey H and F, respectively.
133 For all experiments in monkey H, PMd and M1 were recorded simultaneously.

134 **9. Decoder training**

135 We chose to use a logistic regression classifier based on our preliminary previous results showing
136 excellent offline prediction accuracy in variable duration tasks¹¹ and because of the direct
137 probabilistic interpretation of its output. Our decision variable (DV) was defined as the log odds
138 ratio of observing a particular behavioral choice (T_1 : rightward choice or T_2 : leftward choice) given
139 the population response \vec{r} :

$$DV = \log \frac{P(T_1|\vec{r})}{P(T_2|\vec{r})} = \beta_0(t) + \sum_{i=1}^n \beta_i(t) \times r_i(t) \quad \text{Eqn. 3}$$

140 Where $r_i(t)$ are the z-scored summed spike counts for each neuron and time window, β_0 is an
 141 intercept term and $\beta_i(t)$ are the classifier weights (one for each unit and epoch). Data from all
 142 electrodes with valid waveforms were combined.

143 For simplicity, we decided to use only 3 different decoders for an entire trial (Fig. 1a), instead of
 144 a different one for each 50 ms time window in the trial¹¹. We applied the first decoder from fixation
 145 up to and including the dots period, the second for the delay period and the third for the post go
 146 cue period. After extensive offline tests on a few sessions, the precise epochs for classifier training
 147 were defined as the following:

- 148 • Dots epoch: from 150 to 1000 ms after dots onset;
- 149 • Delay epoch: from 250 to 350 ms after dots offset;
- 150 • Post-go cue epoch: from 200 to 400 ms after onset of the go cue.

151 LASSO regularization was applied to prevent over-fitting when calculating each set of β weights.
 152 A Lambda parameter constraining the L1 norm of the β vectors was calculated separately for each
 153 of the 3 decoders using 10-fold cross validation on the corresponding time epochs listed above.
 154 For each decoder the Lambda value with minimum cross-validation error was chosen. Extended
 155 Data Fig. 6b shows beta weights for an example set of 3 decoders for monkey H sorted by epoch

156 and ranked by magnitude. Positive weights correspond to rightward preferring channels while
157 negative weights correspond to leftward preferring channels. LASSO regularization sets weights
158 of channels with little or no predictive activity to zero. On average, 56% (78%) of the channels
159 were included in the dots epoch decoder, 34% (65%) in the delay epoch, and 44% (74%) in the
160 post-go epoch for monkey H (F) (Supplementary Table 5).

161 The linear classifier was determined offline using recently collected data (from real-time
162 experiments). All 50 ms samples of neural data during the selected period (above) for each epoch
163 were used to train the classifier. The classifier was trained on 90% of the trials and tested on 10%
164 of the trials using 10-fold cross-validation. The weights from one of the cross-validation folds were
165 then used in the upcoming real-time experiments. Decisions to train new decoders were based on
166 experimenter judgment in attempts to optimize performance: if a substantial decrease in real-time
167 decoding performance and/or an increase in the DV offset at baseline was observed, a new
168 classifier was trained and used in the following session. New classifiers were typically used every
169 5 sessions, but some proved to be stable over up to 14 sessions (Extended Data Fig. 6a).

170 **10. Real-time decoding**

171 An essential requirement to compute a real-time read-out of neural activity is the ability to
172 continuously and (nearly) instantaneously access and perform computations on the neural activity
173 being recorded. To accomplish this, the spikes for each channel were temporally smoothed using
174 a causal half-Gaussian kernel with 50 ms standard deviation (to mitigate spurious Poisson
175 fluctuations) and summed for the most recent 50 ms. These smooth spike counts were then stored
176 in a 192x1 (96x1 for monkey F) vector of neural activity and z-scored individually for each

177 channel, using previously calculated μ (mean) and σ (standard deviation) vectors. Z-scoring neural
178 activity was crucial to ensure a reliable and stable real-time readout by preventing the highest firing
179 channels from dominating it. Finally, the z-scored neural activity was projected onto a previously
180 calculated linear decoder (a set of β weights, one for each channel) to obtain our linear readout of
181 internal decision state: a real time decision variable (DV)².

182 The value of the DV was updated every 10 ms, reflecting the neural activity of the preceding 50
183 ms. Because we used a half-gaussian kernel, data preceding the 50 ms window also influenced our
184 DV estimate (with more recent spikes carrying more weight). 95% of the data contributing to the
185 spike counts was limited to the last 100 ms (i.e an additional 50 ms in the past to each 50 ms
186 window). The sign of the DV was used to predict the upcoming behavioral choice: positive for
187 rightward choices and negative for leftward choices. Prediction accuracy was calculated for each
188 time point using the real-time DV and quantified as the fraction of trials in which the classifier
189 correctly predicted the monkey's upcoming choice. Baseline prediction accuracy was defined as
190 the average accuracy for the first 80 ms of the random dots stimulus, and latency as the first of 6
191 consecutive 10 ms time bins with a prediction accuracy value above baseline, Wilcoxon Sign Rank
192 test $p < 0.001$.

193 In addition to using the DV sign to predict choices, the DV value and its history on a single trial
194 could be used (if desired) to impose conditions for termination of the random dots stimulus
195 (experiments 1 and 2) or presentation of a motion pulse (experiment 3), effectively closing the
196 loop on the experiment.

197 While the β weights were not updated online (during the course of one experiment), the μ and σ

198 vectors for each epoch were learned continuously during the course of the experiment, due to
199 changing recording conditions and signals from day to day. The μ and σ vectors were initialized
200 at the beginning of the session using the values calculated offline when training the most recent
201 decoder. Once the session started, the initial μ and σ vectors were blended with online calculated
202 values for the first 25 trials, using a blending factor α :

$$\alpha_j = \max((25 - j)/25, 0), \text{ where } j \text{ is the trial number.} \quad \text{Eqn. 4}$$

203 For trial j , sample number t and for a given epoch in trial, the μ and σ vectors were defined as a
204 weighted mixture between the initial values $\mu_{\text{initial}}(\text{epoch})$ and $\sigma_{\text{initial}}(\text{epoch})$ and the estimate of the
205 current session's values $\mu_{\text{current}}(t, \text{epoch})$ and $\sigma_{\text{current}}(t, \text{epoch})$:

$$\mu_{\text{blended}}(t, \text{epoch}) = \alpha_j * \mu_{\text{initial}}(\text{epoch}) + (1 - \alpha_j) * \mu_{\text{current}}(t, \text{epoch}) \quad \text{Eqn. 5}$$

206

$$\sigma_{\text{blended}}(t, \text{epoch}) = \alpha_j * \sigma_{\text{initial}}(\text{epoch}) + (1 - \alpha_j) * \sigma_{\text{current}}(t, \text{epoch}) \quad \text{Eqn. 6}$$

207 After the first 25 trials α was set to zero which implies the μ and σ vectors kept being continuously
208 updated throughout the session but were no longer blended with values from the previous days.

209 The update rule for $\mu_{\text{current}}(t, \text{epoch})$ was:

$$\boldsymbol{\mu}_{\text{current}}(t, \text{epoch}) = \frac{[\boldsymbol{\mu}_{\text{current}}(t-1, \text{epoch})] * K + \boldsymbol{r}}{K + 1} \quad \text{Eqn. 7}$$

210
$$K = N_{\text{samples}}(t, \text{epoch})$$

211 where \boldsymbol{r} is the most recently sampled vector of spike counts and K is the current number of samples
 212 of spike count vectors obtained so far for this particular epoch.

213 The update rule for $\boldsymbol{\sigma}_{\text{current}}(t, \text{epoch})$ was:

$$\begin{aligned} & \boldsymbol{\sigma}_{\text{current}}(t, \text{epoch}) \\ & = \sqrt{\frac{K-1}{K} * \boldsymbol{\sigma}_{\text{current}}^2(t-1, \text{epoch}) + \frac{1}{K} * (\boldsymbol{r} - \boldsymbol{\mu}_{\text{current}}(t, \text{epoch}))^2} \quad \text{Eqn. 8} \end{aligned}$$

214 After updating the $\boldsymbol{\mu}_{\text{current}}(t, \text{epoch})$ and $\boldsymbol{\sigma}_{\text{current}}(t, \text{epoch})$ vectors, the number of samples for the
 215 corresponding epoch was also updated: $\overline{N_{\text{epoch}}}$

$$N_{\text{samples}}(t, \text{epoch}) = N_{\text{samples}}(t-1, \text{epoch}) + 1 \quad \text{Eqn. 9}$$

216 Importantly, even though we had only 3 different decoders (Fig. 1a) we effectively used 5 different
 217 epochs in experiments 1 and 2: Fixation, Targets, Dots, Delay and Post Go-Cue. The Dots decoder
 218 was also used in the Fixation and Targets epochs, but because average firing rates are different
 219 between these, different $\boldsymbol{\mu}$ and $\boldsymbol{\sigma}$ vectors had to be used. In experiment 3, there was no delay period,
 220 and we continued to use the Dots decoder in the Post Go-Cue period in order to track the post-
 221 pulse evolution of the DV. Every 50 ms sample of neural data for a given epoch was used to update

222 the corresponding μ and σ vectors as described above. We let the μ and σ vectors converge for
223 ~200-300 trials, in the beginning of each experimental session, before starting any closed-loop
224 experiments. One way to check for this convergence was to monitor the average DV value for the
225 first 150 ms of the Dots epoch. Since we verified through offline analyses that no systematic pre-
226 planning activity towards one of the two targets was present in PMd or M1 during this time
227 window, we expected the average DV value in this period to be ~0. Under these conditions a DV
228 offset (systematic difference between the average DV value and 0) suggested that an artificial bias
229 was being introduced in the decoded DV through inaccurately estimated μ and σ vectors, leading
230 the algorithm to under or overestimate the contribution of certain channels to the DV. Having
231 similar μ and σ vectors across days (Extended Data Fig. 6c-d) sped up their online convergence
232 thus increasing the number of trials available for closed-loop experiments within a session.

233 Using a single decoder for an entire epoch was far more efficient to implement than using a
234 different decoder for each time point (as it reduced the number of μ and σ vectors that had to be
235 learned online). In addition, and as demonstrated in an offline analysis (Extended Data Fig. 3c-e),
236 there is no performance difference between a single and multiple decoders for dots and go cue
237 epochs for either brain area. Because choice modulation in PMd/M1 changes dramatically around
238 the peri-movement period a single decoder for an entire trial was not feasible (Extended Data Fig.
239 4).

240 In the end, our method yielded a reliable real-time decision state read out and required only ~18%
241 (15%) of trials in a session for calculating the values of μ and σ for monkey H (F), leaving the
242 remainder available for imposing neurally contingent conditions in closed-loop. The real time
243 decoder was run on two separate PCs (server and client) using the Simulink Real-Time/xPC

244 platform (Mathworks, Massachusetts).

245 **11. Closed-loop experiments**

246 *11.1 Experiment 1: Virtual boundaries*

247 On each trial we set a virtual threshold, or boundary (B), for the magnitude of the DV during the
248 dots epoch. If the DV on the current trial reached B or $-B \pm$ tolerance, the dots presentation was
249 terminated and the monkey asked to report its decision. If the bound was not reached on a given
250 trial, stimulus presentation continued to a preset duration for that trial which was randomly
251 sampled from an exponential distribution ranging from 500-1200 ms. Closed-loop trials for which
252 the boundary was not reached were effectively indistinguishable from open-loop trials.

253 Typically, 5 values for boundaries spanning 0.5 to 5 (DV units) were used every session and one
254 of them was randomly assigned on each trial (uniform distribution). The tolerance used was \pm
255 0.25 DV units. We imposed a minimum duration for all trials to avoid spurious bound crossings,
256 which could be problematic for low bound values in particular. In all sessions the minimum
257 duration was 250 ms, a conservative estimate of the latency for choice related signals driven by
258 the visual stimulus to appear in PMd and M1.

259 After the minimum stimulus duration was reached, the DV was assessed every 10 ms to determine
260 whether it fell within ± 0.25 DV units of the boundary chosen for the current trial (B or $-B$). If so,
261 the stimulus was terminated within 34 ms of the boundary being met (see Methods section 11.4),
262 and we defined $DV_{\text{termination}}$ as the DV value that triggered stimulus termination (thus, $DV_{\text{termination}}$

263 took values within ± 0.25 DV units of the boundary $\pm B$ assigned on each trial). If the bound for the
264 particular trial was not reached, the presentation continued up to the maximum stimulus duration
265 selected for that trial which had been obtained by randomly sampling from an exponential
266 distribution: 500-1200 ms (median 670 ms).

267 Finally, we randomly interleaved closed-loop trials from both experiments 1 and 2 on 70% of the
268 trials (see breakdown below) with open-loop trials on 30% of the trials. In open-loop trials no DV-
269 dependent termination condition was imposed. The motivation for interleaving closed-loop and
270 open-loop trials was to make it extremely hard for the monkey to learn that accelerating the
271 dynamics of choice related signals¹¹ (potentially by recruiting more choice related neurons or
272 increasing their modulation) and thus hitting bounds sooner could potentially increase its reward
273 rate. Not accounting for this possibility could lead to an undesirable change in the monkey's
274 strategy during the course of the closed-loop experiments, which could become problematic when
275 combining data across days.

276 *11.2 Experiment 2: CoM detection*

277 Under our logistic regression framework, the signature of a putative CoM is a sign change of the
278 decision variable. Since these sign changes could happen at any time during the trial, capturing
279 them required not only monitoring the most recent state of the DV, but its history throughout the
280 trial. Because there was noise in our DV estimation and DVs usually started close to 0 at the
281 beginning of the trial we imposed selection criteria to detect likely CoMs based on the neural data.
282 A necessary feature for all potential CoMs was a zero crossing in the sign of the DV: change of
283 DV sign from negative to positive reflected a change in the likelihood of a rightward decision from

284 less than 50% to greater than 50%, and vice versa for the opposite change in sign. To eliminate
285 zero crossings resulting solely from measurement noise, we imposed four additional criteria:

286 • Minimum DV value after zero crossing;

287 • Minimum DV value with opposite sign before zero crossing;

288 • Minimum duration of DV sign stability after zero crossing;

289 • Minimum duration of DV sign stability before zero crossing;

290 The minimum DV values before and after zero crossing were symmetrical for most sessions, as
291 were the periods of minimum duration of DV sign stability (negative or positive values for all time
292 points). If a zero crossing was detected and all four criteria were met, the stimulus presentation
293 was interrupted and the animal was virtually immediately (within 34 ms or less, see Methods
294 section 11.4) prompted to report a decision. The exact parameters used for each session can be
295 found in Supplementary Table 4.

296 By sweeping the parameter space we could test zero crossings that differed in magnitude and
297 stability. Analogously to the virtual boundary experiment, if the minimums were not met and a
298 CoM thus not detected, the stimulus presentation continued uninterrupted for a random duration
299 ranging from 500-1200 ms, selected prior to the start of the trial. A minimum stimulus duration of
300 250 ms was also in place.

301 Because putative CoMs are quite rare², in the first set of experiments we devoted 70% of the
302 closed-loop trials to detect them leaving the remaining 30% as virtual boundary trials. The exact
303 fraction of trials with CoM depends dramatically on how we parameterize them. The longer the
304 minimum periods of consistent sign and the higher the minimum DV value in the initial
305 commitment stage, the rarer they become. Running both experiments on the same sessions ensured
306 that the mapping from DV to choice likelihood was held during the CoM experiments and provided
307 the most faithful indirect validation of initial commitment we could obtain.

308 *11.3 Experiment 3: Motion pulse perturbation*

309 In this experiment, motion pulses were introduced on some trials with motion coherences near or
310 below psychophysical threshold. No motion pulses were presented for suprathreshold coherences
311 based on the results of a pilot experiment (not shown) in which pulses presented at suprathreshold
312 coherences were more perceptually salient and led to changes in the animals' strategy. As in
313 Experiment 1, on each trial we set a virtual boundary (B) for the magnitude of the DV during the
314 dots epoch. In this experiment, 100% of trials with dots coherence at or near psychophysical
315 threshold were treated as closed-loop trials (this corresponds to trials with maximum unsigned
316 coherence of 6.4% for monkey H and 12.8% for monkey F; psychophysical thresholds were
317 measured using the Weibull function described above in Eqn. 1, Methods section 6). Low-
318 coherence trials in which the boundary was not reached (per the criteria below) and trials with
319 suprathreshold dots coherences were all effectively open-loop.

320 If the DV on a closed-loop trial reached B or $-B \pm$ tolerance (± 0.25 DV units), after a minimum
321 stimulus duration of 50 ms, a 200-ms motion pulse was presented, followed immediately by

322 termination of the visual stimulus and presentation of the cue for the monkey to report its decision.
323 If not, dots presentation continued for a pre-assigned duration drawn randomly from an
324 exponential distribution of 500-1200 ms. Four integer values for boundaries (spanning 1 to 4 DV
325 units) were used every session, and one of them was randomly assigned on each trial (uniform
326 distribution).

327 Motion pulses were 200-ms periods of additional dots stimulus presentation with small additive
328 average coherence ($\pm 2\%$ or 4.5% from the initial dots coherence on the same trial for monkey H
329 and F, respectively, where positive coherence values indicate rightward motion); thus pulses
330 effectively randomly added either a small amount of rightward or leftward motion evidence to the
331 stimulus. Pulse strength was calibrated in pilot experiments, in which we converged upon
332 coherence shifts that slightly but significantly biased each animal's behavior, without being overtly
333 perceptually salient (biases were measured using the logistic regression on rightward choice
334 described above in Eqn. 2, Methods section 6). Animals were rewarded for correct reaches in the
335 direction of the coherence of the initial dots stimulus (randomly assigned on 0% coherence trials),
336 regardless of the pulse direction.

337 *11.4 Estimated latency for real time closed-loop setup*

338 To validate our setup, we measured the latency between a neural condition being met and the
339 corresponding task change being implemented. We tested this latency by generating simulated DV
340 steps in the same model used to detect when DV triggering conditions were met in the real
341 experiments. We used these simulated steps to trigger the onset of a bright light on the touchscreen
342 in front of a photodetector, again within the same code used to run the task and generate the stimuli

343 in the real experiments. We then passed both the simulated DV and the photodetector output
344 signals into an oscilloscope, triggered the display on the “DV” steps, and manually measured the
345 delay to onset of the bright dot. Almost all measured delays were within 2 frames, or 26 ms.

346 *11.5 Estimated trial count savings for real time closed-loop setup*

347 The real time setup allowed for precise experimental control over which DV values or DV history
348 to use to trigger a modification in the task (stimulus termination or pulse). However, it could be
349 argued that given enough data, similar trials would have been captured simply by either
350 terminating the stimulus (as in experiments 1 and 2) or presenting the pulse (as in experiment 3)
351 at a random point in the trial and then back sorting them offline (by DV value or history after the
352 data is collected). Note that such an open-loop experiment would require an alternative algorithm
353 for stimulus modification, e.g. random timing within a pre-specified interval. It is conceivable that
354 the sampling algorithm itself could affect the animals’ behavioral strategy and/or the average
355 evolution of the DV over the course of a trial, which could impact our trial count yield in unknown
356 ways.

357 To estimate how much more trial-count efficient it was to use our real time setup compared to
358 offline back-sorting trials where the stimulus was presented for a random duration, we used the
359 CoM experiment as a case study given how rare change of mind events are (10.6% (17.2%) of
360 trials for monkey H (F)).

361 For simplicity, we focused on sessions 1, 2 and 3 from Monkey F, which all have the same (and
362 intermediate) CoM requirements (Supplementary Table 4). We started by calculating the yield

363 from the real time experiment in closed-loop as the ratio between detected CoM trials and trials in
364 which CoMs were checked (i.e. all closed-loop trials in which the stimulus could be terminated if
365 the conditions dictated by the CoM parameters were met, Supplementary Table 4):

$$\text{Yield}_{\text{CL}} = \# \text{CoMs detected} / \# \text{trials CoMs checked} = 11.91\% \quad \text{Eqn. 10}$$

366 To calculate the yield for offline back-sorting trials we used the open-loop trials in the same
367 sessions, which were terminated after a random stimulus duration. Importantly, the stimulus
368 duration on these open-loop trials was sampled from the same distribution as for the closed-loop
369 trials in which CoMs were checked, which allows for a fair yield comparison. We calculated the
370 yield from offline back-sorting as the ratio between the number of trials that would have met all
371 the criteria for CoMs for the same session and the total number of open-loop trials:

$$\text{Yield}_{\text{OL}} = \# \text{Valid putative CoMs} / \# \text{Open loop trials} = 1.85\% \quad \text{Eqn. 11}$$

372 Since the goal would be to probe the new choice preference shortly after the zero crossing (putative
373 change of mind), not many hundreds of ms later, we only considered CoM trials that were (closed-
374 loop) or that would have been (open-loop) terminated within 150 ms of the zero crossing. This
375 cutoff value corresponded to the 82nd percentile of post zero crossing durations for the closed-loop
376 trials analysed in these sessions.

377 In this analysis, Yield_{CL} was 6.43 times higher than Yield_{OL} . This result implies that had we not
378 used a real time setup in closed-loop we would have had to collect 6.43 times the number of trials
379 (and thus sessions) to obtain the same number of events. This would in turn mean collecting

380 around 100 sessions/monkey just for experiments 1 and 2 (assuming the same 30%/70% trial split
381 used in the real time experiments), rendering this experiment practically unfeasible.

382 **12. Logistic regression analysis of neural and behavioral data**

383 *12.1 Comparison of motion energy, mean DV and single trial DV as predictors of choice – Closed-* 384 *loop experiment 1*

385 To contextualize the choice prediction accuracy of our single-trial DV in the first closed-loop
386 experiment, we compared it to the performance of two other variables: the average motion energy
387 of the stimulus within each trial, and mean DV across trials (for the corresponding stimulus
388 coherence and time of stimulus termination). Motion energy quantifies the strength of the visual
389 stimulus along the horizontal axis and was calculated by convolving the positions of the dots with
390 spatio-temporal filters as previously described⁹. Due to the stochastic nature of the random dots
391 kinematograms, even within the same stimulus coherence level and direction, the motion energy
392 fluctuates from moment to moment and from trial to trial. For this reason, motion energy provides
393 a finer description of the visual evidence favoring a specific choice than stimulus coherence, and
394 we used it as a regressor here to quantify how stimulus fluctuations on individual trials predict
395 choice behavior.

396 We performed the model comparison analysis in two complementary ways. First, we built three
397 models (nested) each containing an additional variable of interest (average Motion Energy, Mean
398 DV and Single trial DV). Second, we trained three separate models (single regressor) each
399 containing only one of the three variables of interest. The first way tests how much additional

400 predictive power each regressor adds to an increasingly complex model and the second way tests
401 how each individual regressor by itself compares to the others.

402 For all models, we performed the following logistic regression on the probability of rightward
403 choice:

$$P_{\text{right}} = \frac{1}{1 + e^{-z}}. \quad \text{Eqn. 12}$$

404 For nested model 1 and single regressor model 1:

$$z = \beta_0 + \beta_{\text{ME}} \times \overline{\text{ME}} \quad \text{Eqn. 12a}$$

405 For nested model 2:

$$z(c, t) = \beta_0 + \beta_{\text{ME}} \times \overline{\text{ME}} + \beta_{\text{Mean DV}} \times \text{Mean DV}(c, t) \quad \text{Eqn. 12b}$$

406 For nested model 3:

$$z(c, t) = \beta_0 + \beta_{\text{ME}} \times \overline{\text{ME}} + \beta_{\text{Mean DV}} \times \text{Mean DV}(c, t) \\ + \beta_{\text{DV}} \times \text{DV}_{\text{termination}} \quad \text{Eqn. 12c}$$

407 For single regressor model 2:

$$z(c, t) = \beta_0 + \beta_{\text{Mean DV}} \times \text{Mean DV}(c, t) \quad \text{Eqn. 12d}$$

408 For single regressor model 3:

$$z = \beta_0 + \beta_{\text{DV}} \times \text{DV}_{\text{termination}} \quad \text{Eqn. 12e}$$

409 Where applicable, c is motion coherence, $\overline{\text{ME}}$ is average motion energy within a single trial and t
410 is elapsed time of stimulus presentation. We used 10-fold cross-validation to test the accuracy of
411 each model. Accuracy was defined as the percentage of correctly predicted choices on held out
412 test data. The regressors can be interpreted as follows:

413 β_0 – A constant term that captures the choice preference bias of the subject. A significant
414 positive (negative) coefficient implies the subject has a right (left) side choice bias that
415 cannot be attributed to the other regressors.

416 $\overline{\text{ME}}$ – Average Motion Energy within a single trial, calculated as described above for each
417 trial in closed-loop experiment 1 and averaged for a specific window of time during the
418 trial defined by latency (offset between DV and ME traces) and window size (averaging
419 period duration). In other words, for regression at time t , ME was averaged between [t -
420 (window size + latency)], t -latency]. To find the best model, 16 pairwise combinations of
421 4 values for latency ([100,150,200,250]) and 4 values for window size ([40,80,120,160])
422 were tested. A significant positive coefficient indicates the subject is using stimulus motion
423 energy information to guide choice in a way that is expected from the structure of the task
424 (increasing probability of a rightward choice as visual evidence changes from strong

425 leftward to strong rightward values). A significant negative coefficient would indicate the
426 opposite. The highest prediction accuracy was obtained using 100 ms latency for both
427 monkeys and 160 (120) ms window size for monkey H (F). For this reason, these were the
428 parameters used for averaging ME in Eqn. 12a-12c.

429 Mean $DV(t)$ - Average DV across all open and closed-loop trials as a function of stimulus
430 duration and stimulus coherence. A significant positive coefficient indicates that the
431 average DV at the time of termination and stimulus coherence for each trial is predictive
432 of choice likelihood with strong positive Mean DV values leading to higher likelihood of
433 a rightward choice and strong negative Mean DV values to higher likelihood of a leftward
434 choice. A significant negative coefficient would indicate the opposite.

435 $DV_{\text{termination}}$ - As defined in Methods section 11.1, the DV value at the time the stimulus
436 was terminated and the go cue presented. A significant positive coefficient indicates that
437 the DV at termination is predictive of choice likelihood beyond the other regressors with
438 strong positive DV values leading to higher likelihood of a rightward choice and strong
439 negative DV values to higher likelihood of a leftward choice. A significant negative
440 coefficient would indicate the opposite.

441 *12.2 Effect of experimental and internal factors on choice – Closed-Loop Experiment 1*

442 For the first closed-loop experiment we investigated the effect of experimental factors (stimulus
443 coherence and stimulus duration) as well as internal factors (DV at termination and DV derivative
444 at termination) on choice behavior in our task.

445 To quantify the effect of these factors on choice behavior, we performed the following logistic
446 regression on the probability of rightward choice:

$$P_{\text{right}}(c) = \frac{1}{1 + e^{-z}}, \quad \text{Eqn. 13}$$

447 where $z = \beta_0 + \beta_{\text{coh}} \times \text{coherence} + \beta_{\text{DV}} \times \text{DV}_{\text{termination}} + \beta_{\text{DV Diff}} \times \text{DV Diff}$
448 $+ \beta_{\text{stimulus duration}} \times \text{stimulus duration}$
449 $+ \beta_{\text{stimulus duration} * D} \times (\text{stimulus duration} * D)$

450 β_0 and $\text{DV}_{\text{termination}}$ are defined as above (see Methods section 12.1). The remaining regressors can
451 be interpreted as follows:

452 Coherence – Signed coherence of the stimulus presented on that trial (positive for rightward
453 stimuli and negative for leftward stimuli). A significant positive coefficient indicates the
454 subject is using coherence information to guide choice in a way that is expected from the
455 structure of the task (increasing probability of a rightward choice as stimulus coherence
456 changes from strong leftward to strong rightward values). The converse would be true for
457 a significant negative coefficient.

458 DV Diff – DV derivative calculated as the change in the DV value over the 50 ms preceding
459 termination. A significant positive coefficient indicates that the DV derivative at
460 termination is predictive of choice likelihood beyond stimulus coherence and $\text{DV}_{\text{termination}}$,
461 with strong positive DV Diff values leading to higher likelihood of a right choice and strong
462 negative values to higher likelihood of a left choice. For a significant negative coefficient

463 the opposite would be true. The DV Diff term is proxy for the direction the DV is heading
464 at and shortly after stimulus termination, presumably due in part to stimulus motion energy
465 already in the visual processing pipeline at the time of termination.

466 Stimulus Duration – Duration of the period between stimulus onset and stimulus offset for
467 each trial. A significant positive/negative coefficient indicates an increased
468 rightward/leftward bias for longer trials.

469 Stimulus Duration * D – Interaction term in which the stimulus duration is multiplied by
470 the sign of the correct choice for a given trial (stimulus direction, D ; +1 for right, -1 for
471 left). A significant positive coefficient indicates improved discrimination performance for
472 longer trials. The converse would be true for a significant negative coefficient.

473 The results from this regression analysis can be found in Supplementary Table 2.

474 *12.3 Offline logistic regression of neural activity on choice*

475 For each session, the responses of all neurons in 90% of the trials were fit with a logistic model
476 that attempted to separate rightward (T_1) and leftward (T_2) upcoming choices. The logistic model
477 was fit in 50 ms windows, advanced in 20 ms steps over the entire trial duration:

$$P(T_1|\vec{r}) = \frac{1}{1 + e^{-(\beta_0(t) + \sum_{i=1}^n \beta_i(t) \times r_i(t))}} \quad \text{Eqn. 14}$$

478

479 Where $P(T_1|\vec{r})$ is the probability of observing a particular behavioral choice (T_1 or rightward
480 choice in this case) given the population response \vec{r} ; $r_i(t)$ are the z-scored spike counts for each
481 neuron and time window, $\beta_0(t)$ is an intercept term and $\beta_i(t)$ are the classifier weights (one for each
482 neuron and time window).

483 The remaining 10% of the trials were tested using the previously trained model and its accuracy
484 was recorded. The same process was followed 10 times for each window (10-fold cross-validation)
485 and the percentage of correctly predicted choices recorded. This process was repeated for
486 consecutive windows displaced by 20 ms and yielding a prediction accuracy trace for each session
487 and brain area. Both correct and error trials were included in this analysis to assure there would
488 not be an imbalance between high coherence trials (more likely to be correct trials) and low
489 coherence trials, which would bias the classifier to perform better on high coherence trials.

490 An L1-regularization technique (LASSO) was used to constrain the norm of the β coefficients
491 fitted by the model to prevent over-fitting. The lambda parameter that determines the strength of
492 the penalty for the L1 norm was calculated for the 50 ms window preceding the go-cue by
493 sweeping through 25 potential values and selecting the value with lower deviance by running 10-
494 fold cross validation. This lambda value was then used for the model for all time points.

495 Finally, a slightly different procedure was used when training a single classifier over an entire
496 epoch. The four epochs used for training the four corresponding classifiers were:

- 497 • Targets epoch: [-150, 350] ms aligned to targets onset;

498 • Dots epoch: [150, dots offset] ms aligned to dots onset;

499 • Go epoch: [-600, 0] ms aligned to go cue;

500 • Peri-movement epoch: [-200, 600] ms aligned to reach;

501 All valid 50 ms samples of neural data during the selected period (above) for each epoch were
502 used as a sample to train the corresponding classifier. Valid 50 ms samples were those fully
503 included within the time limits of each epoch. We used the 10-fold cross-validation and LASSO
504 regularization methods explained above. However, the regularization parameter lambda was
505 calculated individually for each epoch through cross-validation and chosen as the value with
506 minimum expected deviance. Accuracy was calculated as fraction of test trials correctly predicted
507 at every 50 ms long window (stepped in 20 ms increments).

508 *12.4 Effect of experimental and internal factors on choice – Closed-loop experiment 2*

509 For the second closed-loop experiment we investigated the effect on choice behavior of the four
510 parameters enforced during the experiment to detect neural candidate CoMs: DV deflections
511 before and after the CoM and duration of the required period of stability in the sign of the DV
512 before and after the CoM. Signed stimulus coherence was also included in the regression given its
513 large influence on choice behavior in our task.

514 To quantify the effect of these factors on choice behavior, we performed the following logistic
515 regression on the probability of rightward choice:

$$P_{\text{right}}(c) = \frac{1}{1 + e^{-z}}, \quad \text{Eqn. 15}$$

516 where $z = \beta_0 + \beta_{\text{coh}} \times \text{coherence} + \beta_{\text{DV}} \times \text{DV}_{\text{termination}} + \beta_{\text{DVmaxopposite}} \times \text{DV}_{\text{maxopposite}}$
 517 $+ \beta_{\text{time post CoM} * D_1} \times \text{time post CoM} * D_1$
 518 $+ \beta_{\text{time pre CoM} * D_2} \times \text{time pre CoM} * D_2$

519 Here, D_1 is the sign of the DV at termination following the CoM, and D_2 is the opposite of D_1 ,
 520 enforced prior to the CoM. β_0 , coherence, and $\text{DV}_{\text{termination}}$ are defined as above (see Methods
 521 section 12.1 and 12.2). The remaining regressors can be interpreted as follows:

522 DV max opposite – Maximum deflection in DV value with sign opposite to $\text{DV}_{\text{termination}}$,
 523 achieved prior to the candidate CoM (zero-crossing). A significant positive coefficient
 524 indicates that the maximum opposite DV deflection decreases the likelihood of the final
 525 choice being congruent with the sign of the DV at termination: a strong early negative
 526 deflection would decrease the odds of a final rightward choice and vice-versa. The
 527 converse would be true for a significant negative coefficient.

528 Time post-CoM * D_1 – Interaction term in which the duration of the stability period after
 529 the zero-crossing is multiplied by the sign of $\text{DV}_{\text{termination}}$ for a given trial. A significant
 530 positive coefficient implies that a longer post-CoM duration is associated with an increased
 531 likelihood that the final choice is congruent with the sign of the DV at termination. The
 532 converse would be true for a significant negative term.

533 Time pre-CoM * D_2 – Interaction term in which the duration of the stability period before
534 the zero-crossing is multiplied by the opposite sign of $DV_{\text{termination}}$ for a given trial. A
535 significant positive coefficient implies that a longer pre-CoM duration is associated with a
536 decreased likelihood of the final choice being congruent with the sign of the DV at
537 termination. The converse would be true for a significant negative coefficient.

538 The results from this regression analysis can be found in Supplementary Table 3.

539 **13. DV variability**

540 Within trial variability was computed by first calculating the difference between consecutive DV
541 values (estimated every 10 ms) for every trial in the datasets collected for experiments 1 and 2
542 (open and closed-loop). This step yielded a DV derivative trace for each trial aligned to dots onset.
543 For each trial these traces were computed only up to the offset of the stimulus and did not include
544 any delay or post go-cue DV data. The DV derivative traces were then sorted and averaged for
545 each choice (Extended Data Fig. 10a-b) or each signed coherence level (Extended Data Fig. 10c-
546 d). Longer trials are increasingly rare due to the shape of our stimulus duration distribution, but
547 this asymmetry does not influence the interpretation of the time course of average DV derivative
548 as this metric only captures *within* trial variability and not *across* trial variance.

549 **14. DV and motion energy correlation**

550 Motion energy (ME) was calculated for each trial in the datasets collected for experiments 1 and
551 2 (open and closed-loop) as described above (Methods section 12.1). The ME trace obtained for

552 each trial captures the strength of the stimulus at every timepoint during the stimulus presentation
553 on individual trials. To evaluate the effect of motion energy on DV we performed a linear
554 regression of single trial DV traces on single trial ME traces. From experiments 1 and 3 we
555 determined that, due to neural latencies, a stimulus fluctuation only exerts an effect on the decoded
556 DV ~ 180 ms later. For this reason, the regression was always performed between DV(t) and ME($t-$
557 180ms) or earlier. In Extended Data Fig. 9a-b, each green trace corresponds to a different way to
558 estimate the motion energy that might affect DV at time t . For the lightest trace and for every
559 timepoint, t , ME was averaged between ($t-180$ ms) and ($t-200$ ms) for every trial and used to regress
560 against DV(t). A separate regression was performed for each timepoint and the resulting variance
561 explained was plotted. The same process then was repeated for every other green trace by
562 progressively increasing the averaging window for ME in 20 ms increments from ($t-180$, $t-200$)
563 ms to ($t-180$, $t-500$) ms. As a control the DV was also regressed against signed coherence for each
564 trial (Extended Data Fig. 9a-b grey traces). This analysis was used to assess how much of the DV
565 variance *across* coherences is explained by motion energy or signed coherence as a function of
566 time.

567 To assess how much DV variance *within* each coherence level could be explained by the motion
568 energy of the stimulus we first sorted the DV traces for each signed coherence level. For each
569 signed coherence level and each timepoint we regressed DV(t) against the instantaneous ME($t-$
570 180ms) for the corresponding trials and calculated the variance explained (Extended Data Fig. 9c-
571 d).

572 Finally, we performed a within-coherence analysis to assess the impact of the motion energy on
573 the direction and magnitude of the DV zero-crossings we captured as putative CoMs (Extended

574 Data Fig. 9e-f). To do this, we calculated the signed DV slope in the 100 ms centered on the DV
575 zero-crossing defining the CoM as the difference between DV at +50 ms and DV at -50 ms aligned
576 to CoM divided by the window size (100 ms) for all CoM trials. We then sorted the slopes by the
577 corresponding signed coherence level. For the trials in each signed coherence level, we regressed
578 the DV slope around the CoM against the ME averaged over the 100 ms preceding the CoM, offset
579 by the 180 ms estimated neural latency described above (ME was averaged between (t^*-180 ms)
580 and (t^*-280 ms) for every trial, where t^* is the time of the CoM zero-crossing).

581 **15. CoM regularities**

582 To test whether the effects of coherence on the number of CoMs were statistically significant we
583 used a bootstrap method to generate 1000 distributions of CoM events with the corresponding
584 coherences by sampling with replacement from the distribution of captured events for each subject
585 separately. For each subject each distribution had the same number of observations as those
586 captured in experiment 2: 985 for monkey H and 1727 for monkey F. For each randomly sampled
587 distribution the number of CoMs for each coherence level was counted. The resulting counts were
588 then regressed against the coherence level they belonged to. CoM count was highly and negatively
589 correlated with coherence for both subjects ($p < 0.001$).

590 To test whether the CoMs were more likely to correct a mistake than lead to one (corrective vs.
591 erroneous), we followed a similar bootstrapping procedure and generated 1000 distributions of
592 CoM events (excluding 0% coherence trials, in which correct and erroneous are undefined). For
593 each randomly sampled distribution the number of CoMs for each condition was counted. The
594 difference between the median counts of corrective and erroneous CoMs was tested by performing

595 a one-sided Wilcoxon rank sum test ($p < 0.001$) testing the hypothesis that corrective counts were
596 higher than erroneous counts.

597 Accurate and stable decoder performance (Extended Data Fig. 6) was crucial for studying rare
598 events such as CoMs, which could not have been characterized adequately using a single session's
599 data.

600 **16. Pulse effects**

601 To quantify the overall behavioral effect of the pulses, we performed the following logistic
602 regression on the probability of a rightward choice:

$$P_{\text{right}}(c) = \frac{1}{1 + e^{-\beta_1 \times (\beta_0 + \beta_{\text{pulse}} D + c)}}, \quad \text{Eqn. 16}$$

603

604 where c is motion coherence, β_1 is the slope parameter, D is the pulse direction, and $-\beta_0$ is the
605 motion strength corresponding to the indifference point.

606 To determine the effect of the pulse on the evolving DV, we first estimated the minimum latency
607 for the visual stimulus information to influence the DV by calculating the first time of significant
608 divergence of rightward vs. leftward DV traces during dots presentation on open-loop trials with
609 stimuli of maximal motion strength ($\pm 25.6\%$, 51.2% coherence for monkey H, F), assessed using
610 a two-sample t-test with correction for a false discovery rate of 0.05^{44} . We refer to this estimate as

611 the “evidence representation latency” (ERL; 170, 180 ms, bootstrapped SEM with 1000 iterations
612 of 58, 63 ms for monkey H, F, respectively). We then measured the evolution of the DV after pulse
613 presentation by calculating the difference between the empirically observed DV at each time point
614 t and the DV at the “pulse evidence representation latency” (PERL, or time of pulse onset plus the
615 ERL):

$$\delta DV(t) = DV(t) - DV_{\text{PERL}}. \quad \text{Eqn. 17}$$

616 DV values for all neural pulse analyses were calculated using the decoder trained from the Dots
617 epoch (see Methods sections 9-10). Because the DV calculated using the Dots decoder became
618 dramatically less reliable in the time bins preceding reach onset (see Extended Data Fig. 4), we
619 truncated these DV traces 150 ms before the single-trial reaction times (RTs). However, because
620 single-trial RTs varied, the number of trials contributing to each data point dropped off with
621 increasing time post-PERL. Therefore, to standardize the time window for analysis of neural pulse
622 effects across trials, δDV traces were terminated no later than 100 ms before the subject’s median
623 RT. Any trials for which the online DV traces were corrupted (truncated) before the end of this
624 time period (0 (1) closed-loop pulse trial for monkey H (F)) were excluded from further analysis.
625 Additionally, there were rare trials in which the online DV traces peaked dramatically over the
626 course of the trial, or with very large offsets from 0 at baseline. These outlier trials, defined as
627 trials with DV offset magnitudes reaching >50 DV units (33 (0) closed-loop pulse trials for monkey
628 H (F)), or DV values reaching magnitudes >100 DV units (4 (3) closed-loop pulse trials for
629 monkey H (F)), were also excluded from further analysis.

630 *16.1 Dissociating the influence of pulse time and DV at pulse onset – Closed-loop experiment 3*

631 We adopted a data-driven, somewhat brute force approach to disentangle the effects of DV_{PERL}
632 and time. The same approach was used separately to analyze the neural and behavioral pulse effects
633 (measured using the change in the DV post-PERL or the monkey's choice, respectively). In
634 essence, we first calculated the condition-averaged response variable (choice or post-PERL change
635 in DV) for each of many conditions specified below, then subtracted the condition-averaged effect
636 from single trial measurements for each trial belonging to each condition. This yielded a set of
637 residual values, one for each trial, that we then combined across all conditions to generate
638 statistical power. To calculate how the time-adjusted pulse effect varied with DV_{PERL} , stimulus
639 duration was included in the definition of the conditions (described below). Thus residuals were
640 calculated separately for each time bin before combining the data to examine the effect of DV_{PERL}
641 on the pulse effect. Similarly, to calculate how the DV-adjusted pulse effect varied with time,
642 DV_{PERL} (not time) was included in the definition of conditions and the individual trials were re-
643 sorted among conditions before subtraction of the condition-averages. This yielded a set of
644 residuals that were calculated separately for each DV_{PERL} bin before being combined across
645 conditions to achieve statistical power. This basic procedure was used to calculate the data of
646 Figure 4e-h and Extended Data Figure 8c-j. We now describe the procedure in detail.

647 The data set for this analysis consisted of all trials with motion pulses in experiment 3 (19111 trials
648 for monkey H and 20358 for monkey F). We define the response variables as follows:

649 Choice: Binary values assigned 1 for rightward choices and 0 for leftward choices.

650 ΔDV : We consider neural pulse effects over the time window from PERL to 100 ms
651 before median RT *or* 150 ms before the single-trial RT (whichever came first; see

652 Methods section 16 above). In order to quantify the single-trial change in DV after the
653 pulse, we define ΔDV as the DV averaged over the last 50 ms of that time window, *minus*
654 the DV averaged in the 50ms pre-PERL. This variable is continuous.

655 We also define the following predictors:

656 Coherence: Signed coherence of the baseline (pre-pulse) dots stimulus. This variable
657 could take one of 7 possible values on a given trial for each subject (3 negative/leftward,
658 one zero, and 3 positive/rightward peri- and subthreshold coherences)

659 DV_{PERL} : The DV at the time of pulse onset plus the estimated PERL for each subject, as
660 described above. Negative values correspond to greater likelihood of a leftward choice, as
661 per conventions in experiments 1 and 2. This variable is continuous but was rounded to
662 the nearest integer for conditionalization as described below. Trials with DV_{PERL} values \leq
663 -5.5 or ≥ 5.5 were discarded (359 trials were discarded for monkey H and 455 for
664 monkey F, corresponding to 1.9% and 2.2% of trials, respectively).

665 Stimulus duration: Duration of the period between stimulus onset and offset for each
666 trial, *including* the 200 ms motion pulse. This variable could take values between 250 and
667 1400 ms. To pool data across subjects, the stimulus duration was normalized to fall in the
668 range [0 1] for each subject before pooling.

669 Pulse direction: Binary values assigned 1 for rightward pulse trials and -1 for leftward
670 pulse trials.

671 We used these variables as follows to calculate, statistically evaluate, and visualize the residual
672 pulse effects described in general terms above.

673 *Step 1: Re-sign variables according to the direction of the baseline stimulus*

674 To reduce the number of conditions in the following analyses, we re-signed all relevant
675 variables according to their congruence with the baseline motion stimulus (this allowed
676 us to use the *unsigned* baseline stimulus coherence in later steps). For trials with leftward
677 (negative) baseline coherences, we flipped the signs (or the values of the binary choice
678 variable) of DV_{PERL} , pulse direction, choice, and ΔDV . The sign-adjusted variables are
679 denoted with daggers. Thus pulse direction[†] is positive (‘congruent’) for trials with
680 rightward pulses on rightward baseline dots *or* leftward pulses on leftward baseline dots;
681 it is negative (‘incongruent’) for trials with rightward pulses on leftward baseline dots *or*
682 leftward pulses on rightward baseline dots. After re-signing, choice[†] takes a value of 1
683 for correct responses and 0 for incorrect responses, because the direction of the baseline
684 stimulus determines the correct response in our task.

685 Pulses were also presented on trials with 0% baseline motion coherence. In that case, we
686 used the sign of DV_{PERL} as a proxy for the momentary directional evidence at the time of
687 pulse onset. Thus, on 0% coherence trials with $DV_{\text{PERL}} < 0$, we flipped the signs of
688 DV_{PERL} , pulse direction, choice, and ΔDV .

689 *Step 2: Sort trials into conditions*

690 For analyses of pulse effects *over time*, to isolate the effects of stimulus duration from
691 those of the pre-pulse motion strength and DV_{PERL} , trials were first sorted into one of 88
692 conditions (11 rounded DV_{PERL} integer values from -5 to 5 * 4 unsigned baseline
693 coherence levels * 2 pulse directions), irrespective of stimulus duration. For analyses of
694 pulse effects *over DV_{PERL}* , to isolate the effects of DV_{PERL} from those of the pre-pulse
695 motion strength and stimulus duration, trials were sorted into one of 64 conditions (eight
696 stimulus duration quantiles * 4 unsigned baseline coherence levels * 2 pulse directions),
697 irrespective of DV_{PERL} . (As mentioned above, we accounted for
698 congruence/incongruence with the baseline dots stimulus by re-signing the variables in
699 Step 1, which eliminates the need to sort by the signed baseline coherence.)

700 *Step 3: Calculate single-trial residual pulse effects around within-condition means*

701 This is the critical step in isolating the predictor variable of interest for each analysis.
702 Residual *behavioral* pulse effects ($choice^{\dagger}_{res}$) were calculated by subtracting the mean
703 $choice^{\dagger}$ for each condition from the single-trial $choice^{\dagger}$ values for trials within that
704 condition. Similarly, residual *neural* pulse effects ($\Delta DV^{\dagger}_{res}$) were calculated by
705 subtracting the mean ΔDV^{\dagger} for each condition from the single-trial ΔDV^{\dagger} values. Note
706 that the conditions were defined differently for analyses of pulse effects over stimulus
707 duration vs. DV_{PERL} , as described in Step 2. We denote these different residuals with the
708 subscript $resT$ for the analyses of pulse effects over stimulus duration, and $resD$ for the
709 effects over DV_{PERL} .

710 *Step 4: Quantify effects*

711 *Statistical analysis*

712 Based on preliminary observations, we expected that pulse effects might decrease at
713 longer stimulus durations, even after accounting for the effects of the baseline motion
714 strength and DV_{PERL} . To test this hypothesis, we re-sorted the residuals ($\text{choice}_{\text{resT}}^{\dagger}$ or
715 $\Delta DV_{\text{resT}}^{\dagger}$) by pulse direction only, and then fit curves to the residual pulse effects
716 ($\text{choice}_{\text{resT}}^{\dagger}$ or $\Delta DV_{\text{resT}}^{\dagger}$) over the full distribution of stimulus durations using the built-in
717 function *fitnlm* in MATLAB.

718 The behavioral and neural pulse effects were fit separately using the following equations
719 for a half-Gaussian curve centered at the minimum stimulus duration:

$$\text{choice}_{\text{resT}}^{\dagger} = \text{dir}_{\text{pulse}} \times (\beta_{\text{baseline}} + \beta_{\text{amplitude}} \times e^{-z}) \quad \text{Eqn. 18}$$

720 and

$$\Delta DV_{\text{resT}}^{\dagger} = \text{dir}_{\text{pulse}} \times (\beta_{\text{baseline}} + \beta_{\text{amplitude}} \times e^{-z}) \quad \text{Eqn. 19}$$

721

$$\text{where } z = \frac{(\text{dur}_{\text{stim}} - \min(\text{dur}_{\text{stim}}))^2}{\beta_{\text{width}}^2}$$

722 Here $\text{dir}_{\text{pulse}}$ is the pulse direction (-1 for stimulus incongruent and +1 for stimulus
723 congruent pulses, as defined above) and dur_{stim} is the stimulus duration. A significant

724 positive fitted amplitude coefficient ($\beta_{\text{amplitude}}$) for a given model implies that the pulse
725 effect decreases over time.

726 Analogously, our preliminary observations suggested that pulse effects might decrease at
727 large $|\text{DV}_{\text{PERL}}|$ values, even after accounting for the effects of the baseline motion
728 strength and the stimulus duration. To test this hypothesis, we re-sorted the residuals
729 ($\text{choice}_{\text{resD}}^{\dagger}$ or $\Delta\text{DV}_{\text{resD}}^{\dagger}$) by pulse direction only, and then fit curves to the residual pulse
730 effects ($\text{choice}_{\text{resD}}^{\dagger}$ or $\Delta\text{DV}_{\text{resD}}^{\dagger}$) over the full distribution of $|\text{DV}_{\text{PERL}}|$ values, again using
731 the MATLAB *fitlm* function.

732 Here, the behavioral and neural pulse effects were fit separately using the following
733 equations for a half-Gaussian curve centered at zero:

$$\text{choice}_{\text{resD}}^{\dagger} = \text{dir}_{\text{pulse}} \times (\beta_{\text{baseline}} + \beta_{\text{amplitude}} \times e^{-z}) \quad \text{Eqn. 20}$$

734 and

$$\Delta\text{DV}_{\text{resD}}^{\dagger} = \text{dir}_{\text{pulse}} \times (\beta_{\text{baseline}} + \beta_{\text{amplitude}} \times e^{-z}) \quad \text{Eqn. 21}$$

735

$$\text{where } z = \frac{|\text{DV}_{\text{PERL}}|^2}{\beta_{\text{width}}^2}$$

736 Here, $\text{dir}_{\text{pulse}}$ is again the pulse direction (-1 for stimulus incongruent and +1 for stimulus
737 congruent pulses, as defined above). A significant positive fitted amplitude coefficient
738 ($\beta_{\text{amplitude}}$) for a given model implies that the pulse effect decreases over $|\text{DV}_{\text{PERL}}|$.

739 We also performed a final complementary statistical analysis of the pulse effects over
740 *signed* DV_{PERL} , hypothesizing that the effects would peak near zero/at lower DV_{PERL}
741 magnitudes. To test this hypothesis, we again took the residual pulse effects re-sorted by
742 pulse direction only, and then fit curves to the residuals ($\text{choice}_{\text{resD}}^{\dagger}$ or $\Delta\text{DV}_{\text{resD}}^{\dagger}$) over the
743 full distribution of signed DV_{PERL} values.

744 Here, the behavioral and neural pulse effects were fit separately using the following
745 equations for a Gaussian curve centered at zero:

$$\text{choice}_{\text{resD}}^{\dagger} = \text{dir}_{\text{pulse}} \times (\beta_{\text{baseline}} + \beta_{\text{amplitude}} \times e^{-z}) \quad \text{Eqn. 22}$$

746 and

$$\Delta\text{DV}_{\text{resD}}^{\dagger} = \text{dir}_{\text{pulse}} \times (\beta_{\text{baseline}} + \beta_{\text{amplitude}} \times e^{-z}) \quad \text{Eqn. 23}$$

747

$$\text{where } z = \frac{(\text{DV}_{\text{PERL}} - \beta_{\text{center}})^2}{\beta_{\text{width}}^2}$$

748 Here, $\text{dir}_{\text{pulse}}$ is the pulse direction as defined above. A significant positive fitted
749 amplitude coefficient ($\beta_{\text{amplitude}}$) for a given model implies that the pulse effect decreases
750 at larger DV_{PERL} magnitudes.

751 *Visualization*

752 To visualize the variation of the pulse effects *over time*, we binned the residuals
753 ($\text{choice}_{\text{resT}}^{\dagger}$ or $\Delta\text{DV}_{\text{resT}}^{\dagger}$) into eight equally populated stimulus duration quantiles and
754 subdivided each duration quantile by pulse direction (minimum trial counts in each
755 subdivided condition: 1122, 1217 for monkey H, F respectively). We calculated the mean
756 residual value in each of these 16 new conditions, and quantified the *overall* pulse effect
757 as the difference between the mean residuals for congruent minus incongruent pulse
758 trials, $\Delta(\text{Mean choice}_{\text{resT}}^{\dagger})$ or $\Delta(\text{Mean } \Delta\text{DV}_{\text{resT}}^{\dagger})$, in each of the 8 stimulus duration bins.
759 Because we did not have both congruent and incongruent pulses on individual trials, and
760 because the number of congruent and incongruent pulses in each condition did not
761 perfectly match, we used a bootstrap procedure to estimate the standard error of these
762 differences between mean residuals. We randomly resampled each distribution, with
763 replacement, to build 1000 new distributions of the same sample size as the original in
764 each of the 32 conditions described above (stimulus duration quantile and pulse
765 direction). For each resampled data set, we calculated the difference between the mean
766 residuals for congruent minus incongruent pulse trials in each of the 8 stimulus duration
767 bins. We used the standard deviation of these 1000 differences of means to estimate the
768 standard error of the mean of the original measurements. In Fig. 4g-h (Extended Data
769 Fig. 8e-f), we plotted in black the mean and bootstrapped SEM of the overall pulse

770 effects in the data (difference between residuals for congruent and incongruent pulse
771 trials) for each of the 8 stimulus duration bins. We then overlaid in blue the model fits
772 (described in “*Statistical analysis*” above) to the actual data in each of the 8 stimulus
773 duration bins.

774 For analyses of pulse effects *over* $|DV_{\text{PERL}}|$, we re-sorted the residuals (choice[†]_{resD} or
775 $\Delta DV^{\dagger}_{\text{resD}}$) by the rounded $|DV_{\text{PERL}}|$ and pulse direction to calculate the mean in each of
776 those 12 new conditions (minimum trial counts in each subdivided condition: 501, 504
777 for monkey H, F respectively). We then quantified the *overall* pulse effect as the
778 difference between the mean residuals for congruent minus incongruent pulse trials for
779 each of the 6 $|DV_{\text{PERL}}|$ values, plotted in black in Fig. 4e-f (Extended Data Fig. 8c-d,
780 $\Delta(\text{Mean choice}^{\dagger}_{\text{resD}})$ or $\Delta(\text{Mean } \Delta DV^{\dagger}_{\text{resD}})$). Error bars for these differences of means
781 were calculated by the bootstrap procedure described above. We overlaid in blue the
782 model fits to the actual data in each of the 6 $|DV_{\text{PERL}}|$ conditions. This procedure was also
783 followed using the 11 rounded signed $DV^{\dagger}_{\text{PERL}}$ values to generate Extended Data Fig. 8g-
784 j (minimum trial counts in each subdivided condition: 149, 151 for monkey H, F
785 respectively).

786 **17. General statistical analysis and reproducibility**

787 Sample sizes were chosen based on our labs' experience and standards in the field. Trial count
788 targets for experiment 3 were based on the effect sizes and sample sizes used in prior
789 psychophysical and electrophysiological studies with similar motion pulses^{9,19}. Unless explicitly
790 described, behavioral data without neural recordings are not reflected in the main body of the

791 manuscript; we only report behavioral data during neural recordings. Any extremely noisy neural
792 recording channels were deactivated at the beginning of a session, and all other channels were
793 used in this study. Trial parameters (including stimulus difficulty, stimulus direction, stimulus
794 duration, pulse direction, delay duration, and DV boundary values) were randomly assigned.
795 Blinding was not relevant for this study because trial parameters within each experiment were
796 automatically randomly assigned and interleaved without cues (subjects could not anticipate, and
797 experimenters did not autonomously control, the parameters used on a given trial). All data
798 collection and analyses were carried out in two subjects and were generally consistent as
799 presented in this article; no additional replications were attempted.

800 **Supplementary Notes**

801 **Supplementary Note 1 – Effect of coherence and stimulus duration on DV and prediction** 802 **accuracy.**

803 Figure 2c (Extended Data Fig. 1d) combines trials across a wide range of coherences and
804 stimulus durations, aggregated across 17 (15) sessions from monkey H (F). To identify
805 experimental factors that might influence the observed relationship between DV at termination and
806 prediction accuracy, we first resorted the same trials in Figure 2c by stimulus coherence. The
807 results show that there is a separation between the curves for high and low coherence trials
808 (Extended Data Fig. 1e) with higher accuracy for high coherence trials. The shift is small but
809 reliable across monkeys. We hypothesized that this difference resulted from motion energy signals
810 already en route from the retina to PMd/M1 (~175 ms latency) when the DV reached stimulus
811 termination. More motion energy signals would be arriving from this neural ‘pipeline’ on high
812 coherence trials, leading to a slightly higher DV than we measured at stimulus termination.

813 To assess this possibility, we measured the derivative of the DV around termination and
814 performed the following two analyses. First, we checked whether DV derivative explained a
815 significant fraction of choice variance beyond DV value alone (see Methods section 12.2). For
816 both monkeys the effect of DV derivative (defined as the DV slope in the last 50 ms of stimulus
817 presentation) was significant ($p = 0.02$, $p = 4.5 \times 10^{-11}$ for monkey H and F, respectively) and the
818 effect was congruent with our hypothesis: stronger positive derivatives predicted higher likelihood
819 of rightward choices and stronger negative derivatives predicted higher likelihood of leftward
820 choices (Supplementary Table 2, “DV diff”). Second, we tested whether high coherence trials were
821 associated with higher DV derivatives at termination by performing linear regression of DV

822 derivatives as a function of signed coherence. For both monkeys signed coherence was strongly
823 predictive of DV slopes: $p = 2.17 \times 10^{-171}$ and $R^2 = 0.23$ for monkey H and $p = 1.57 \times 10^{-105}$ and R^2
824 $= 0.16$ for monkey F. These results confirm that DV derivative is predictive of choice beyond DV
825 alone and show that higher coherence trials are associated with higher DV derivatives. The data
826 are consistent with our hypothesis above that the DV continues to evolve under the influence of
827 ‘pipeline’ sensory information for a short interval following stimulus termination, resulting in
828 somewhat better prediction accuracy than expected from the DV at termination, especially at high
829 coherences.

830 Sorting trials by duration (Extended Data Fig. 1f) reveals a different effect: the centers of
831 the quantiles are strongly shifted to the right (higher DV magnitudes) for longer stimuli compared
832 to shorter stimuli. This effect is expected from multiple sequential sampling models^{9,15,17,19}. In drift
833 diffusion models, for example, diffusion to high decision bounds requires more time than for low
834 bounds given constant stimulus coherence¹². However, we tested whether stimulus duration *per se*
835 was a significant predictor of choice independently of DV value by including two additional
836 regressors in our logistic model of choice: stimulus duration (representing choice bias as a function
837 of time) and an interaction term between stimulus duration and direction (representing increased
838 sensitivity to stimulus coherence as function of time). Neither regressor was significant for either
839 monkey ($p > 0.05$, Supplementary Table 2), implying that the likelihood of making one or the other
840 choice depended on DV value independently of the time required to reach that value.

841

842 **Supplementary Note 2 – Interpretation and validation of individual neural CoMs**

843 By the most demanding standard, a true CoM requires a full commitment to an initial
844 choice that is later reversed. We cannot assert unequivocally that this was the case for each putative

845 CoM that we captured from the neural data since only one behavioral choice is reported per trial.
846 However, in aggregate the median magnitude of DV deflection prior to the zero crossing was 2.1
847 (2.2) for monkey H (F), indicating that the subjects were expected to have chosen their initial
848 preference ~90% of the time had the trial stopped then. In addition, across all trials, we correctly
849 predict the final choice in 80% (92%) of the trials for monkey H (F). Putting these two results
850 together for the pre- and post-CoM likelihood of choice strongly suggests that the vast majority of
851 captured CoMs in our experiment are indeed real CoMs, even though we cannot say with 100%
852 confidence that the subject changed its mind on any individual trial.

853

854 **Supplementary Note 3 – Effect of DV trajectory parameters on choice likelihood.**

855 We formally tested the hypothesis that choice likelihood was influenced by some aspect of
856 the DV trajectory history by regressing choice as a function of three additional parameters (in
857 addition to the DV at termination) that were enforced and monitored in this experiment (see
858 Methods section 12.4): maximum DV deflection before sign change, and duration of sign stability
859 before and after DV sign change. For monkey F, no additional factor was choice predictive,
860 whereas for monkey H both the duration of sign stability before and after the CoM were also choice
861 predictive (Supplementary Table 3) as suggested by Extended Data Fig. 7a.

862

863 **Supplementary Note 4: Intra-trial fluctuations in DV are likely driven by multiple factors,** 864 **including the visual stimulus and noise sources.**

865 Interestingly, we frequently observed large, behaviorally relevant fluctuations in the DV—
866 equal in amplitude to many CoM fluctuations—even when the DV remained on one side of the
867 discriminant hyperplane in non-CoM trials (e.g. Figs. 1e and 2b). We wondered whether these DV

868 fluctuations were related to moment-to-moment stochastic variations in motion strength of the
869 stimulus within single trials. While average motion energy explains a large portion of DV variance
870 across coherence levels (Extended Data Fig. 9a-b, Methods 14), our data show that stochastic,
871 momentary stimulus fluctuations are not the dominant cause of DV fluctuations within single trials
872 (Extended Data Fig. 9c-d, Methods 14). Even in CoM trials, moment-to-moment motion strength
873 fluctuations within coherence levels explain a very small proportion in the variance of the signed
874 DV slope during the 100 ms centered around the zero-crossing defining the CoM, which
875 determines the direction of the CoM (Extended Data Fig. 9e-f). Considered as a whole, the analyses
876 in ED Fig. 9 suggest that single-trial DV fluctuations, including neurally defined CoMs, are likely
877 shaped by a variety of factors, some dependent on the visual stimulus while others (including
878 multiple possible noise sources) are independent of the stimulus. Further experiments will be
879 needed to address the source(s) of these fluctuations and their relationship with fluctuations in
880 other brain areas⁴⁵ as well as other cognitive processes including motor preparation and
881 execution^{46,47}, attention, motivation, and confidence.

882

883 **Supplementary Note 5 – Computational models**

884 As mentioned in the Discussion, models of the decision-making process that are compatible
885 with our finding of absorbing decision bounds include linear integration to a bound (such as a
886 simple stopping criterion) or more complex nonlinear integration processes^{15,22,23,25}. However, our
887 finding that pulses are less effective when presented later in the trial cannot be explained without
888 extensions of or alternatives to these models, such as those incorporating time-varying decision
889 bounds or other “urgency” signals (which could be implemented via gain modulation rather than
890 collapsing bounds)^{16-18,20,21,27,28,48,49}.

891 The bistable attractor dynamics and leaky-competing accumulator models can also be
892 extended to permit commitment to a choice prior to their attractor states, by adding decision bounds
893 to these models. Proper parameterization of these extended models could replicate our
894 experimental observations. Interestingly, similar extensions have been used to explain behavior in
895 reaction-time tasks^{15,22,23} but they are commonly avoided for fixed and variable duration tasks—a
896 practice that should be revised in light of our results. Our results are also compatible with discrete
897 attractor networks with a nonselective ramping input as proposed in Inagaki et al. 2019²⁴.

898 Finally, we cannot completely rule out a time-dependent termination of *attention* to the
899 visual stimulus in combination with a model that otherwise lacks a time dependent *decision*
900 termination criterion. In independent behavioral experiments with the same two monkeys¹¹, we
901 estimated the psychophysical integration times to be ~530 ms and ~680 ms for monkeys H and F
902 respectively, which is equivalent to the first ~3(~6) duration bins for monkey H(F) in Figure 4g,h.
903 Thus, while variable attention might account for some of the decrease in pulse effect for longer
904 duration trials in Figure 4g,h, it is unlikely to account for the more pronounced decrease in shorter
905 duration trials.

906

907 **Supplementary Note 6 – Caveats and future directions**

908 In considering our results as a whole, it is reasonable to consider two caveats. First, the
909 impressive choice prediction accuracy achieved in this study using a linear decoder does not imply
910 that the brain’s decision formation process is also linear. In principle, a linear decoder could predict
911 binary choices quite well even if the true neural process underlying decision formation were
912 nonlinear, depending on the form of the nonlinearity (see, e.g., Sussillo et al. 2016⁵⁰ for an example
913 of a linear neural to kinematic decoder which only slightly underperforms a more powerful

914 nonlinear recurrent neural network). Regardless, our linear DV is tightly linked to choice behavior
915 (e.g. Fig. 2c), showing that variations in the DV magnitude meaningfully track the ongoing process
916 of decision formation despite any nonlinearities in the underlying neural mechanism. Second, it
917 is possible in principle that our decoded DV might in part reflect movement preparation or even
918 small changes in muscle tone over the course of individual trials⁵¹⁻⁵³. However, given the
919 previously reported lack of EMG modulation during an enforced delay period in cued reaching
920 tasks⁵⁴⁻⁶⁰, it is unlikely that such movement-related signals explain the bulk of the variance in our
921 decoded DV.

922 Future work could build upon our neurally contingent feedback approach with experiments
923 designed to interrogate additional aspects of the neural computations underlying decision
924 formation. For example, one could imagine closed-loop experiments that dynamically titrate
925 stimulus information from moment to moment during the trial to “clamp” the DV at particular
926 values to further dissociate the DV from elapsed time in a given trial, or perhaps drive the DV to
927 particular values earlier or later in a trial to better understand the effects of DV history on the
928 upcoming choice (and its susceptibility to changes in the stimulus).

929 **Supplementary Tables**

Results for Models with 1 regressor + bias term		
Choice Prediction Accuracy at termination for LR models with the following regressors:	Monkey H	Monkey F
Signed Coherence	77.80%	73.60%
Motion Energy	74.50%	71.50%
Mean DV	78.70%	72.90%
Single Trial DV	88.40%	83.40%
Results for nested models + bias term		
Choice Prediction Accuracy at termination for LR models with the following regressors:	Monkey H	Monkey F
Motion Energy	74.50%	71.50%
Motion Energy, Mean DV	78.10%	73.20%
Motion Energy, Mean DV, Single Trial DV	89.30%	85.90%

930
 931 **Supplementary Table 1 – Prediction accuracy results for single regressor and nested**
 932 **regressor logistic regression models for both subjects (see Methods section 12.1).**

933

934

Logistic Regression on Choice

Predictor	Monkey H			Monkey F		
	Beta value	95% CI	p-value	Beta value	95% CI	p-value
Bias	-0.1613	[-0.2987 , -0.02381]	0.02147	-0.1197	[-0.2413 , 0.001806]	0.0535
Coherence	2.747	[2.268 , 3.227]	2.757e-29	1.885	[1.55 , 2.22]	2.912e-28
DV Termination	2.073	[1.779 , 2.367]	1.887e-43	1.708	[1.52 , 1.895]	1.626e-71
DV Diff	0.2989	[0.0488 , 0.5491]	0.01916	0.577	[0.4062 , 0.7478]	3.586e-11
Stimulus duration	0.1199	[-0.01489 , 0.2548]	0.08125	0.07474	[-0.04713 , 0.1966]	0.2293
Stimulus duration* stimulus direction	-0.02803	[-0.2053 , 0.1492]	0.7566	0.0466	[-0.1132 , 0.2064]	0.5677

935

936 **Supplementary Table 2 – Coefficients obtained from logistic regression on choice – virtual**

937 **boundary experiment (see Methods section 12.2).**

938

939

940

Logistic Regression on Choice after CoM

Predictor	Monkey H			Monkey F		
	Beta Value	95% CI	p-value	Beta Value	95% CI	p-value
Bias	-0.2876	[-0.4658 , -0.1093]	0.001568	-0.192	[-0.3773 , -0.00662]	0.04235
Coherence	1.304	[1.005 , 1.602]	1.189e-17	1.284	[0.96 , 1.608]	8.028e-15
DV Termination	1.67	[1.158 , 2.182]	1.623e-10	2.46	[1.872 , 3.049]	2.51e-16
DV max opposite	0.06967	[-0.4461 , 0.5854]	0.7912	-0.2387	[-0.8959 , 0.4185]	0.4766
Time after CoM * sign(DV Termination)	0.7065	[0.2317 , 1.181]	0.003544	0.02265	[-0.3835 , 0.4288]	0.913
Time before CoM * sign(DV max opposite)	0.7471	[0.3083 , 1.186]	0.0008467	-0.1427	[-0.6394 , 0.3539]	0.5732

941

942 **Supplementary Table 3– Coefficients obtained from logistic regression on choice - change of mind**

943 **experiment (see Methods section 12.4).**

Subject	Session	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	
Monkey H	tminPre(s)	0.1	0.1	0.1	0.1	0.05	0.05	0.05	0.05	0.05	0.05	0.15	0.15	0.15	0.1	0.05	0.15	0.15	
	tminPost(s)	0.1	0.1	0.1	0.1	0.05	0.05	0.05	0.05	0.05	0.05	0.15	0.15	0.15	0.15	0.15	0.05	0.1	
	DVminPre	2	1.5	3	1	2	1	1.5	3	1.5	1	2	2	2	2	2	2	2	2
	DVminPost	2	1.5	3	1	2	1	1.5	3	0.1	0.1	2	2	2	2	2	2	2	2
	tminDots(s)	0.25	0.25	0.25	0.25	0.25	0.25	0.25	0.25	0.25	0.25	0.25	0.25	0.25	0.25	0.25	0.25	0.25	0.25
Monkey F	tminPre(s)	0.1	0.1	0.1	0.1	0.1	0.1	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.15	0.15			
	tminPost(s)	0.1	0.1	0.1	0.1	0.1	0.1	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.1			
	DVminPre	2	2	2	1.5	3	1	2	3	1.5	1	1	1.5	1	2	2			
	DVminPost	2	2	2	1.5	3	1	2	3	1.5	1	1	0.1	0.1	2	2			
	tminDots(s)	0.25	0.25	0.25	0.25	0.25	0.25	0.25	0.25	0.25	0.25	0.25	0.25	0.25	0.25	0.25	0.25		

944

945 **Supplementary Table 4 – Parameters used for virtual boundaries and CoM closed loop**

946 **experiments (monkeys H and F).**

	Number of Channels with non-zero weights for the Decoders					
	Monkey H - 192 channels available			Monkey F - 96 channels available		
Decoder	Monkey H - Dots	Monkey H - Delay	Monkey H - Post Go	Monkey F - Dots	Monkey F - Delay	Monkey F - Post Go
D1	122	77	92	71	53	60
D2	117	54	51	78	72	76
D3	131		98	67	53	74
D4	70		108	80	73	66
D5	79		86	67		73
D6	120		84	71		60
D7	112		96	80		80
D8	100		97	79		78
D9	113		81	74		67
D10	124		70	75		60
D11	121		66	80		73
D12	88		74	83		86
D13	89		86	79		75
D14				74		57
D15				66		75
D16				77		77
D17				70		72
D18				83		82
D19				73		65
D20				71		69
D21				71		69
D22				70		79

948

949 **Supplementary Table 5 – Number of channels with non-zero weights for all decoders used**950 **in both subjects.**

- 952 44 Benjamini, Y. & Hochberg, Y. *Controlling The False Discovery Rate - A Practical And*
953 *Powerful Approach To Multiple Testing*. Vol. 57 (1995).
- 954 45 Semedo, J. D., Zandvakili, A., Machens, C. K., Yu, B. M. & Kohn, A. Cortical Areas
955 Interact through a Communication Subspace. *Neuron* **102**, 249-259.e244,
956 doi:<https://doi.org/10.1016/j.neuron.2019.01.026> (2019).
- 957 46 Seideman, J. A., Stanford, T. R. & Salinas, E. Saccade metrics reflect decision-making
958 dynamics during urgent choices. *Nature Communications* **9**, 2907, doi:10.1038/s41467-
959 018-05319-w (2018).
- 960 47 Musall, S., Kaufman, M. T., Juavinett, A. L., Gluf, S. & Churchland, A. K. Single-trial
961 neural dynamics are dominated by richly varied movements. *bioRxiv*, 308288,
962 doi:10.1101/308288 (2019).
- 963 48 Murphy, P. R., Boonstra, E. & Nieuwenhuis, S. Global gain modulation generates time-
964 dependent urgency during perceptual choice in humans. *Nat Commun* **7**, 13526,
965 doi:10.1038/ncomms13526 (2016).
- 966 49 Niyogi, R. K. & Wong-Lin, K. Dynamic Excitatory and Inhibitory Gain Modulation Can
967 Produce Flexible, Robust and Optimal Decision-making. *PLoS Computational Biology* **9**,
968 e1003099, doi:10.1371/journal.pcbi.1003099 (2013).
- 969 50 Sussillo, D., Stavisky, S. D., Kao, J. C., Ryu, S. I. & Shenoy, K. V. Making brain-
970 machine interfaces robust to future neural variability. *Nature Communications* **7**, 13749,
971 doi:10.1038/ncomms13749.
972 <https://www.nature.com/articles/ncomms13749#supplementary-information> (2016).
- 973 51 Gold, J. I. & Shadlen, M. N. Representation of a perceptual decision in developing
974 oculomotor commands. *Nature* **404**, 390-394, doi:10.1038/35006062 (2000).
- 975 52 Spivey, M. J., Grosjean, M. & Knoblich, G. Continuous attraction toward phonological
976 competitors. *Proceedings of the National Academy of Sciences of the United States of*
977 *America* **102**, 10393, doi:10.1073/pnas.0503903102 (2005).
- 978 53 Selen, L. P. J., Shadlen, M. N. & Wolpert, D. M. Deliberation in the Motor System: Reflex
979 Gains Track Evolving Evidence Leading to a Decision. *The Journal of Neuroscience* **32**,
980 2276, doi:10.1523/JNEUROSCI.5273-11.2012 (2012).
- 981 54 Churchland, M. M., Yu, B. M., Ryu, S. I., Santhanam, G. & Shenoy, K. V. Neural
982 Variability in Premotor Cortex Provides a Signature of Motor Preparation. *The Journal of*
983 *Neuroscience* **26**, 3697, doi:10.1523/JNEUROSCI.3762-05.2006 (2006).
- 984 55 Afshar, A. *et al.* Single-Trial Neural Correlates of Arm Movement Preparation. *Neuron*
985 **71**, 555-564, doi:10.1016/j.neuron.2011.05.047 (2011).
- 986 56 Kaufman, M. T. *et al.* Roles of monkey premotor neuron classes in movement
987 preparation and execution. *Journal of neurophysiology* **104**, 799-810,
988 doi:10.1152/jn.00231.2009 (2010).
- 989 57 Kaufman, M. T., Churchland, M. M. & Shenoy, K. V. The roles of monkey M1 neuron
990 classes in movement preparation and execution. *Journal of neurophysiology* **110**, 817-
991 825, doi:10.1152/jn.00892.2011 (2013).
- 992 58 Kaufman, M. T., Churchland, M. M., Ryu, S. I. & Shenoy, K. V. Cortical activity in the null
993 space: permitting preparation without movement. *Nature Neuroscience* **17**, 440-448,
994 doi:10.1038/nn.3643 (2014).
- 995 59 Kaufman, M. T. *et al.* The Largest Response Component in the Motor Cortex Reflects
996 Movement Timing but Not Movement Type. *eneuro* **3**, ENEURO.0085-0016.2016,
997 doi:10.1523/ENEURO.0085-16.2016 (2016).

998 60 Elsayed, G. F., Lara, A. H., Kaufman, M. T., Churchland, M. M. & Cunningham, J. P.
999 Reorganization between preparatory and movement population responses in motor
1000 cortex. *Nature Communications* **7**, 13239, doi:10.1038/ncomms13239
1001 <https://www.nature.com/articles/ncomms13239#supplementary-information> (2016).