1 **TITLE**

2 Slowly evolving dopaminergic activity modulates the moment-to-moment probability of reward-

3 related self-timed movements.

4

5 **AUTHOR NAMES**

6 Allison E. Hamilos[1*], Giulia Spedicato[1], Ye Hong[1], Fangmiao Sun[2], Yulong Li[2] &

7 John A. Assad[1,3*]

8

9 **AFFILIATIONS**

10 [1]Department of Neurobiology, Harvard Medical School, Boston, Massachusetts, 02115, USA.

11 [2]State Key Laboratory of Membrane Biology, Peking University School of Life Sciences,

12 Beijing, 100871, P.R. China.

13 [3]Istituto Italiano di Tecnologia, Genova, Italy.

14 *For correspondence: A.E.H. (allison_hamilos@hms.harvard.edu) or J.A.A.

15 (jassad@hms.harvard.edu).

16

17 **ABSTRACT**

18 Clues from human movement disorders have long suggested that the neurotransmitter dopamine

19 plays a role in motor control, but how the endogenous dopaminergic system influences

20 movement is unknown. Here we examined the relationship between dopaminergic signaling and

21 the timing of reward-related movements in mice. Animals were trained to initiate licking after a

22 self-timed interval following a start-timing cue; reward was delivered in response to movements

23 initiated after a criterion time. The movement time was variable from trial-to-trial, as expected

24 from previous studies. Surprisingly, dopaminergic signals ramped-up over seconds between the

25 start-timing cue and the self-timed movement, with variable dynamics that predicted the

26 movement/reward time on single trials. Steeply rising signals preceded early lick-initiation,

27 whereas slowly rising signals preceded later initiation. Higher baseline signals also predicted

28 earlier self-timed movements. Optogenetic activation of dopamine neurons during self-timing

29 did not trigger immediate movements, but rather caused systematic early-shifting of movement

30 initiation, whereas inhibition caused late-shifting, as if modulating the probability of movement.

31 Consistent with this view, the dynamics of the endogenous dopaminergic signals quantitatively

32 predicted the moment-by-moment probability of movement initiation on single trials. We

33 propose that ramping dopaminergic signals, likely encoding dynamic reward expectation, can

34 modulate the decision of when to move.

35

36 **INTRODUCTION**

37 What makes us move? Empirically, a few hundred milliseconds before movement, thousands of

38 neurons in the motor system suddenly become active in concert, and this neural activity is

39 relayed via spinal and brainstem neurons to recruit muscle fibers that power movement (*Shenoy*

40 *et al., 2013*). Yet just before this period of intense neuronal activity, the motor system is largely

41 quiescent. How does the brain suddenly and profoundly rouse motor neurons into the

42 coordinated action needed to trigger movement?

43

44 In the case of movements made in reaction to external stimuli, activity evoked first in sensory

45 brain areas is presumably passed along to appropriate motor centers to trigger this coordinated

46 neural activity, thereby leading to movement. But humans and animals can also self-initiate

47　movement without overt, external input (*Deecke, 1996; Hallett, 2007; Lee and Assad, 2003;*

48　*Romo et al., 1992*). For example, while reading this page, you may decide without prompting to

49　reach for your coffee. In that case, the movement cannot be clearly related to an abrupt,

50　conspicuous sensory cue. What "went off" in your brain that made you reach for your coffee at

51　this *particular* moment, as opposed to a moment earlier or later?

52

53　Human movement disorders may provide clues to this mystery. Patients and animal models of

54　Parkinson's Disease experience difficulty self-initiating movements, exemplified by

55　perseveration (*Hughes et al., 2013*), trouble initiating steps when walking (*Bloxham et al.,*

56　*1984*), and problems timing movements (*Malapani et al., 1998; Meck, 1986, 2006; Mikhael and*

57　*Gershman, 2019*). In contrast to these self-generated actions, externally cued reactions are often

58　less severely affected in Parkinson's, a phenomenon sometimes referred to as "paradoxical

59　kinesia" (*Barthel et al., 2018; Bloxham et al., 1984*). For example, patients' gait can be

60　normalized by walking aids that prompt steps in reaction to visual cues displayed on the ground

61　(*Barthel et al., 2018*).

62

63　Because the underlying neuropathophysiology of Parkinson's includes the loss of midbrain

64　dopaminergic neurons (DANs), the symptomatology of Parkinson's suggests DAN activity plays

65　an important role in deciding when to self-initiate movement. Indeed, pharmacological

66　manipulations of the neurotransmitter dopamine causally and bidirectionally influence

67　movement timing (*Dews and Morse, 1958; Lustig and Meck, 2005; Meck, 1986; Mikhael and*

68　*Gershman, 2019; Schuster and Zimmerman, 1961*). This can be demonstrated in the context of

69　*self-timed* movement tasks, in which subjects reproduce a target-timing interval by making a

70 movement following a self-timed delay that is referenced to a start-timing cue (*Malapani et al.,*

71 *1998*). Species across the animal kingdom, from rodents and birds to primates, can learn these

72 tasks and produce self-timed movements that occur, on average, at about the target time,

73 although the exact timing exhibits considerable variability from trial-to-trial (*Gallistel and*

74 *Gibbon, 2000; Meck, 2006; Mello et al., 2015; Merchant et al., 2013; Rakitin et al., 1998;*

75 *Remington et al., 2018; Schuster and Zimmerman, 1961; Sohn et al., 2019; Wang et al., 2018*).

76 In such self-timed movement tasks, decreased dopamine availability/efficacy (*e.g.*, Parkinson's,

77 neuroleptic drugs) generally produces late-shifted movements (*Malapani et al., 1998; Meck,*

78 *1986, 2006; Merchant et al., 2013*), whereas high dopamine conditions (*e.g.*, amphetamines)

79 produce early-shifting (*Dews and Morse, 1958; Schuster and Zimmerman, 1961*).

80

81 Although exogenous dopamine manipulations can influence timing behavior, it remains unclear

82 whether endogenous DAN activity is involved in determining when to move. DANs densely

83 innervate the striatum, where they modulate the activity of spiny projection neurons of the direct

84 and indirect pathways, which are thought to exert a push-pull influence on movement centers

85 (*Albin et al., 1989; DeLong, 1990; Freeze et al., 2013; Grillner and Robertson, 2016*). Most

86 studies on endogenous DAN activity have focused on reward-related signals, but there are also

87 reports of movement-related DAN signals. For example, phasic bursts of dopaminergic activity

88 have been observed just prior to movement onset (within ~500 ms; *Coddington and Dudman,*

89 *2018, 2019; da Silva et al., 2018; Dodson et al., 2016; Howe and Dombeck, 2016; Wang and*

90 *Tsien, 2011*), and dopaminergic signals have been reported to reflect more general encoding of

91 movement kinematics (*Barter et al., 2015; Engelhard et al., 2019; Parker et al., 2016*).

92 However, optogenetic activation of dopamine neurons—within physiological range—does not

93   elicit immediate movements (*Coddington and Dudman, 2018, 2019*). We hypothesized that

94   rather than overtly triggering movements, the ongoing activity of nigrostriatal DANs could

95   influence movement initiation over longer timescales by controlling or modulating the moment-

96   by-moment decision of *when* to execute a planned movement.

97

98   To test this hypothesis, we trained mice to make a movement (lick) after a self-timed interval

99   following a start-timing cue. The mice learned the timed interval, but, as observed in other

100   species, the exact timing of movement was highly variable from trial-to-trial, spanning seconds.

101   We exploited this inherent variability by examining how moment-to-moment nigrostriatal DAN

102   signals differed when animals decided to move relatively early versus late. We found that

103   dopaminergic signals "ramped up" during the timing interval, with variable dynamics that were

104   highly predictive of trial-by-trial movement timing, even seconds before the movement occurred.

105   Because reward was delivered at the time of movement, the ramping dopaminergic signals likely

106   related to the animal's expectation of when reward would be available in response to movement.

107   Furthermore, optogenetic DAN manipulation during the timing interval produced bidirectional

108   changes in the probability of movement timing, with activation causing a bias toward earlier self-

109   timed movements and suppression causing a bias toward later self-timed movements. These

110   combined observations suggest a novel role for the dopaminergic system in the timing of

111   movement initiation, wherein slowly evolving dopaminergic signals, likely driven by reward

112   expectation, can modulate the moment-to-moment probability of whether a reward-related

113   movement will occur.

114

RESULTS

We trained head-fixed mice to make self-timed movements to receive juice rewards (*Figure 1A*). Animals received an audio/visual start-timing cue and then had to decide when to first-lick in the absence of further cues. Animals only received juice if they waited a proscribed interval following the cue before making their first-lick (>3.3 s in most experiments). As expected from previous studies, the distribution of first-lick timing was broadly distributed over several seconds, and exhibited the canonical scalar property of timing, as described by Weber's Law (*Figure 1B* and *Figure 1—figure supplement 1A-B; Gallistel and Gibbon, 2000*). We note this variability in timing was not imposed on the animal by training it to reproduce a variety of target intervals (*e.g.*, 2 *vs.* 5 s), but is rather a natural consequence of timing behavior, even for a single target interval.

Our main objective was to exploit the inherent variability in self-timed behavior to examine how differences in neural activity might relate to variability in movement timing. Nonetheless, the trained animals well-understood the timing contingencies of the task. In self-timed movement tasks in which a *single* movement is used to assess timing, the distributions of movement times (in both rodents and monkeys) tend to anticipate the target interval, even at the expense of reward on many trials (*Eckard and Kyonka, 2018; Kirshenbaum et al., 2008; Lee and Assad, 2003*). In these paradigms, however, once a movement occurs, it removes future opportunities to move, which creates premature "bias" in the raw timing distributions (*Anger, 1956*). To correct this bias, movement times must be normalized by the (ever-diminishing) number of opportunities to move at each timepoint (*Jaldow et al., 1990*). This yields the hazard function (the conditional probability of movement given that movement has not already occurred, as a function of time),

138 which is equivalent to the instantaneous probability of movement. For example, on the first day

139 of training, our animals displayed fairly flat hazard functions, indicating a uniform instantaneous

140 probability of movement over time—*i.e.*, the animals did not yet understand the timing

141 contingency (*Figure 1C-D*). However, after training, the hazard function for our animals peaked

142 near the target time (either 3.3 or 5 s), suggesting an accurate latent timing process reflected in

143 the instantaneous movement probability (*Figure 1E*). Mice trained on a variant of the self-timed

144 movement task without lamp-off/on events showed no systematic differences in their timing

145 distributions (*Figure 1—figure supplement 1C*), suggesting that the mice referenced their timing

146 to the start-timing cue rather than the lamp-off event.

147

148 When mice were fully trained, we employed fiber photometry to record the activity of

149 genetically-defined DANs expressing the calcium-sensitive fluorophore GCaMP6f (12 mice,

150 substantia nigra pars compacta (SNc); *Figure 1—figure supplement 2*). We controlled for

151 mechanical/optical artifacts by simultaneously recording fluorescence modulation of a co-

152 expressed, calcium-insensitive fluorophore, tdTomato. We also recorded bodily movements with

153 neck-muscle EMG, high-speed video, and a back-mounted accelerometer.

154

155 **DAN signals ramp up slowly between the start-timing cue and self-timed movement**

156 DAN GCaMP6f fluorescence typically exhibited brief transients following cue onset and

157 immediately before first-lick onset (*Figure 2A*), as observed in previous studies (*Coddington*

158 *and Dudman, 2018; da Silva et al., 2018; Dodson et al., 2016; Howe and Dombeck, 2016;*

159 *Schultz et al., 1997*). However, during the timed interval, we observed slow "ramping up" of

160 fluorescence over seconds, with a minimum after the cue-aligned transient and maximum just

161   before the lick-related transient. The relatively fast intrinsic decay kinetics of GCaMP6f ($t_{1/2}$

162   <100 ms at 37°; *Helassa et al., 2016*) should not produce appreciable signal integration over the

163   seconds-long timescales of the ramps we observed.

164

165   We asked whether this ramping differed between trials in which the animal moved relatively

166   early or late. Strikingly, when we averaged signals pooled by movement time, we observed

167   systematic differences in the steepness of ramping that were highly predictive of movement

168   timing (*Figure 2B-C*). Trials with early first-licks exhibited steep ramping, whereas trials with

169   later first-licks started from lower fluorescence levels and rose more slowly toward the time of

170   movement. The fluorescence ramps terminated at nearly the same amplitude, regardless of the

171   movement time. Ramping dynamics were not evident in control tdTomato signals (*Figure 2C*),

172   indicating that the ramping in the GCaMP6f signals was not an optical artifact. The quantitative

173   relationship between GCaMP6f dynamics and movement time will be addressed in a subsequent

174   section of this paper.

175

176   **Higher pre-cue DAN signals are correlated with earlier self-timed movements**

177   In addition to ramping dynamics, average DAN GCaMP6f signals were correlated with first-lick

178   timing even before cue-onset, with higher baseline fluorescence predicting earlier first-licks

179   (*Figure 2B-C*). This correlation began before the lamp-off event (the 2 s "Baseline" period

180   before lamp-off; Pearson's r = -0.63 (95% CI=[-0.92, -0.14]), n=12 mice) and grew stronger

181   during the "Lamp-Off Interval" between lamp-off and the cue (Pearson's r = -0.89 (95% CI=[-

182   -0.98, -0.68]), n=12 mice; *Figure 2—figure supplement 1A-B*). This correlation was

183   independent of the duration of the lamp-off interval (*Figure 2—figure supplement 1C*). Because

184   dF/F correction methods can potentially distort baseline measurements, we rigorously tested and

185   validated three different dF/F methods, and we also repeated analyses with raw fluorescence

186   values compared between pairs of sequential trials with different movement times (*Figure 2—*

187   *figure supplement 2*; see *Methods*). All reported results, including the systematic baseline

188   differences, were robust to dF/F correction.

189

190   In principle, the amplitude of the baseline signal on a given trial n could be related to the

191   animal's behavior during the baseline interval or the outcome of the previous trial. To test this,

192   we performed four-way ANOVA to compare the main effects of the following factors on the pre-

193   cue signal (averaged for each trial between lamp-off and the start-timing cue, the "lamp-off

194   interval" (LOI), n=12 mice): 1) presence or absence of spontaneous licking during the LOI; 2)

195   outcome of the previous trial (rewarded or unrewarded); 3) upcoming movement time on trial n

196   (categorized as <3.3 s or >3.3 s to provide a simple binary proxy for movement time); and 4)

197   session number (to account for signal variability across animals and daily sessions). Although the

198   effects of LOI-licking and previous trial outcome were statistically significant ($F(1,18282)=10.7$,

199    $p=0.008$, $\eta_p^2=5.9\cdot10^{-4}$ and $F(1,18282)=281.2$, $p=7.5\cdot10^{-47}$, $\eta_p^2=0.015$, respectively), the

200   upcoming movement time had an independent, statistically significant effect ($F(1,18282)=63.4$,

201    $p=5.9\cdot10^{-6}$, $\eta_p^2=0.0035$). This raises the possibility of an additional source of variance in baseline

202   dopaminergic activity that is independent from previous trial events, but potentially influences

203   the upcoming movement time on that trial.

204

205   **Ramping dynamics in other dopaminergic areas and striatal dopamine release**

206   We found similar ramping dynamics in SNc DAN axon terminals in the dorsolateral striatum

207   (DLS; *Figure 2—figure supplement 3A-B*) at a location involved in goal-directed licking

208   behavior (*Sippy et al., 2015*). Ramping was also present in GCaMP6f-expressing DAN cell

209   bodies in the ventral tegmental area (VTA, *Figure 2—figure supplement 3C*), reminiscent of

210   mesolimbic ramping signals described in goal-oriented navigation tasks (*Howe et al., 2013; Kim*

211   *et al., 2019*).

212

213   To determine if these movement timing-related signals are available to downstream targets that

214   may be involved in movement initiation, we monitored dopamine release in the DLS with two

215   complementary florescent dopamine sensors (dLight1.1 and DA_{2m}) expressed broadly in striatal

216   cells (*Figure 3 and Figure 2—figure supplement 3D-E*). The decay kinetics of the two

217   extracellular dopamine sensors differ somewhat (*Patriarchi et al., 2018; Sun et al., 2020*), which

218   we confirmed (dLight1.1 $t_{1/2}$~75 ms, DA_{2m} $t_{1/2}$~125 ms; *Figure 3—figure supplement 1*), yet

219   both revealed similar timing-related ramping dynamics on average (*Figure 3 inset*). These

220   combined data argue that the seconds-long dopaminergic ramping signals were not artifacts of

221   sluggish temporal responses of the various fluorescent sensors and were ultimately expressed as

222   ramp-like increases in dopamine release in the striatum.

223

224   **First-lick timing-predictive DAN signals are not explained by ongoing body movements**

225   The systematic ramping dynamics and baseline differences were not observed in the tdTomato

226   optical control channel nor in any of the other movement-control channels, at least on average

227   (*Figure 4*), making it unlikely that ramping dynamics resulted from optical artifacts.

228   Nevertheless, because DANs show transient responses to salient cues and movements

229 (*Coddington and Dudman, 2018; da Silva et al., 2018; Dodson et al., 2016; Howe and*

230 *Dombeck, 2016; Schultz et al., 1997*), it is possible that fluorescence signals could reflect the

231 superposition of dopaminergic responses to multiple task events, including the cue, lick, ongoing

232 spurious body movements, and hidden cognitive processes like timing. For example, accelerating

233 spurious movements could, in principle, produce motor-related neural activity that "ramps up"

234 during the timed interval, perhaps even at different rates on different trials.

235

236 We thus derived a nested generalized linear encoding model of single-trial GCaMP6f signals

237 (*Engelhard et al., 2019; Park et al., 2014; Runyan et al., 2017*), a data-driven, statistical

238 approach designed to isolate and quantify the contributions of task events (timing-independent

239 predictors) from processes predictive of movement timing (timing-dependent predictors; *Figure*

240 *5A-B* and *Figure 5—figure supplement 1A-D*). The model robustly detected task-event

241 GCaMP6f kernels locked to cue, lick and EMG/accelerometer events, but these timing-

242 independent predictors alone were insufficient to capture the rich variability of GCaMP6f signals

243 for trials with different first-lick times, especially the timing-dependent ramp-slope and baseline

244 offset (n=12 mice, *Figure 5C* and *Figure 5—figure supplement 1E-G*). In contrast, two timing-

245 dependent predictors robustly improved the model: 1) a baseline offset with amplitude linearly

246 proportional to first-lick time; and 2) a "stretch" feature representing percentages of the timed

247 interval (*Figure 5B-C* and *Figure 5—figure supplement 1E*). The baseline offset term fit a

248 baseline level inversely proportional to movement time, and the temporal stretch feature

249 predicted a ramping dynamic from the time of the cue up to the first-lick, whose slope was

250 inversely proportional to first-lick time. Similar results were obtained for SNc DAN axon

251 terminals in the DLS, VTA DAN cell bodies, and extracellular striatal dopamine release (*Figure*

252 *5—figure supplement 1H*).

253

254 We note that the stretch feature of this GLM makes no assumptions about the underlying shape

255 of the dopaminergic signal; it only encodes percentages of timing intervals to allow for temporal

256 "expansion" or "contraction" to fit whatever shape(s) were present in the data. In particular, the

257 stretch feature cannot produce ramping unless ramping is present in the signal *and* temporally

258 scales with the length of the interval. Because this feature empirically found a ramp (although

259 not constrained to do so), the stretch aspect indicated that the underlying ramping process took

260 place at different rates for trials with different movement times, at least on average.

261

262 In contrast to the GCaMP6f model, when the same GLM was applied to the tdTomato control

263 signal, the timing-independent predictors (which could potentially cause optical/mechanical

264 artifacts—cue onset, first-lick, EMG/accelerometer) improved the model, but timing-dependent

265 predictors did not (*Figure 5C* and *Figure 5—figure supplement 1F-H*). In addition, separate

266 principal component (PC) analysis revealed ramp-like and baseline-offset-like components that

267 explained as much as 93% of the variance in DAN signals during the timing interval

268 (mean: 66%, range: 16-93%), but similar PCs were not present when tdTomato control signals

269 were analyzed with PCA (mean variance explained: 4%, range: 1.6-15%, *Figure 5—figure*

270 *supplement 2*).

271

272 **Single-trial DAN ramping and baseline signals predict movement timing**

273 Given that ramping and baseline-offset signals were not explained by nuisance movements or

274 optical artifacts, we asked whether DAN GCaMP6f fluorescence could predict first-lick timing

275 on single trials. Using a simple threshold-crossing decoding model (*Maimon and Assad, 2006*),

276 we found that single-trial GCaMP6f signals were predictive of first-lick time even for low

277 thresholds intersecting the "base" of the ramp, with the predictive value of the model

278 progressively improving for higher thresholds (n=12 mice: mean R^2 low=0.54, mid=0.71,

279 high=0.82 (95% CI: low=[0.44,0.64], mid=[0.68,0.75], high=[0.76,0.87]); analysis for one

280 mouse shown in *Figure 6A*). We will return to this observation in more detail in the upcoming

281 section on single-trial dynamics.

282

283 To more thoroughly determine the independent, additional predictive power of DAN baseline

284 and ramping signals over other task variables (*e.g.*, previous trial first-lick time and reward

285 outcome, *etc*.), we derived a nested decoding model for first-lick time (*Figure 6A*). In this model,

286 the pre-cue "baseline" was divided into two components: the pre-lamp-off intertrial interval

287 signal ("ITI") and the lamp-off to cue interval signal ("LOI"). All predictors contributed to the

288 predictive power of the model. However, even when we accounted for the contributions of prior

289 trial history, tdTomato artifacts and baseline GCaMP6f signals, GCaMP6f threshold-crossing

290 time robustly dominated the model and absorbed much of the variance explained by baseline

291 dopaminergic signals, alone explaining 10% of the variance in first-lick time on average (range:

292 1-27%, *Figure 6B-D*). Alternate formulations of the decoding model produced similar results

293 (*Figure 6—figure supplement 1*).

294

295 **Characterizing single-trial dopaminergic dynamics**

296 Although the threshold-crossing analysis made no assumptions about the underlying dynamics of

297   the GCaMP6f signals on single-trials, in principle, ramping dynamics in *averaged* neural signals

298   could be produced from individual trials with a single, discrete "step" occurring at different

299   times on different trials. Ramping has long been observed in averaged neural signals recorded

300   during perceptual decision tasks in monkeys, and there has been considerable debate over

301   whether single-trial responses in these experiments are better classified as "ramps" or a single

302   "step" (*Latimer et al., 2015, 2016; Shadlen et al., 2016; Zoltowski et al., 2019; Zylberberg and*

303   *Shadlen, 2016*). It has even been suggested that different sampling distributions can produce

304   opposite model classifications in ground-truth synthetic datasets (*Chandrasekaran et al., 2018*).

305

306   We attempted to classify single-trial dynamics as a discrete stepping or ramping process with

307   hierarchical Bayesian models implemented in probabilistic programs (*Figure 6—figure*

308   *supplement 2A-B*). However, like the perceptual decision-making studies, we also found

309   ambiguous results, with about half of single-trials best classified by a linear ramp and half best

310   classified by a discrete step dynamic (*Figure 6—figure supplement 2C*). Nonetheless, three

311   separate lines of evidence suggest that single trials are better characterized by slowly evolving

312   ramps:

313

314   First, the relationship of threshold-crossing time to first-lick time is different for the step *vs.*

315   ramp models when different threshold levels are sampled (*Maimon and Assad, 2006*), as

316   schematized in *Figure 6—figure supplement 3A*: Increasing slope of this relationship is

317   consistent with ramps on single trials, but not with a discrete step, which would be expected to

318   have the same threshold-crossing time regardless of threshold level (*Figure 6—figure*

319   *supplement 3B*). We found that the slope of this relationship increased markedly as the threshold

320 level was increased, consistent with the ramp model (n=12 mice: mean slope low=0.46, mid=0.7,

321 high=0.82 (95% CI: low=[0.37,0.54], mid=[0.66,0.73], high=[0.74,0.88]), *Figure 6—figure*

322 *supplement 3C*).

323

324 Second, if single trials involve a step change occurring at different times from trial-to-trial, then

325 aligning trials on that step should produce a clear step on average (rather than a ramp; *Latimer et*

326 *al., 2015*). We thus aligned single-trial GCaMP6f signals according to that optimal step position

327 determined from a Bayesian step model fit for each trial and then averaged the step-aligned

328 signals across trials. The averaged signals did not resemble a step function, but rather yielded a

329 sharp transient superimposed on a "background" ramping signal (*Figure 6—figure supplement*

330 *4A*). Step-aligned tdTomato and EMG averages showed a small inflection at the time of the step,

331 but neither signal showed background ramping. This suggests that the detected "steps" in the

332 GCaMP6f signals were likely transient movement artifacts superimposed on the slower ramping

333 dynamic rather than *bona fide* steps.

334

335 Third, the ideal step model holds that the step occurs at different times from trial-to-trial,

336 producing a ramping signal when trials are averaged together. In this view, the trial-by-trial

337 variance of the signal should be maximal at the time at which 50% of the steps have occurred

338 among all trials, and the signal should be minimal at the beginning and end of the interval (when

339 no steps or all steps have occurred, respectively). We thus derived the optimal step time for each

340 trial using the Bayesian step model, and then calculated variance as a function of time within

341 pools of trials with similar movement times. The signal variance showed a monotonic downward

342 trend during the timed interval, with a minimum variance at the time of movement rather than at

343 the point at which 50% of steps had occurred among trials, inconsistent with the discrete step

344 model (*Figure 6—figure supplement 4B*).

345

346 Taken together, we did not find evidence for a discrete step dynamic on single trials; on the

347 contrary, our observations concord with slow ramping dynamics on single trials. Regardless, our

348 GLM movement-time decoding approaches in *Figure 6* did not make any assumptions about

349 underlying single-trial dynamics.

350

351 **Moment-to-moment DAN activity causally controls movement timing**

352 Because dopaminergic ramping signals robustly predicted first-lick timing and were apparently

353 transmitted via dopamine release to downstream striatal neurons, ramping DAN activity may

354 causally determine movement timing. However, because the animals could expect reward within

355 a few hundred milliseconds of the first-lick, it is also possible that the dopaminergic ramps could

356 instead serve as a "passive" monitor of reward expectation without influencing movement

357 initiation. To distinguish these possibilities, we optogenetically activated or inhibited DANs (in

358 separate experiments) on 30% of randomly-interleaved trials (*Figure 7A* and *Figure 7—figure*

359 *supplement 1*). For activation experiments, we chose light levels that elevated DAN activity

360 within the physiological range observed in our self-timed movement task, as assayed by

361 simultaneous photometry in the DLS with a fluorescent sensor of released dopamine (dLight1.1,

362 *Figure 7—figure supplement 2*). DAN activation significantly early-shifted the distribution of

363 self-timed movements on stimulated trials compared to unstimulated trials (12 mice; 2-sample

364 Kolmogorov-Smirnov (KS) Test, D=0.078 (95% CI: [0.067,0.093]), p=$2.8 \cdot 10^{-26}$), whereas

365 inhibition produced significant late-shifting compared to unstimulated trials (4 mice; 2-sample

366 KS Test, D=0.051 (95% CI: [0.034,0.077]), p=$3.1 \cdot 10^{-4}$; *Figure 7B and Figure 7—figure*

367 *supplement 3A*). Stimulation of mice expressing no opsin produced no consistent effect on

368 timing (5 mice; 2-sample KS Test, D=0.017 (95% CI: [0.015,0.040]), p=0.62). The direction of

369 these effects was consistent across all animals tested in each category (*Figure 7B*).

370 Complementary analysis methods revealed consistent effects (bootstrapped difference in median

371 first-lick times between categories: Δ(activation - no-opsin) = -0.22 s (95% CI=[-0.32 s,-0.12

372 s]), Δ(inhibition – no-opsin) = +0.19 s (95% CI=[+0.09 s,+0.30 s]), *Figure 7C-D*; bootstrapped

373 comparison of difference in area under the cdf curves: Δ(activation – no-opsin) = -0.31 dAUC

374 (95% CI=[-0.47 dAUC,-0.15 dAUC]), Δ(inhibition – no-opsin) = +0.23 dAUC (95%

375 CI=[+0.08 dAUC,+0.37 dAUC]), *Figure 7—figure supplement 3B*; bootstrapped difference in

376 mean first-lick times between categories: Δ(activation – no-opsin) = -0.34 s (95%

377 CI=[-0.49 s,-0.19 s]), Δ(inhibition – no-opsin) = +0.24 s (95% CI=[+0.09 s,+0.39 s]), *Figure 7—*

378 *figure supplement 3C*). Similar effects were obtained with activation of SNc DAN axon

379 terminals in the DLS (2 mice, *Figure 7—figure supplement 3A-B*). Because these exogenous

380 manipulations of DAN activity modulated movement timing on the same trial as the

381 stimulation/inhibition, this suggests that the *endogenous* dopaminergic ramping we observed

382 during the self-timed movement task likewise affected movement initiation in real time, rather

383 than serving solely as a passive monitor of reward expectation.

384

385 Recent studies have shown that physiological ranges of optogenetic DAN activation (as assayed

386 by simultaneous recordings from DANs) fail to elicit overt movements (*Coddington and*

387 *Dudman, 2018*). We likewise found that optogenetic DAN activation did not elicit immediate

388 licking outside the context of the task (*Figure 7—figure supplement 4A*). Additionally,

389   optogenetic DAN inhibition did not reduce the rate of spontaneous licking outside the context of

390   the task (*Figure 7—figure supplement 4B*). In both cases, we used the same light levels that had

391   elicited the robust shifts in timing behavior during the self-timed movement task. In other control

392   experiments, we purposefully drove neurons into non-physiological activity regimes during the

393   task by applying higher activation light levels. Over-stimulation caused large, immediate,

394   sustained increases in DLS dopamine (*Figure 7—figure supplement 2*), comparable in

395   amplitude to the typical reward-related dopamine transients on interleaved, unstimulated trials.

396   These non-physiological manipulations resulted in rapid, nonpurposive body movements and

397   disrupted performance of the task. Together, these results suggest that the optogenetic effects on

398   timing in *Figure 7* did not result from direct, immediate triggering or suppression of movement,

399   nor from non-physiological dopamine release due to over-stimulation.

400

401   **Linking endogenous DAN signals to the moment-to-moment probability of movement**

402   **initiation**

403   Optogenetic manipulations of DAN activity in the physiological range appeared to modulate the

404   *probability* of initiating the pre-potent, self-timed movement. Given that endogenous DAN

405   signals increased during the timing interval of the self-timed movement task, we reasoned that

406   the probability of movement should likewise increase over the course of the timed interval. We

407   thus derived a nested probabilistic movement-state decoding model to explore the link between

408   DAN signals and movement propensity (*Figure 8A*). We applied a GLM based on logistic

409   regression, in which we classified each moment of time as either a non-movement (0) or

410   movement (1) state (*Figure 8A-B*), and we examined how well various parameters could predict

411   the probability of transitioning from the non-movement state to the movement state. Unlike the

412   decoding model in **Figure 6**, which considers a single threshold-crossing time, the probabilistic

413   approach takes into account continuous DAN signals. Initial model selection included previous

414   trial history (movement time and reward outcome history) in addition to the DAN GCaMP6f

415   signal, but Bayesian Information Criterion (BIC) analysis indicated that the instantaneous

416   GCaMP6f signal alone was a robustly significant predictor of movement state, whereas previous

417   trial outcomes were insignificant contributors and did not further improve the model (*Figure 8—*

418   *figure supplement 1*). We thus only considered the DAN GCaMP6f signal as a predictor in

419   subsequent analyses.

420

421   The continuous DAN GCaMP6f signal was indeed predictive of current movement state at any

422   time *t*, and it served as a significant predictor of movement state, up to at least 2 seconds in the

423   past (*Figure 8C*). However, the signals became progressively more predictive of the current

424   movement state as time approached *t*. That is, the dopaminergic signal levels closer to time *t*

425   tended to absorb the behavioral variance explained by more distant, previous signal levels

426   (*Figure 8C*), reminiscent of how threshold crossing time absorbed the variance explained by the

427   baseline dopaminergic signal in the movement-timing decoding model (*Figure 6B-C*). This

428   observation is consistent with a diffusion-like ramping process on single trials, in which the most

429   recent measurement gives the best estimate of whether there will be a transition to the movement

430   state (but is difficult to reconcile with a discrete step process on single trials, consistent with the

431   results in *Figure 6—figure supplements 3-4*).

432

433   We applied the fitted instantaneous probabilities of transitioning to the movement state to derive

434   a fitted hazard function for each behavioral session (*Figure 8D*). The DAN GCaMP6f signals

435 were remarkably predictive of the hazard function, both for individual sessions and on average,

436 explaining 65% of the variance on average (n=12 mice). Conversely, when the model was fit on

437 the same data in which the timepoint identifiers were shuffled, this predictive power was

438 essentially abolished, explaining only 5% of the variance on average (*Figure 8E*).

439

440 Together, these results demonstrate that slowly evolving dopaminergic signals are predictive of

441 the moment-to-moment probability of movement initiation. When combined with the

442 optogenetics results, they argue that dopaminergic signals causally modulate the moment-to-

443 moment probability of the pre-potent movement. In this view, trial-by-trial variability in the

444 DAN signal gives rise to trial-by-trial differences in movement timing in the self-timed

445 movement task.

446

447 **DISCUSSION**

448 We made two main findings. First, both baseline and slowly ramping DAN signals were

449 predictive of the timing of the first-lick. Second, optogenetic modulation of DANs affected the

450 timing of movements on the *concurrent trial*, suggesting that DANs can play a "real time" role

451 in behavior. These observations raise two (presumably separable) questions of interpretation: 1)

452 what is the mechanistic origin of ramping DAN signals in the self-timed movement task, and 2)

453 how do DAN signals affect self-timed movements in real time?

454

455 **The origin of ramping DAN signals**

456 A number of studies have reported short-latency (<500 ms) modulations in DAN activity

457 following reward-predicting sensory cues and immediately preceding movements (*Coddington*

458 *and Dudman, 2018; da Silva et al., 2018; Dodson et al., 2016; Howe and Dombeck, 2016;*

459 *Schultz et al., 1997*), similar to the sensory- and motor-related transients we observed within

460 ~500 ms of the cue and first-lick. However, the ramping DAN signals we observed during self-

461 timing were markedly slower, unfolding over *seconds* and preceding the first-lick by as long as

462 10 s. Previous studies have reported similarly slow ramping dopaminergic signals in other

463 behavioral contexts, including goal-directed navigation toward rewarded targets (*Howe et al.,*

464 *2013*); multi-step tasks leading to reward (*Hamid et al., 2016; Howard et al., 2017; Mohebi et*

465 *al., 2019*); and passive observation of dynamic visual cues indicating proximity to reward (*Kim*

466 *et al., 2019*). A common feature in these experiments and our self-timed movement task is that

467 trials culminated in the animal's receiving reward. Thus, parsimony suggests that dopaminergic

468 ramping could reflect reward expectation. However, dopaminergic ramping is generally *absent*

469 in Pavlovian paradigms, in which animals learn to expect passive reward delivery at a fixed time

470 following a conditioned stimulus (*Menegas et al., 2015; Tian et al., 2016; Schultz et al., 1997;*

471 *Starkweather et al., 2017*). (One exception is a report of ramping activity in monkey DANs

472 during a Pavlovian paradigm with reward uncertainty (*Fiorillo et al., 2003*); however, ramping

473 was not subsequently reproduced under similar conditions, either in monkeys (*Fiorillo, 2011;*

474 *Matsumoto and Hikosaka, 2009; Tobler et al., 2005*) or rodents (*Hart et al., 2015; Tian and*

475 *Uchida, 2015*). Thus, while dopaminergic ramping is likely related to reward expectation, the

476 preponderance of evidence suggests that reward expectation *alone* is insufficient to cause DAN

477 ramping.

478

479 To reconcile these disparate findings, Gershman and colleagues proposed a formal model in

480 which dopaminergic ramping encodes reward expectation in the form of an "ongoing" reward-

481   prediction error (RPE) that arises from resolving uncertainty of one's position in the value

482   landscape (*i.e.*, one's spatial-temporal distance to reward delivery/omission). For example,

483   uncertainty is resolved if animals are provided visuospatial cues indicating proximity to reward

484   (*Howe et al., 2013; Kim et al., 2019*). In contrast, because animals can only imprecisely estimate

485   the passage of time, the animal is uncertain of when reward will be delivered/omitted in

486   Pavlovian tasks. The RPE model holds that this temporal uncertainty flattens the Pavlovian value

487   landscape, thereby flattening dopaminergic ramping to the degree that it is obscured (*Gershman,*

488   *2014*; *Kim et al., 2019; Mikhael and Gershman, 2019; Mikhael et al., 2019*). Though both our

489   task and Pavlovian tasks involve timing, the key difference may be that the animal actively

490   determines when reward will be delivered/omitted in the self-timed movement task—just after it

491   moves. Certainty in the timing of reward relative to its own movement would resolve the

492   animal's uncertainty of its position in the value landscape, and may thus explain why

493   dopaminergic ramping occurs prominently in the self-timed movement task, but not in Pavlovian

494   tasks (*Hamilos and Assad, 2020*). Though the RPE model provides a plausible explanation for

495   our findings, dopaminergic ramping signals are also consistent with broader views of "reward

496   expectation," such as tracking value as animals approach reward (*Hamid et al., 2016; Mohebi et*

497   *al., 2019*). In a companion theoretical paper (*Hamilos and Assad, 2020)*, we explore the reward-

498   expectation-based computational framework in more detail, including a reconciliation of

499   apparently contradictory DAN signals reported in the context of a perceptual timing task (*Soares*

500   *et al., 2016*).

501

How do DAN signals affect movement in real time?

We found that trial-by-trial variability in ramping dynamics explained the precise timing of self-timed licks. However, because the animals could expect reward shortly after the first-lick, the ramping dopaminergic signal might serve as a passive monitor of reward expectation rather than causally influencing the timing of movement initiation. To distinguish these possibilities, we optogenetically manipulated SNc DAN activity. We found that exciting or inhibiting DANs altered the timing of the first-lick on the *concurrent* trial, in a manner suggesting an increase/decrease in the probability of movement, respectively. This suggests that endogenous DAN signaling could play a causal role in the initiation of reward-related movements in real time—but by what mechanism?

One possibility is that endogenous or exogenous DAN signals could increase the animal's motivation or heighten its expectation of reward, which then secondarily influences reward-related movement. There is some evidence that might support this view. Phillips *et al.* found that electrical stimulation of the VTA in rats elicited approach behavior for self-delivery of intravenous cocaine; however, the electrical stimulation could have activated non-DAN fibers/pathways *via* the VTA (*Phillips et al., 2003*). Hamid *et al.* found that selective optogenetic stimulation of DANs could shorten the latency for rats to engage in a port-choice task—but only if the rat was disengaged from the task; if the rat was already engaged in task performance, the latency became slightly *longer* (*Hamid et al., 2016*).

In contrast to these equivocal findings, a large body of evidence suggests that selective optogenetic stimulation or inhibition of DANs generally does *not* affect reward-related

525 movements *on the same trial*. First, we ourselves could not evoke licking (nor inhibit

526 spontaneous licking) outside the context of our self-timed movement task (*Figure 7—figure*

527 *supplement 4*). Our mice were thirsty and perched near their usual juice tube, but offline DAN

528 stimulation/inhibition did not alter licking behavior, even though we applied the same optical

529 power that altered movement probability during the self-timed movement task. Numerous studies

530 have also examined the effects of optogenetic modulation of DANs in Pavlovian conditioning

531 paradigms, with the general finding that DAN modulation affects conditioned movements on

532 *subsequent* trials or sessions—a learning effect—but not on the *same* trial (*Coddington and*

533 *Dudman, 2018, 2021; Lee et al., 2020; Maes et al., 2020; Morrens et al., 2020; Pan et al., 2021;*

534 *Saunders et al., 2018*). For example, Lee *et al.* found that optogenetic inhibition of mouse DANs

535 at the same time as an olfactory conditioned stimulus had no effect on anticipatory licking on the

536 concurrent trial, even though inhibition at the time of reward delivery reduced the probability

537 and rate of anticipatory licking on subsequent trials (*Lee et al., 2020*). Thus, the preponderance

538 of evidence argues against a simple scheme whereby modulating DAN activity leads to a change

539 in motivation that automatically evokes or suppresses reward-related movements in real time.

540 The fact that we observed robust, concurrent optogenetic modulation of movement timing in our

541 experiment suggests that additional factors were at play for self-timed movements.

542

543 One possibility is that during self-timing, exogenous (optogenetic) stimulation of DANs summed

544 with the endogenous ramping DAN signal, leading to supra-heightened motivation to obtain

545 reward. However, when we deliberately over-stimulated DANs—eliciting even higher dopamine

546 signals in the DLS (*Figure 7—figure supplement 2*)—we observed "dyskinetic" body

547 movements rather than purposive licking. An alternative possibility is that the explicit timing

548   requirement of the self-timed movement task made it particularly responsive to dopaminergic

549   modulation. A long history of pharmacological and lesion experiments suggests that the

550   dopaminergic system modulates timing behavior (*Meck, 2006; Merchant et al., 2013*). Broadly

551   speaking, conditions that increase/decrease dopamine availability or efficacy speed/slow the

552   "internal clock," respectively (*Dews and Morse, 1958; Mikhael and Gershman, 2019; Schuster*

553   *and Zimmerman, 1961*; *Malapani et al., 1998; Meck, 1986, 2006; Merchant et al., 2013*). The

554   dopaminergic ramping signals we observed also bear resemblance to Pacemaker-Accumulator

555   models of neural timing, in which a hypothetical accumulator signals that an interval has elapsed

556   when it reaches a threshold level (*Gallistel and Gibbon, 2000; Lustig and Meck, 2005; Meck,*

557   *2006*). To "self-time" a movement also implies that the movement is prepared and pre-potent

558   during the timing period, potentially making the relevant neural motor circuits more sensitive to

559   dopaminergic modulation.

560

561   Regardless of the detailed mechanism, our results provide a link between dopaminergic signaling

562   and the initiation of self-timed movements. Though endogenous dopaminergic ramping likely

563   reflects reward expectation, we propose that these reward-related ramping signals can influence

564   the timing of movement initiation, at least in certain behavioral contexts. This framework also

565   provides a link between two seemingly disparate roles that have been proposed for the

566   dopaminergic system—reward/reinforcement-learning on one hand, and movement modulation

567   on the other.

568

569   Importantly, we are not suggesting that DANs *directly* drive movement (like corticospinal or

570   corticobulbar neurons). To the contrary, outside of the context of the self-timed movement task,

571   we could not evoke reward-related movements by activating DANs. Even during the self-timed

572   movement task, DAN stimulation did not elicit *immediate* movements: first-lick times still

573   spanned a broad distribution from trial-to-trial. Moreover, dopaminergic ramping does not

574   invariably lead to movement. For example, Kim *et al.* found dopaminergic ramping in the

575   presence of visual cues that signaled proximity to reward, independent of reward-related

576   movements (*Kim et al., 2019*). Consequently, we propose that when a movement is pre-potent

577   (as in our self-timed movement task), dopaminergic signaling can modulate the *probability* of

578   initiating that movement. Consistent with this view, we found that the endogenous ramping

579   dynamics were highly predictive of the moment-by-moment probability of movement (as

580   captured by the hazard function), with DAN signals becoming progressively better predictors as

581   the time of movement onset approached.

582

583   This view of dopaminergic modulation of movement probability could be related to classic

584   findings from extrapyramidal movement disorders, in which dysfunction of the nigrostriatal

585   pathway produces aberrations in movement initiation rather than paralysis or paresis (*Bloxham*

586   *et al., 1984; Fahn, 2011; Hallett and Khoshbin, 1980*). That is, movements *do* occur in

587   extrapyramidal disorders, but at inappropriate times, either too little/late (*e.g.*, Parkinson's), or

588   too often (*e.g.*, dyskinesias). Based on the deficits observed in Parkinsonian states (*e.g.*,

589   perseveration), this role may extend to behavioral transitions more generally, *e.g.*, starting new

590   movements or stopping ongoing movements (*Guru et al., 2020*).

591

Is DAN ramping also present before "spontaneous" movements?

We have suggested that the ramping DAN signals in the self-timed movement task could be related to reward expectation coupled with the explicit timing requirement of the task. However, when we averaged DAN signals aligned to "spontaneous" licks during the ITI, we also observed noisy, slow ramping signals building over seconds up to the time of the next lick, with a time course related to the duration of the inter-*lick* interval (*Figure 8—figure supplement 2*). This observation raises the possibility that slowly evolving DAN signals may be related to the generation of self-initiated movements more generally—although our highly trained animals may have also been "rehearsing" timed movements between trials, and/or expecting reward even for spontaneous licks.

Relationship to setpoint and stretching dynamics in other neural circuits

We found that DAN signals predict movement timing via two low-dimensional signals: a baseline offset and a ramping dynamic that "stretches" depending on trial-by-trial movement timing. Intriguingly, similar stretching of neural responses has been observed before self-timed movement in other brain areas in rats and primates, including the dorsal striatum (*Emmons et al., 2017; Mello et al., 2015; Wang et al., 2018*), lateral interparietal cortex (*Maimon and Assad, 2006*), presupplementary and supplementary motor areas (*Mita et al., 2009*), and dorsomedial frontal cortex (DMFC; *Remington et al., 2018; Sohn et al., 2019; Wang et al., 2018; Xu et al., 2014*). In the case of DMFC, applying dimensionality reduction to the population responses revealed two lower-dimensional characteristics that resembled our findings in DANs: 1) the speed at which the population dynamics unfolded was scaled ("stretched") to the length of the produced timing interval (*Wang et al., 2018*), and 2) the population state at the beginning of the

615 self-timed movement interval ("setpoint") was correlated with the timed interval (*Remington et*

616 *al., 2018; Sohn et al., 2019*). Recurrent neural network models suggested variation in stretching

617 and setpoint states could be controlled by (unknown) tonic or monotonically-ramping inputs to

618 the cortico-striatal system (*Remington et al., 2018; Sohn et al., 2019; Wang et al., 2018*). We

619 found that DANs exhibit both baseline (*e.g.*, "setpoint") signals related to timing, as well as

620 monotonically ramping input during the timing interval. Thus, through their role as diffusely-

621 projecting modulators, DANs could potentially orchestrate variations in cortico-striatal dynamics

622 observed during timing behavior. Ramping DAN signals could also be related to the slow

623 ramping signals that have been observed in the human motor system in anticipation of self-

624 initiated movements, *e.g.*, readiness potentials in EEG recordings (*Deecke, 1996; Libet et al.,*

625 *1983*).

626

627 **Possible relationship to motivational/movement vigor**

628 In operant tasks in which difficulty is systematically varied over blocks of trials, increased inter-

629 trial dopamine in the nucleus accumbens has been associated with higher average reward rate

630 and decreased latency to engage in a new trial, suggesting a link between dopamine and

631 "motivational vigor," the propensity to invest effort in work (*Hamid et al., 2016; Mohebi et al.,*

632 *2019*). Intriguingly, we observed the *opposite* relationship in the self-timed movement task:

633 periods with higher average reward rates had *lower* average baseline dopaminergic signals and

634 later first-lick times. Moreover, for a given first-lick time (*e.g.*, 3.5-3.75 s), we did not detect

635 differences in baseline (or ramping) signals during periods with different average reward rates,

636 such as near the beginning or end of a session. This difference between the two tasks may be due

637 to their opposing strategic constraints: in the aforementioned experiments, faster trial initiation

638   increased the number of opportunities to obtain reward, whereas earlier first-licks tended to

639   decrease reward acquisition in our self-timed movement task.

640

641   The basal ganglia have also been implicated in controlling "movement vigor," generally

642   referring to the speed, force or frequency of movements (*Bartholomew et al., 2016; Dudman*

643   *and Krakauer, 2016; Panigrahi et al., 2015; Turner and Desmurget, 2010; Yttri and Dudman,*

644   *2016*). The activity of nigrostriatal DANs has been shown to correlate with these parameters

645   during movement bouts and could promote more vigorous movement via push-pull interactions

646   with the direct and indirect pathways (*Barter et al., 2015; da Silva et al., 2018; Mazzoni et al.,*

647   *2007; Panigrahi et al., 2015*). Movement vigor might also entail earlier self-timed movements,

648   mediated by moment-to-moment increases in dopaminergic activity.

649

650   If moving earlier is a signature of greater movement vigor, then earlier self-timed movements

651   might also be executed with greater force/speed. We looked for movement-related vigor signals,

652   examining both the amplitude of lick-related EMG signals and the latency between lick initiation

653   and lick-tube contact. We detected no consistent differences in these force- or speed-related

654   parameters as a function of movement time; on the contrary, the EMG signals were highly

655   stereotyped irrespective of the first-lick time (data not shown). It is possible that vigor might

656   affect movement timing without affecting movement kinematics/dynamics—but, if so, the

657   distinction between "timing" and "vigor" would seem largely semantical.

658

Overall view

We have posited that dopaminergic ramping reflects reward expectation, a common element of behavioral paradigms that reveal slow dopaminergic ramping. Furthermore, our optogenetic manipulations indicate that dopaminergic signals do not directly trigger movements, but rather act as if modulating the probability of the pre-potent self-timed movement. Taken together, these observations suggest that as DAN activity ramps up, the probability of movement likewise increases. In this view, different rates of increase in DAN activity lead to shorter or longer elapsed intervals before movement, on average. This framework leaves open the question of what makes movement timing "probabilistic." One possibility is that recurrent cortical-basal ganglia–thalamic circuits could act to generate movements "on their own," without direct external triggers (*e.g.*, a "go!" cue). By providing crucial modulation of these circuits, DANs could tune the propensity to make self-timed movements—and pathological loss of DANs could reduce the production of such movements. Future experiments should address how dynamic dopaminergic input influences downstream motor circuits involved in self-timed movements.

SOURCE DATA

Source data files have been provided for all main figures and supplements:

Figure 1—source data
Figure 2—source data 1
Figure 2—source data 2
Figure 2—source data 3
Figure 2—source data 4
Figure 3—source data
Figure 4—source data
Figure 5—source data
Figure 6—source data
Figure 7—source data
Figure 8—source data
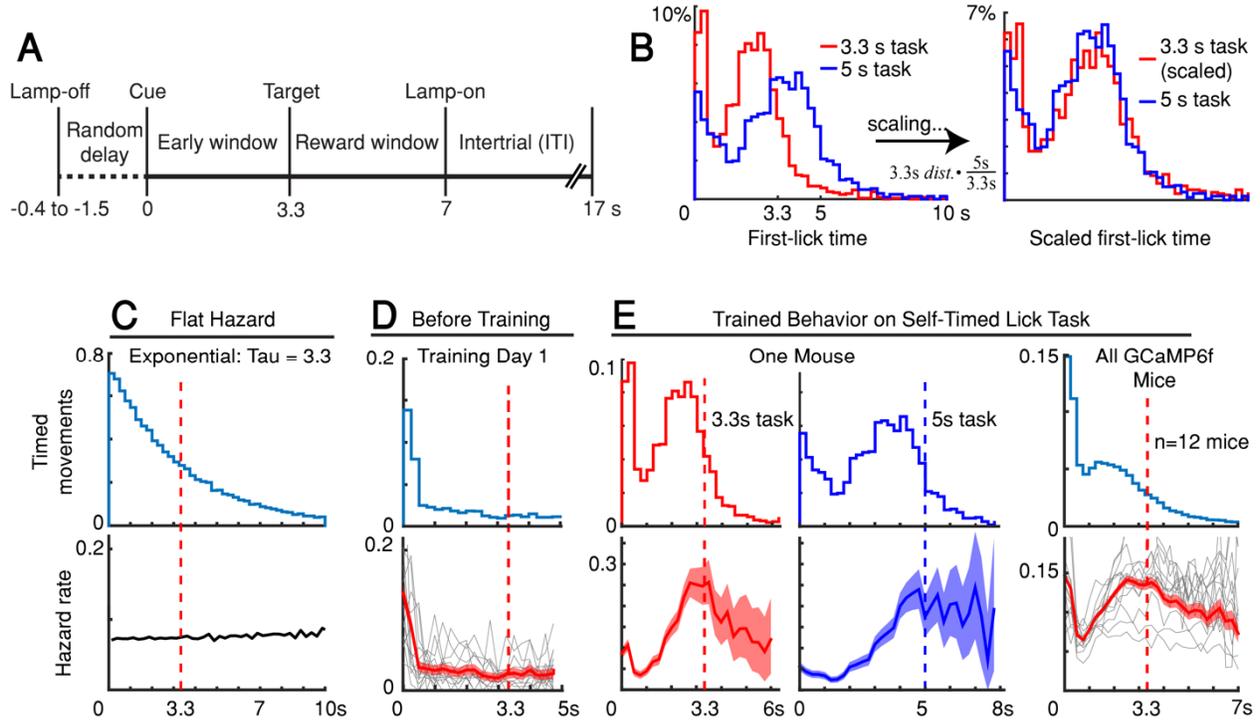
Figure 1. Self-timed movement task. (**A**) Task schematic (3.3 s version shown). (**B**) First-lick timing distributions generated by the same mouse exhibit the scalar property of timing (Weber's Law). Red: 3.3 s target time (4 sessions); Blue: 5 s target time (4 sessions). For all mice, see *Figure 1—figure supplement 1B*. (**C-E**) Hazard-function analysis. Time=0 is the start-timing cue; dashed vertical lines are target times. (**C**) Uniform instantaneous probability of movement over time is equivalent to a flat hazard rate (bottom) and produces an exponential first-lick timing distribution (top). (**D**) Before Training: First day of exposure to the self-timed movement task. Top: average first-lick timing distribution across mice; bottom: corresponding hazard functions. Gray traces: single session data. Red traces: average among all sessions, with shading indicating 95% confidence interval produced by 10,000x bootstrap procedure. (**E**) Trained Behavior: Hazard functions (bottom) computed from the first-lick timing distributions for the 3.3 s- and 5 s tasks (top) reveal peaks at the target times. Right: average first-lick timing distribution and hazard functions for all 12 GCaMP6f photometry animals. See also *Figure 1—figure supplements 1-2*. Source data: *Figure 1—source data*.
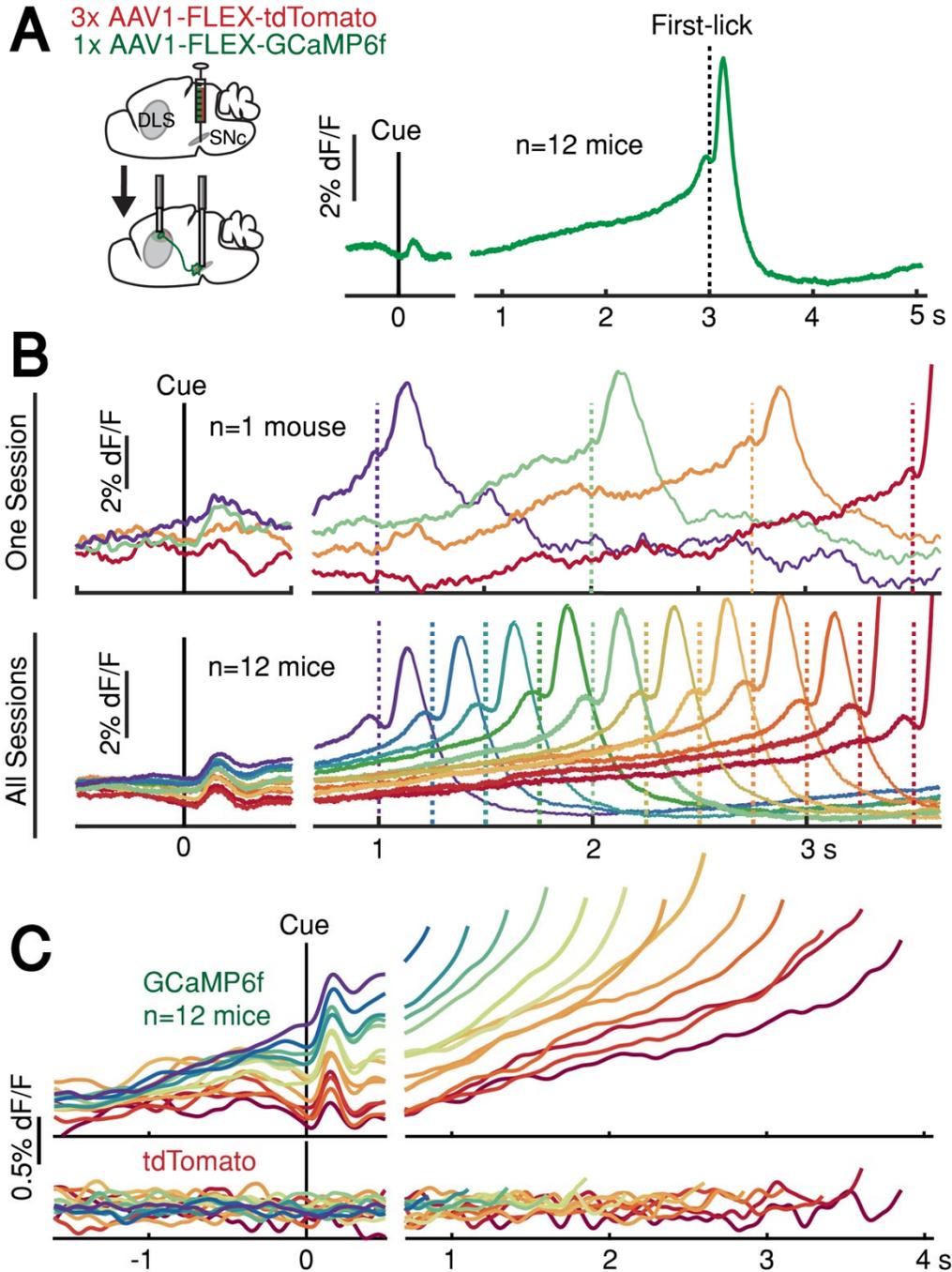
Figure 2. SNc DAN signals preceding self-timed movement. (**A**) Left: surgical strategy for GCaMP6f/tdTomato fiber photometry. Right: average SNc DAN GCaMP6f response for first-licks between 3-3.25 s (12 mice). Data aligned separately to both cue-onset (left) and first-lick (right), with the break in the time axis indicating the change in plot alignment. (**B**) Average SNc DAN GCaMP6f responses for different first-lick times (indicated by dashed vertical lines). (**C**) Comparison of average DAN GCaMP6f and tdTomato responses on expanded vertical scale. Traces plotted up to 150 ms before first-lick. See also *Figure 2—figure supplements 1-3*. Source data: *Figure 2—source data 1*.
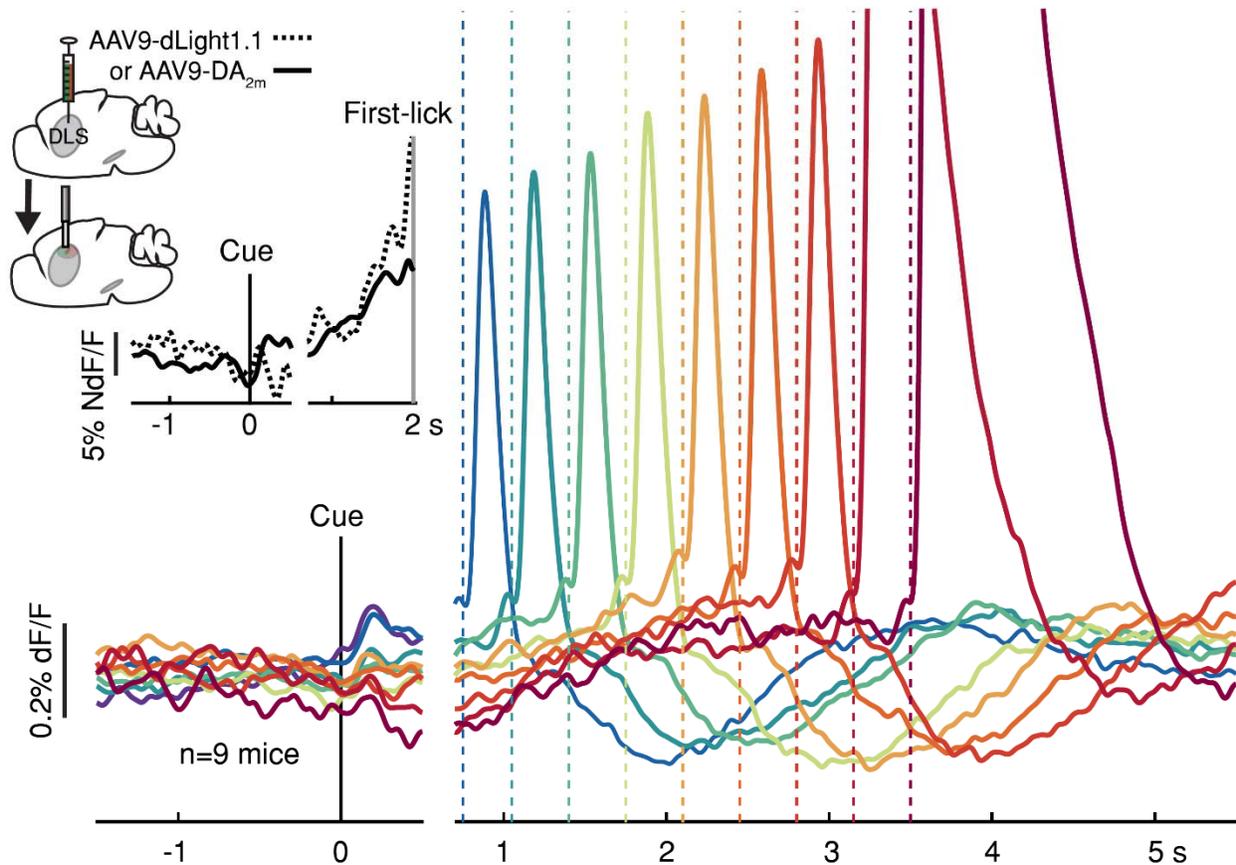
Figure 3. Striatal dopamine release during the self-timed movement task. Photometry signals averaged together from DA_{2m} signals (n=4 mice) and dLight1.1 signals (n=5 mice) recorded in DLS. Axis break and plot alignment as in *Figure 2*. Dashed lines: first-lick times. Inset, left: surgical strategy. Inset, right: Comparison of dLight1.1 and DA_{2m} dynamics. Expanded vertical scale to show ramping in the average signals for DA_{2m} (solid trace) and dLight1.1 (dashed trace) up until the time of the first-lick (first-lick occurred between 2-3 s after the cue for this subset of the data). See also: *Figure 3—figure supplement 1*. Source data: *Figure 3—source data*.
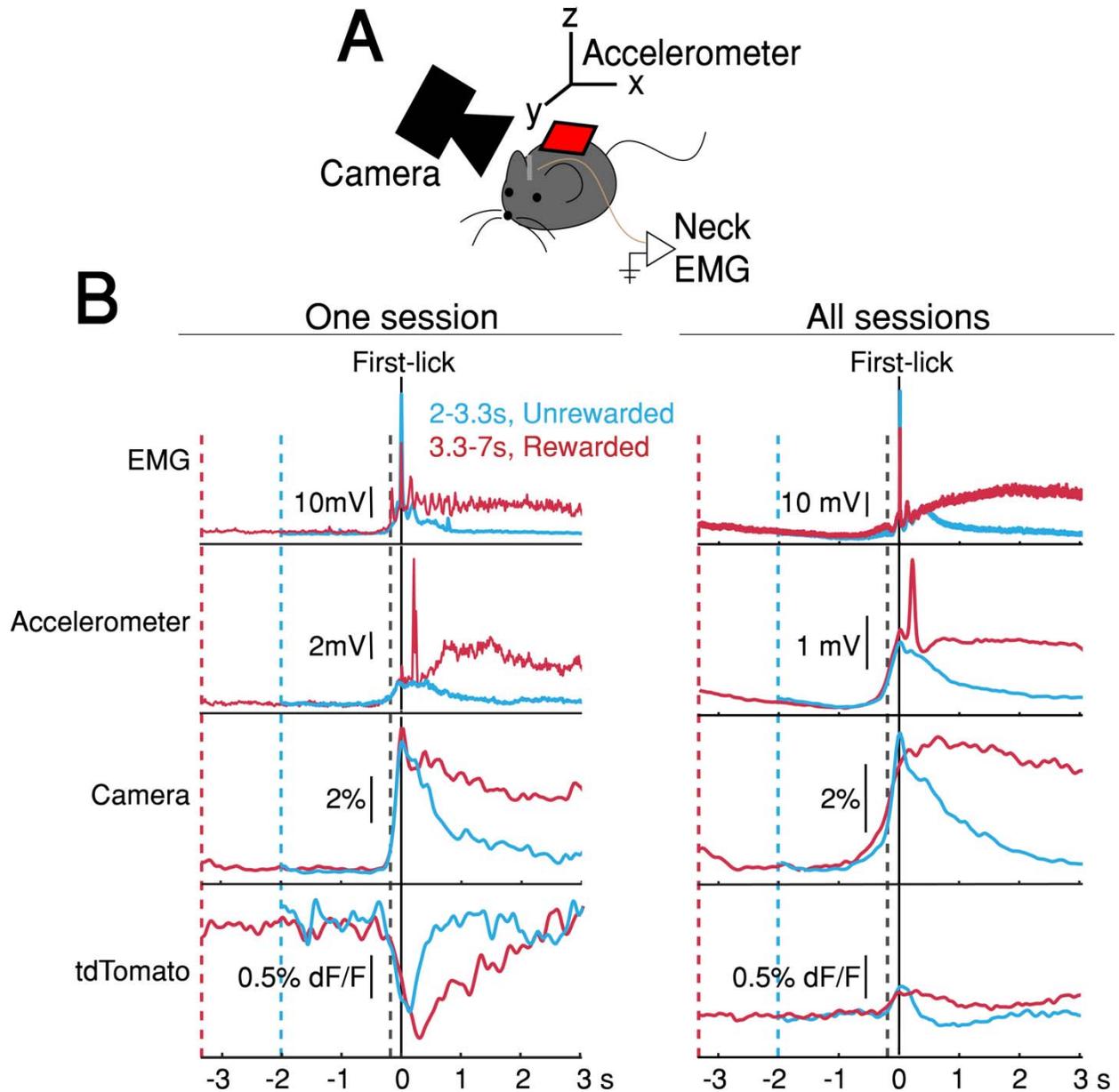
Figure 4. Movement controls reliably detected movements, but there were no systematic differences in movement during the timing interval. **(A)** Schematic of movement-control measurements. **(B)** First-lick-aligned average movement signals on rewarded (red) and unrewarded (blue) trials. Pre-lick traces begin at the nearest cue-time (dashed red, dashed blue). Left: one session; Right: all sessions. Dashed grey line: time of earliest-detected movement on most sessions (150 ms before first-lick). Average first-lick-aligned tdTomato optical artifacts showed inconsistent excursion directions (up/down) even within the same session; signals for each artifact direction shown in *Figure 4—figure supplement 1*. Source data: *Figure 4—source data*.
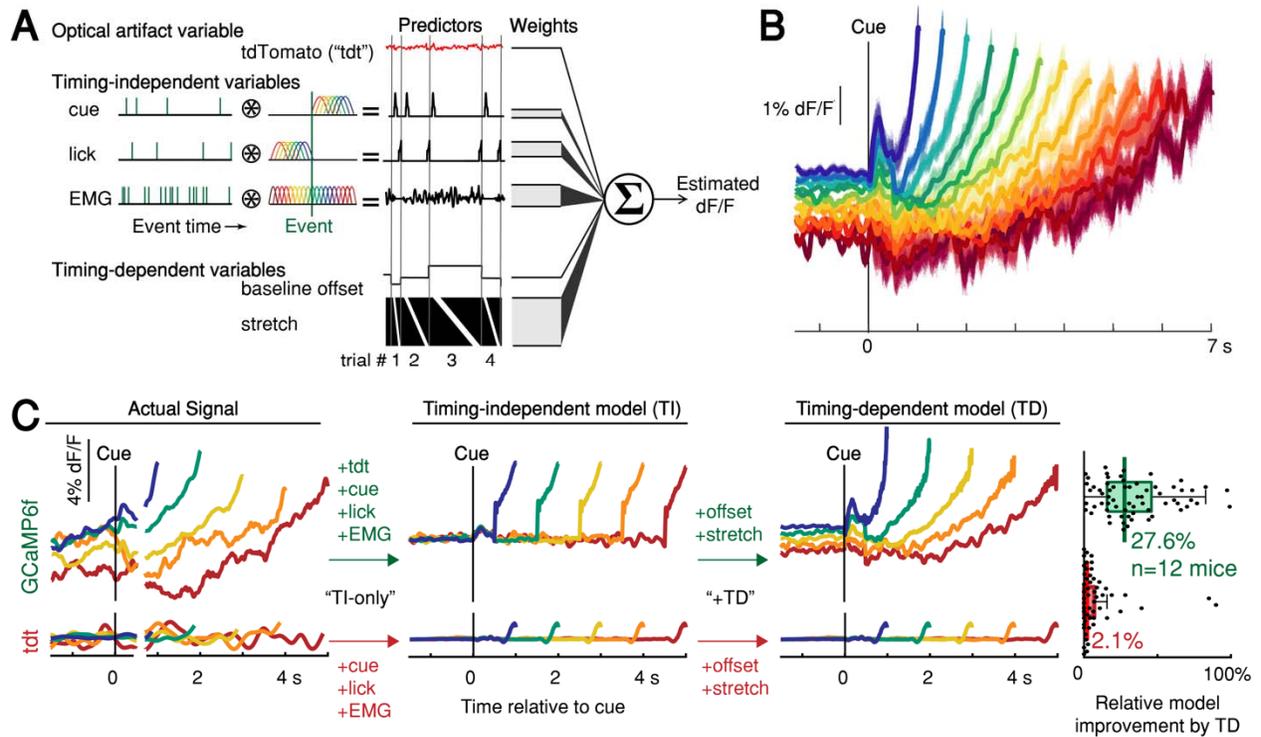
Figure 5. Contribution of optical artifacts, task variables and nuisance bodily movements to SNc GCaMP6f signals. **(A)** Nested encoding model comparing the contribution of timing-independent predictors (TI) to the contribution of timing-dependent predictors (TD). **(B)** Predicted dF/F signal for one session plotted up to time of first-lick. Model error simulated 300x (shading). **(C)** Nested encoding model for one session showing the actual recorded signal (1st panel), the timing-independent model (2nd panel), and the full, timing-dependent model with all predictors (3rd panel). Top: GCaMP6f; Bottom: tdTomato (tdt). Right: relative loss improvement by timing-dependent predictors (grey dots: single sessions, line: median, box: lower/upper quartiles, whiskers: 1.5x IQR). See also *Figure 5—figure supplement 1*. Source data: *Figure 5—source data*.
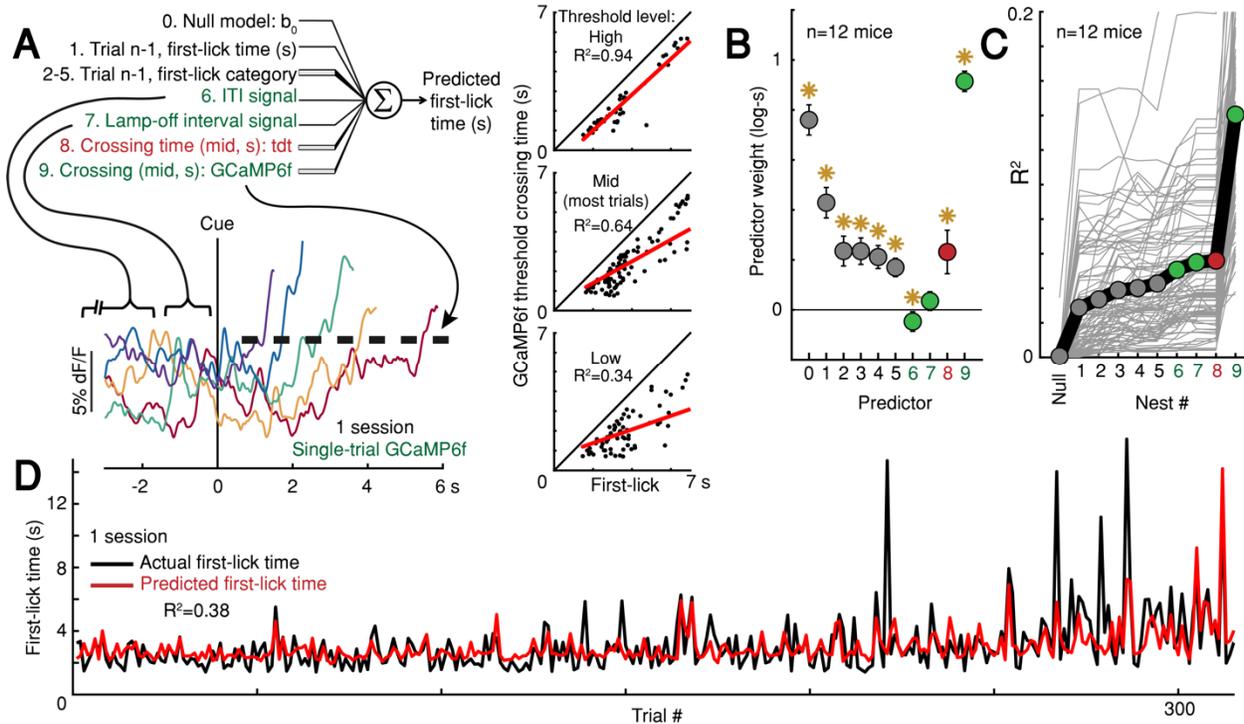
Figure 6. Single-trial DAN signals predict first-lick timing. **(A)** Schematic of nested decoding model. Categories for n-1th trial predictors: 2) reaction, 3) early, 4) reward, 5) ITI first-lick (see *Methods*). Bottom: single-trial cue-aligned SNc DAN GCaMP6f signals from one session (6 trials shown for clarity). Traces plotted up to first-lick. Right: threshold-crossing model. Low/Mid/High label indicates threshold amplitude. Dots: single trials. **(B)** Model weights. Error bars: 95% CI, *: $p < 0.05$, 2-sided t-test. Numbers indicate nesting-order. **(C)** Variance explained by each model nest. Grey lines: single sessions; thick black line: average. For model selection, see *Figure 6—figure supplement 1C*. **(D)** Predicted *vs*. actual first-lick time, same session as *6A*. See also *Figure 6—figure supplements 1-4*. Source data: *Figure 6—source data*.

36
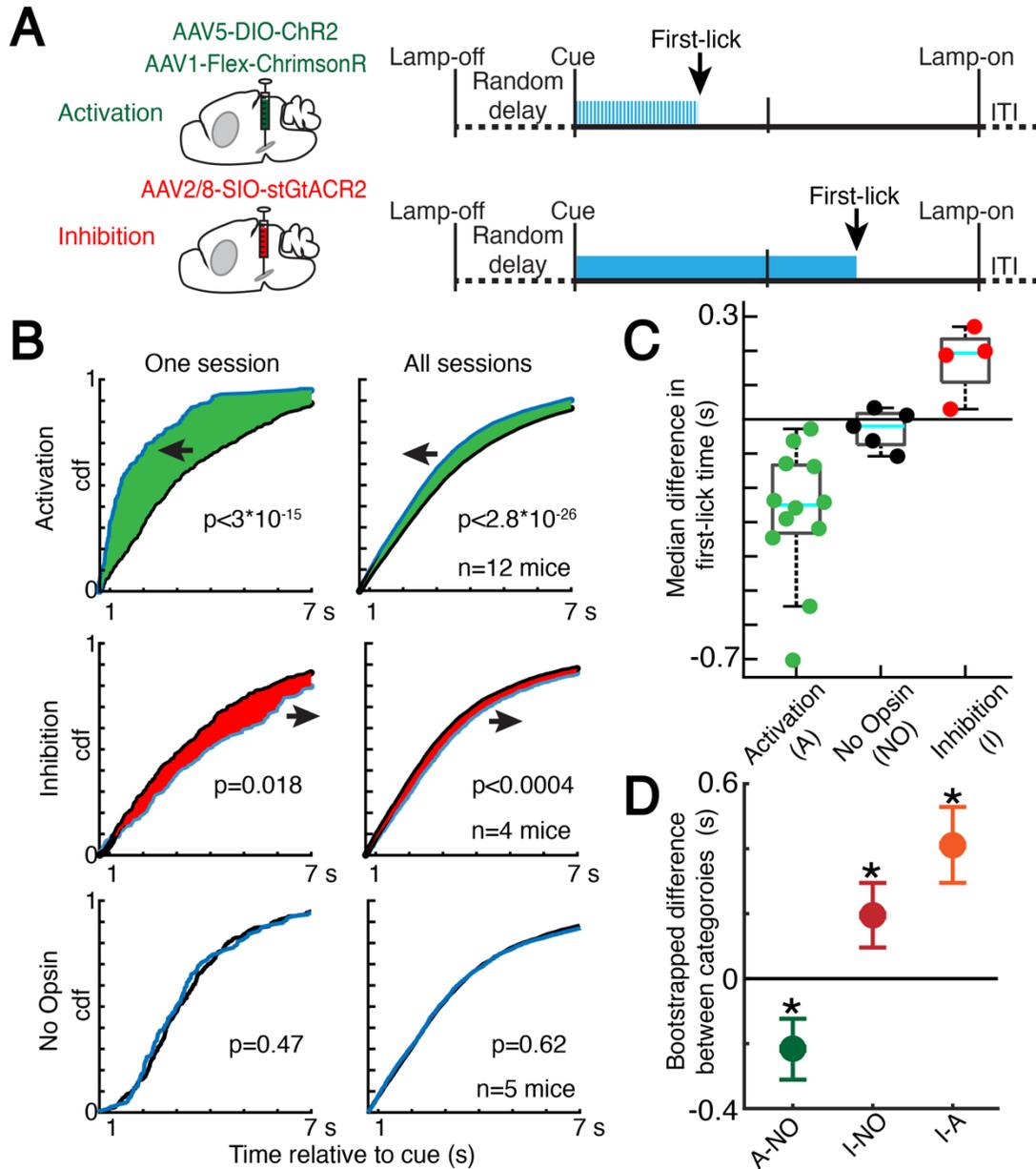
Figure 7. Optogenetic DAN manipulation systematically and bidirectionally shifts the timing of self-timed movements. **(A)** Strategy for optogenetic DAN activation or inhibition. Mice were stimulated from cue-onset until first-lick or 7 s. **(B)** Empirical continuous probability distribution functions (cdf) of first-lick times for stimulated (blue line) versus unstimulated (grey line) trials**.** Arrow and shading show direction of effect. P-values calculated by Kolmogorov-Smirnov test (for other metrics, see *Figure 7—figure supplements 1-2*). **(C)** Median 1,000,000x bootstrapped difference in first-lick time, stimulated-minus-unstimulated trials. Box: upper/lower quartile; line: median; whiskers: 1.5x IQR; dots: single mouse. **(D)** Comparison of median first-lick time difference across all sessions. Error bars: 95% confidence interval (*: p<0.05, 1,000,000 bootstrapped median difference in first-lick time between sessions of different stimulation categories). See also *Figure 7—figure supplements 1-4*. Source data: *Figure 7—source data*.
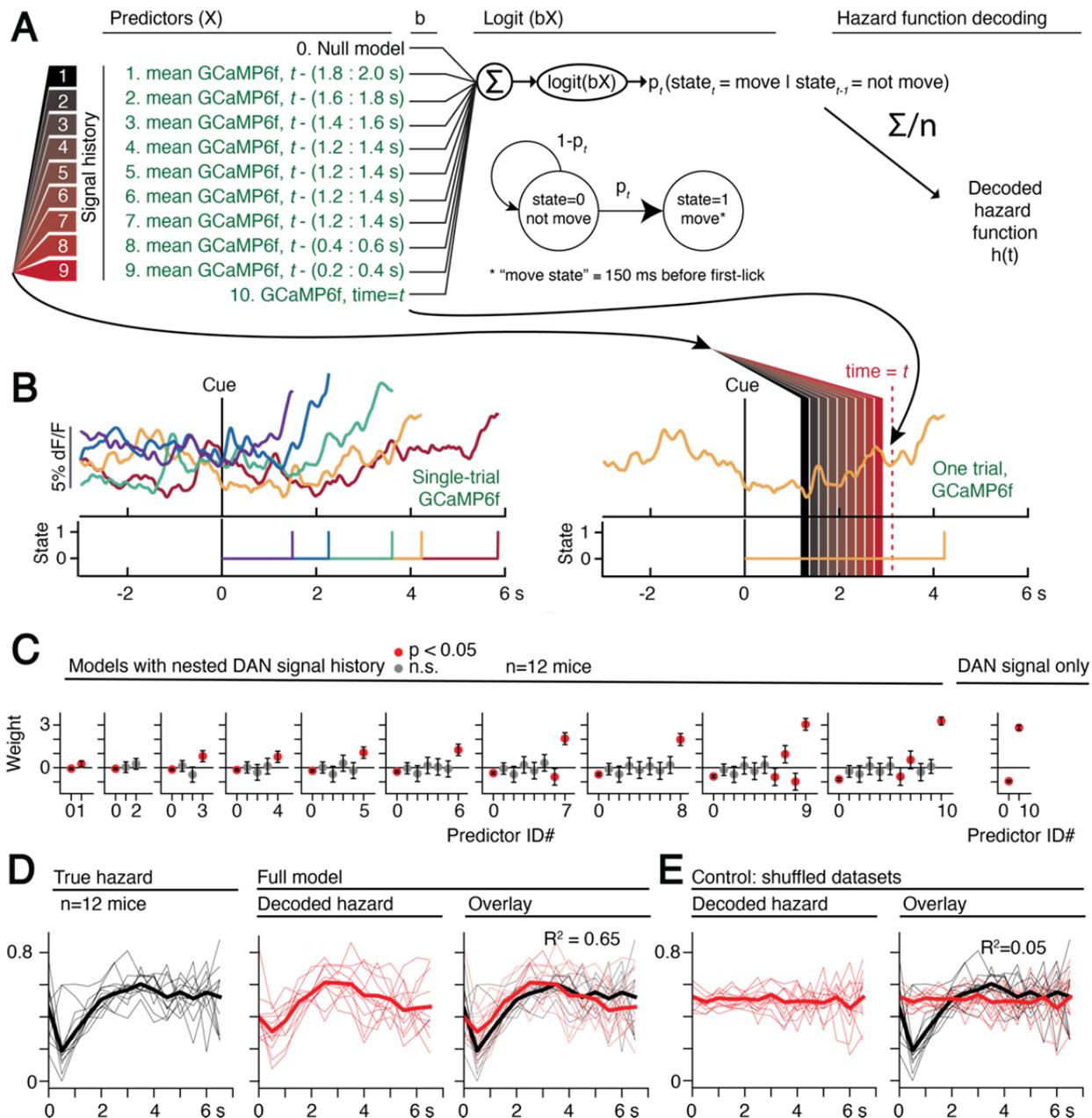
Figure 8. Single-trial dynamic dopaminergic signals predict the moment-to-moment probability of movement initiation. **(A)** Probabilistic movement-state model schematic. **(B)** Single-trial DAN GCaMP6f signals at SNc from one session. First-lick time truncated 150 ms before movement detection to exclude peri-movement signals. Bottom: Movement states for the trials shown as a function of time. Diagram on the right schematizes the model predictors relative to an example time=t on a single trial. **(C)** Nested model fitted coefficients. **(D)** Decoded hazard functions from full model (with all 10 predictors). Thick line=mean. n=12 mice. **(E)** Hazard function fitting with shuffled datasets abolished the predictive power of the model (same 12 mice). See also *Figure 8—figure supplements 1-2*. Source data: *Figure 8—source data*.

Key Resources Table				
Reagent type (species) or resource	**Designation**	**Source or reference**	**Identifiers**	**Additional information**
strain, strain background (*M. musculus*)	DAT-Cre	The Jackson Laboratory, Bar Harbor, ME	B6.SJL-*Slc6a3$^{tm1.1(cre)Bkmm}$*/J RRID:IMSR_JAX:020080	Cre expression in dopaminergic neurons
strain, strain background (*M. musculus*)	Wild-type	The Jackson Laboratory, Bar Harbor, ME	C57BL/6 RRID:IMSR_JAX:000664	
other	tdTomato ("tdt")	UNC Vector Core, Chapel Hill, NC	AAV1-CAG-FLEX-tdT	Virus, for control photometry expression
other	gCaMP6f	Penn Vector Core, Philadelphia, PA	AAV1.Syn.Flex.GCaMP6f.WPRE.SV40	Virus, for photometry expression
other	DA2m	Vigene, Rockville, MD	AAV9-hSyn-DA4.4(DA2m)	Virus, for photometry expression
other	dLight1.1	Lin Tian Lab; Children's Hospital Boston Viral Core, Boston, MA	AAV9.hSyn.dLight1.1.wPRE	Virus, for photometry expression
other	turboRFP	Penn Vector Core	AAV1.CB7.CI.TurboRFP.WPRE.rBG	Virus, for control photometry expression

other	ChR2	UNC Vector Core, Chapel Hill, NC	AAV5-EF1a-DIO-hChR2(H134R)-EYFP-WPRE-pA	Virus, for opsin expression
other	ChrimsonR	UNC Vector Core, Chapel Hill, NC	AAV1-hSyn-FLEX-ChrimsonR-tdT	Virus, for opsin expression
other	stGtACR2	Addgene/Janelia Viral Core, Ashburn, VA	AAV2/8-hSyn1-SIO-stGtACR2-FusionRed	Virus, for opsin expression
software, algorithm	Matlab	Mathworks	Matlab2018B	For most analyses
software, algorithm	Julia Programming Language	The Julia Project	Julia 1.5.3	For probabilistic models
software, algorithm	Gen.jl	The Gen Team	Gen.jl	For probabilistic models

784

Animals

786 Adult male and female hemizygous DAT-cre mice (*Backman et al.,*

787 *2006*; B6.SJL-*Slc6a3*$^{tm1.1(cre)Bkmm}$/J, RRID:IMSR_JAX:020080; The Jackson Laboratory, Bar

788 Harbor, ME) or wild-type C57BL/6 mice were used in all experiments (> 2 months old at the

789 time of surgery; median body weight 23.8g, range 17.3-31.9 g). Mice were housed in standard

790 cages in a temperature and humidity-controlled colony facility on a reversed night/day cycle

791 (12 h dark/12 h light), and behavioral sessions occurred during the dark cycle. Animals were

792 housed with enrichment objects provided by the Harvard Center for Comparative Medicine

793   (IACUC-approved plastic toys/shelters, e.g., Bio-Huts, Mouse Tunnels, Nest Sheets, *etc*.) and

794   were housed socially whenever possible (1-5 mice per cage). All experiments and protocols were

795   approved by the Harvard Institutional Animal Care and Use Committee (IACUC protocol

796   #05098, Animal Welfare Assurance Number #A3431-01) and were conducted in accordance

797   with the National Institutes of Health Guide for the Care and Use of Laboratory Animals.

798

799   **Surgery**

800   Surgeries were conducted under aseptic conditions and every effort was taken to minimize

801   suffering. Mice were anesthetized with isoflurane (0.5-2% at 0.8 L/min). Analgesia was provided

802   by *s.c.* 5 mg/kg ketoprofen injection during surgery and once daily for 3 d postoperatively

803   (Ketofen, Parsippany, NJ). Virus was injected (50 nL/min) and the pipet remained in place for 10

804   min before removal. 200 µm, 0.53 NA blunt fiber optic cannulae (Doric Lenses, Quebec, Canada)

805   or tapered fiber optic cannulae (200 µm, 0.60 NA, 2 mm tapered shank, OptogeniX, Lecce, Italy)

806   were positioned at SNc, VTA or DLS and secured to the skull with dental cement (C&B

807   Metabond, Parkell, Edgewood, NY). Neck EMG electrodes were constructed from two Teflon-

808   insulated 32G stainless steel pacemaker wires attached to a custom socket mounted in the dental

809   cement. Sub-occipital neck muscles were exposed by blunt dissection and electrode tips

810   embedded bilaterally.

811

812   **Stereotaxic coordinates (from bregma and brain surface)**

813   Viral Injection:

814   <u>SNc</u>: 3.16 mm posterior, +/- 1.4 mm lateral, 4.2 mm ventral

815   <u>VTA</u>: 3.1 mm posterior, +/-0.6 mm lateral, 4.2 mm ventral

816   DLS: 0 mm anterior, +/- 2.6 mm lateral, 2.5 mm ventral.

817   Fiber Optic Tips:

818   SNc/VTA: 4.0 mm ventral (photometry) or 3.9 mm ventral (optogenetics).

819   DLS: 2.311 mm ventral (blunt fiber) or 4.0 mm ventral (tapered fiber)

820

821   **Virus**

822   Photometry:

823   tdTomato ("tdt"): AAV1-CAG-FLEX-tdT (UNC Vector Core, Chapel Hill, NC), 100 nL

824   used alone or in mixture with other fluorophores (below), working concentration

825              $5.3*10^{12}$ gc/mL

826   gCaMP6f (at SNc or VTA): 100 nL AAV1.Syn.Flex.GCaMP6f.WPRE.SV40

827   ($2.5*10^{13}$ gc/mL, Penn Vector Core, Philadelphia, PA). Virus was mixed in a 1:3 ratio

828   with tdt (200 nL total)

829   DA_{2m} (at DLS): 200-300 nL AAV9-hSyn-DA4.4(DA2m) (working concentration:

830   *ca.* $3*10^{12}$ gc/mL, Vigene, Rockville, MD) + 100 nL tdt

831   dLight1.1 (at DLS): 300 nL AAV9.hSyn.dLight1.1.wPRE bilaterally at DLS (*ca.*

832              $9.6*10^{12}$ gc/mL, Children's Hospital Boston Viral Core, Boston, MA) + 100 nL

833   AAV1.CB7.CI.TurboRFP.WPRE.rBG (*ca.* $1.01*10^{12}$ gc/mL, Penn Vector Core)

834   Optogenetic stimulation/inhibition (all bilateral at SNc):

835   ChR2: 1000 nL AAV5-EF1a-DIO-hChR2(H134R)-EYFP-WPRE-pA ($3.2*10^{13}$ gc/mL,

836   UNC Vector Core, Chapel Hill, NC)

837   ChrimsonR +/- dLight1.1: 700 nL AAV1-hSyn-FLEX-ChrimsonR-tdT ($4.1*10^{12}$ gc/mL,

838   UNC Vector Core, Chapel Hill, NC) +/- 400-550 nL AAV9-hSyn-dLight1.1

839   bilaterally at DLS (*ca.* 10^{13} gc/mL, Lin Tian Lab, Los Angeles, CA)

840   stGtACR2: 300 nL 1:10 AAV2/8-hSyn1-SIO-stGtACR2-FusionRed (working

841   concentration $4.7*10^{11}$ gc/mL, Addgene/Janelia Viral Core, Ashburn, VA)

842

843   **Water-deprivation and acclimation**

844   Animals recovered for 1 week postoperatively before water deprivation. Mice received daily

845   water supplementation to maintain ≥80% initial body weight and fed *ad libitum*. Mice were

846   habituated to the experimenter and their health was monitored carefully following guidelines

847   reported previously (*Guo et al., 2014*). Training commenced when mice reached the target

848   weight (~8-9 d post-surgery).

849

850   **Histology**

851   Mice were anesthetized with >400 mg/kg pentobarbital (Somnasol, Henry Schein Inc, Melville,

852   NY) and perfused with 10 mL 0.9% sodium chloride followed by 50mL ice-cold 4%

853   paraformaldehyde in 0.1 M phosphate buffer. Brains were fixed in 4% paraformaldehyde at 4°C

854   for >24 hr before being transferred to 30% sucrose in 0.1 M phosphate buffer for >48 hr. Brains

855   were sliced in 50 μm coronal sections by freezing microtome, and fluorophore expression was

856   assessed by light microscopy. The sites of viral injections and fiber optic placement were

857   mapped with an Allen Mouse Brain Atlas.

858

859 **Behavioral rig, data acquisition and analysis**

860 A custom rig provided sensory cues, recorded events and delivered juice rewards under the

861 control of a Teensy 3.2 microprocessor running a custom Arduino state-system behavioral

862 program with MATLAB serial interface. Digital and analog signals were acquired with a CED

863 Power 1400 data acquisition system/Spike2 software (Cambridge Electronic Design Ltd,

864 Cambridge, England). Photometry and behavioral events were acquired at 1,000 Hz; movement

865 channels were acquired at 2,000 Hz. Video was acquired with FlyCap2 or Spinnaker at 30 fps

866 (FLIR Systems, Wilsonville, OR). Data were analyzed with custom MATLAB statistics

867 packages.

868

869 **Self-timed movement task**

870 Mice were head-fixed with a juice tube positioned in front of the tongue. The spout was placed

871 as far away from the mouth as possible so that the tongue could still reach it to discourage

872 compulsive licking (*Guo et al., 2014*), ~1.5 mm ventral and ~1.5 mm anterior to the mouth.

873 During periods when rewards were not available, a houselamp was illuminated. At trial start, the

874 houselamp turned off, and a random delay ensued (0.4-1.5 s) before a cue (simultaneous LED

875 flash and 3300 Hz tone, 100 ms) indicated start of the timing interval. The timing interval was

876 divided into two windows, early (0-3.333 s in most experiments; 0-4.95 s in others) and reward

877 (3.333-7 s; 4.95-10 s), followed by the intertrial interval (ITI, 7-17 s; 10-20 s). The window in

878 which the mouse first licked determined the trial outcome (early, reward, or no-lick). An early

879 first-lick caused an error tone (440 Hz, 200 ms) and houselamp illumination, and the mouse had

880 to wait until the full timing interval had elapsed before beginning the ITI. Thus there was no

881 advantage to the mouse of licking early. A first-lick during the reward window caused a reward

44

882   tone (5050 Hz, 200 ms) and juice delivery, and the houselamp remained off until the end of the

883   trial interval. If the timing interval elapsed with no lick, a time-out error tone played (131 Hz, 2

884   s), the houselamp turned on, and ITI commenced. During the ITI and pre-cue delay ("lamp-off

885   interval"), there was no penalty for licking.

886

887   Mice learned the task in 3 stages (*Figure 1—figure supplement 1A*). On the first 1-4 days of

888   training, mice learned a beginner-level task, which was modified in two ways: 1) to encourage

889   participation, if mice did not lick before 5 s post-cue, they received a juice reward at 5 s; and 2)

890   mice were not penalized for licking in reaction to the cue (within 500 ms). When the mouse

891   began self-triggering ≥50% of rewards (day 2-6 of training), the mouse advanced to the

892   intermediate-level task, in which the training reward at 5 s was omitted, and the mouse had to

893   self-trigger all rewards. After completing >250 trials/day on the intermediate task (usually day 4-

894   7 of training), mice advanced to the mature task, with no reaction licks permitted. All animals

895   learned the mature task and worked for ~400-1,500 trials/session.

896

897   **Hazard function correction of survival bias in the timing distribution**

898   The raw frequency of a particular response time in the self-timed movement task is "distorted"

899   by how often the animal has the chance to respond at that time (*Anger, 1956*). This bias was

900   corrected by calculating the hazard function, which takes into account the number of response

901   opportunities the animal had at each timepoint. The hazard function is defined as the conditional

902   probability of moving at a time, *t*, given that the movement has not yet occurred (referred to as

903   "IRT/Op" analysis in the old Differential Reinforcement of Low Rates (DRL) literature). The

904   hazard function was computed by dividing the number of first-movements in each 250 ms bin of

905   the first-lick timing histogram by the total number of first-movements occurring at that bin-time

906   or later—the total remaining "opportunities."

907

908   **Online movement monitoring**

909   Movements were recorded simultaneously during behavior with four movement-control

910   measurements: neck EMG (band-pass filtered 50-2,000 Hz, 60 Hz notch, amplified 100-1,000x),

911   back-mounted accelerometer (SparkFun Electronics, Boulder, CO), high-speed camera

912   (30 Hz, FLIR Systems, Wilsonville, OR), and tdTomato photometry. All control signals

913   contained similar information, and thus only a subset of controls was used in some sessions.

914

915   **Photometry**

916   Fiber optics were illuminated with 475 nm blue LED light (Plexon, Dallas, TX)

917   (SNc/VTA: 50 µW, DLS: 35 µW) measured at patch cable tip with a light-power meter

918   (Thorlabs, Newton, NJ). Green fluorescence was collected via a custom dichroic mirror (Doric

919   Lenses, Quebec, Canada) and detected with a Newport 1401 Photodiode (Newport Corporation,

920   Irvine, CA). Fluorescence was allowed to recover ≥1 d between recording sessions. To avoid

921   crosstalk in animals with red control fluorophore expression, the red channel was recorded at one

922   of the 3 sites (SNc, VTA, or DLS, 550 nm lime LED, Plexon, Dallas, TX) while GCaMP6f,

923   dLight1.1 or DA_{2m} was recorded simultaneously only at the other implanted sites.

924

925 **dF/F**

926 Raw fluorescence for each session was pre-processed by removing rare singularities (single

927 points >15 STD from the mean) by interpolation to obtain F(t). To correct photometry signals for

928 bleaching, dF/F was calculated as:

929

$$\frac{dF}{F}(t) = \frac{F(t) - F_0(t)}{F_0(t)}$$

930

931

932 where $F_0(t)$ is the 200 s moving average of F(t) (*Figure 2—figure supplement 2A*). We tested

933 several other complementary methods for calculating dF/F and all reported results were robust to

934 dF/F method (see *Methods: dF/F method characterization and validation*). To ensure dF/F

935 signal processing did not introduce artifactual scaling or baseline shifts, we also tested several

936 complementary techniques to isolate undistorted F(t) signals where possible and quantified the

937 amount of signal distortion when perfect isolation was not possible (see *Methods: dF/F method*

938 *characterization and validation*, below, and *Figure 2—figure supplement 2C*).

939

940 **dF/F method characterization and validation**

941 dF/F calculations are intended to reduce the contribution of slow fluorescence bleaching to fiber

942 photometry signals, and many such methods have been described (*Kim et al., 2019; Mohebi et*

943 *al., 2019; Soares et al., 2016*). However, dF/F methods have the potential to introduce artifactual

944 distortion when the wrong method is applied in the wrong setting. Thus, to derive an appropriate

945 dF/F method for use in the context of the self-timed movement task, we characterized and

946 quantified artifacts produced by 4 candidate dF/F techniques.

947

948   *Detailed description of complementary dF/F methods.*

949   1. <u>Normalized baseline</u>: a commonly used dF/F technique in which each trial's

950   fluorescence is normalized to the mean fluorescence during the 5 s preceding the trial.

951   2. <u>Low-pass digital filter</u>: F_0 is the low-pass, digital infinite impulse response

952   (IIR)-filtered raw fluorescence for the whole session (implemented in MATLAB with

953   the built-in function *lowpass* with $f_c=5 \cdot 10^{-5}$ Hz, steepness=0.95).

954   3. <u>Multiple baseline</u>: a variation of Method 1, in which each trial's fluorescence is

955   normalized by the mean fluorescence during the 5 s preceding the current trial, as

956   well as 5 trials before the current trial and 5 trials after the current trial.

957   4. <u>Moving average</u>: F_0 is the 200 s moving average of the raw fluorescence at each point

958   (100 s on either side of the measured timepoint).

959

960   Although *normalized baseline* (Method 1) is commonly used to correct raw fluorescence signals

961   (F) for bleaching, this technique assumes that baseline activity has no bearing on the trial

962   outcome; however, because the mouse decides when to move in the self-timed movement task, it

963   is possible that baseline activity may differ systematically with the mouse's choice on a given

964   trial. Thus, normalizing F to the baseline period would obscure potentially physiologically-

965   relevant signals. More insidiously, if baseline activity *does* vary systematically with the mouse's

966   timing, normalization can also introduce substantial amplitude scaling and y-axis shifting

967   artifacts when correcting F with this method (*Figure 2—figure supplement 2C*, middle panels).

968   Thus, Methods 2-4 were designed and optimized to isolate photometry signals minimally

969   distorted by bleaching signals and systematic baseline differences during the self-timed

48

970   movement task. Methods 2-4 produced the same results in all statistical analyses, and the moving

971   average method is shown in all figures.

972

973   *Isolating minimally-distorted photometry signals with paired trial analyses of raw fluorescence.*

974   Although slow bleaching prevents comparison of raw photometry signals (F) at one time in a

975   behavioral session with those at another time, the time-course of appreciable bleaching was slow

976   enough in the reported behavioral sessions that minimal bleaching occurred over the course of

977   3 trials (~1 min, *Figure 2—figure supplement 2A*). Thus, to observe the most minimally-

978   distorted photometry signals possible, we compared F between *pairs of consecutive* trials

979   (*Figure 2—figure supplement 2B-C*). We compared F baseline signals between all paired trials

980   in which an early trial (unrewarded first-lick between 0.7-2.9 s; abbreviated as "E") was

981   followed by a rewarded trial (first-lick between 3.4-7 s; abbreviated as "R"); this two-trial

982   sequence is thus referred to as an "ER" comparison. To ensure systematic differences did not

983   result from subtle bleaching in the paired-trial interval, we reversed the ordering contingency and

984   also compared all Rewarded trials preceding Early trials ("RE" comparison). The same

985   systematic relationship between baseline signals and first-lick time was found for paired trials

986   analyzed by raw F (*Figure 2—figure supplement 2C, left panels*).

987

988   *Quantification of artifactual amplitude scaling/baseline shifts introduced by dF/F processing.*

989   Each Candidate dF/F Method was applied to the same Paired Trial datasets described above. The

990   resulting paired-fluorescence datasets were normalized after processing (minimum dF/F=0,

991   maximum=1). The amount of distortion introduced by dF/F was quantified with a Distortion

992   Index (DI), which was calculated as:

993 $$\text{Distortion Index, } DI(t) = abs(F(t) - dF/F(t))$$

994 where $F(t)$ and $dF/F(t)$ are the normalized, paired-trial raw fluorescence signal or dF/F signal at

995 time t, respectively. t spanned from the beginning of the n-1th trial (-20 s) to the end of the nth

996 trial (20 s), aligned to the cue of the nth trial (*Figure 2—figure supplement 2C, bottom panels*).

997 The DI shown in plots has been smoothed with a 200 ms moving average kernel for clarity.

998

999 As expected, normalizing fluorescence to the baseline period (*normalized baseline*) erased the

1000 correlation of baseline dF/F signals with first-lick time (*Figure 2—figure supplement 2C,*

1001 middle panels). More insidiously, this also resulted in distortion of GCaMP6f dynamics *during*

1002 the timing interval, evident in the diminished difference between E-signals compared to R-

1003 signals relative to the shapes observed in the raw fluorescence paired-trial comparison (*Figure*

1004 *2—figure supplement 2C, middle-bottom panel*). However, dF/F Methods 2-4 visually and

1005 quantitatively recapitulated the dynamics observed in the raw fluorescence comparison (*Figure*

1006 *2—figure supplement 2C, right panels*).

1007

1008 These results were corroborated by time-in-session permutation tests in which datasets for single

1009 sessions were divided into thirds (beginning of session, middle of session, and end of session).

1010 The differences between baseline and ramping dynamics observed in whole-session averages

1011 were present even within these shorter blocks of time within the session (*i.e*., faster ramping and

1012 elevated baseline signals on trials with earlier self-timed licks). Furthermore, permutation tests in

1013 which the block identity (begin, middle, end) was shuffled showed that this pattern held when

1014 trials with earlier first-licks from the end of the session were compared with trials with later first-

1015 licks from the beginning of the session (and *vice versa*).

1016

Normalized dF/F for comparing dopamine sensor signals

1017

1018 DA_{2m} was about twice as bright as dLight1.1, and thus generally yielded larger and less noisy

1019 dF/F signals. To compare the two extracellular dopamine sensors in the same plot, dF/F was

1020 normalized for each signal by the amplitude of its lick-related transient. dF/F was calculated as

1021 usual, and then the mean baseline-to-transient peak amplitude was measured for trials with first-

1022 licks occurring between 2-3 s. Percentage NdF/F is reported as the percentage of this amplitude.

1023

Dopamine sensor kinetics

1024

1025 dLight1.1 is an extracellular dopamine sensor derived from the dopamine-1-receptor, and has

1026 fast reported kinetics: rise $t_{1/2}$ = 9.5 ± 1.1 ms, decay $t_{1/2}$ = 90 ± 11 ms (*Patriarchi et al., 2018*).

1027 DA_{2m} is a new extracellular dopamine indicator derived from the dopamine-2-receptor, which

1028 provides brighter signals. DA_{2m} signals have been reported to decay slowly in slice preparations

1029 but are much faster *in vivo*, presumably because endogenous dopamine-clearance mechanisms

1030 are preserved: reported rise $t_{1/2}$ ~50 ms, decay $t_{1/2}$ ~360 ms in freely behaving mice; decay $t_{1/2}$

1031 ~190 ms in head-fixed *drosophila* (*Sun et al., 2020*). To estimate the dopamine-sensor kinetics

1032 in our head-fixed mice, we examined the phasic fluorescence transient occurring on unrewarded

1033 first-licks (0.5-3.3 s), which showed a stereotyped fast rise and decay with both sensors (*Figure*

1034 *2—figure supplement 3D-E*). While the transient was somewhat complex (reminiscent of phasic

1035 burst-pause responses sometimes observed for movement-related DAN activity (*Coddington and*

1036 *Dudman, 2018, 2019*), we measured the time for average fluorescence to decay from the peak of

1037 the transient to half the baseline-to-peak amplitude. We found decay $t_{1/2}$~75 ms for dLight1.1 and

1038 $t_{1/2}$~125 ms for DA_{2m} (*Figure 3—figure supplement 1*). Given that the dopaminergic ramping

1039 signals in our study evolved over several seconds, the kinetics of both dopamine sensors are thus

1040 fast enough that they should not have caused appreciable distortion of the slow ramping

1041 dynamics.

1042

1043 **Pearson's correlation of baseline/lamp-off to cue interval signals to first-lick time.**

1044 The mean SNc GCaMP6f signal during the "baseline" (2 s interval before the lamp-off event) or

1045 minimum lamp-off interval ("LOI;" -0.4 s to 0 s, the cue-time) was compared to the first-lick

1046 time for pooled trials in *Figure 2C* by calculating the Pearson correlation coefficient. There were

1047 at least 700 trials in each pooled set of trials (0.75-4 s included).

1048

1049 **DAN signal encoding model**

1050 To test the independent contribution of each task-related input to the photometry signal and

1051 select the best model, we employed a nested fitting approach, in which each dataset was fit

1052 multiple times (in "nests"), with models becoming progressively more complex in subsequent

1053 nests. The nests fit to the GCaMP6f photometry data employed the inputs $X^{(j)}$ at each jth nest:

1054 Null Model: $X^{(0)} = x_0$

1055 Nest 1: $X^{(1)} = X^{(0)} + $ tdTomato (tdt)

1056 Nest 2: $X^{(2)} = X^{(1)} + $ cue + first-lick

1057 Nest 3: $X^{(3)} = X^{(2)} + $ EMG/accelerometer

1058 Nest 4: $X^{(4)} = X^{(3)} + $ time-dependent baseline offset

1059 Nest 5: $X^{(5)} = X^{(4)} + $ stretch representing percentages of interval

1060 Overfitting was penalized by ridge regression, and the optimal regularization parameter for each

1061 nest was obtained by 5-fold cross-validation to derive the final model fit for each session. Model

1062 improvement by each input was assessed by the percentage loss improvement at the nest where

1063 the input first appeared compared to the prior nest. The loss improvement of Nest 1 was

1064 compared to the Null Model (the average of the photometry timeseries). The nested model of tdt

1065 control photometry signals was the same, except Nest 1 was omitted.

1066

1067 The GLM for each nest takes the form:

1068 $$Y = \theta X^{(j)}$$

1069 Where Y is the $1 \times n$ vector of the photometry signal across an entire behavioral session (n is the

1070 total number of sampled timepoints); $X^{(j)}$ is the $d \times n$ design matrix for nest j, where the rows

1071 correspond to the d_j predictors for nest j and the columns correspond to each of the n sampled

1072 timepoints of Y; and θ is the $d \times 1$ vector of fit weights.

1073

1074 Y is the concatenated photometry timeseries taken from trial start (lights off) to the time of first

1075 lick. Because of day-to-day/mouse-to-mouse variation (ascribable to many possible sources, *e.g.*,

1076 different neural subpopulations, expression levels, behavioral states, *etc.*), each session was fit

1077 separately.

1078

1079 The d_j design matrix predictors were each scaled (maximum amplitude 1) and grouped by input

1080 to the model. The timing-independent inputs were: 1. Null offset (x_0, 1 predictor), 2. tdt (1

1081 predictor), 3. cue (24 predictors), 4. first-lick (28 predictors), and 5. EMG/accelerometer (44

1082 predictors). The timing-dependent inputs were: 6. timing-dependent baseline offset (1 predictor),

1083 7. stretch (500 predictors).

1084

1085   To reduce the number of predictors, the cue, first-lick and EMG/accelerometer predictors

1086   (*Figure 5—figure supplement 1C*) were composed from sets of basis kernels as described

1087   previously (*Park et al., 2014; Runyan et al., 2017*). The cue basis kernels were spaced 0-500 ms

1088   post-cue and first-lick basis kernels were spaced -500 ms-0 ms relative to first-lick, the typically-

1089   observed windows of stereotypical sensory and motor-related neural responses. For nuisance

1090   movements (EMG/accelerometer), events were first discretized by thresholding (*Figure 5—*

1091   *figure supplement 1B*) and then convolved with basis kernels spanning -500 to 500 ms around

1092   the event. This window was consistent with the mean movement-aligned optical artifact

1093   observed in the tdt channel. The timing-dependent baseline offset was encoded as a constant

1094   offset spanning from lamp-off until first-lick, with amplitude taken as linearly proportional to the

1095   timed interval on the current trial. The timing-dependent stretch input was composed of 500

1096   predictors, with each predictor containing 1's tiling 0.05% of the cue-to-lick interval, and 0's

1097   otherwise (*Figure 5—figure supplement 1D*). Importantly, the stretch was not constrained in

1098   any way to form ramps.

1099

1100   Basis sets were optimized to minimize Training Loss, as calculated by mean squared error of the

1101   unregularized model:

1102 $$\mathrm{argmin}_X{}^{(j)}(\text{Training Loss}(\Theta) = 1/n * (Y - \Theta X^{(j)})^2)$$

1103

1104   Superfluous basis set elements that did not improve Training Loss compared to the Null Model

1105   were not included in the final model. Goodness of the training fit was assessed by Akaike

1106   Information Criterion (AIC), Bayesian Information Criterion (BIC), R^2, and Training Loss. The

1107   optimal, regularized model for each nest/session was selected by 5-fold cross-validation in which

1108   the regularization parameter, λ_j, was optimized for minimal average Test Loss:

1109 $\text{argmin}_{\lambda_j} (\text{Test Loss}(\Theta, \lambda_j) = 1/n * (Y - \Theta X^{(j)})^2 + \lambda_j |\Theta|^2)$

1110

1111   Test Loss for each optimal model was compared across nests to select the best model for each

1112   session. Models were refit with the optimal λ_j to obtain the final fit.

1113

1114   Model error was simulated 1,000 times by redrawing Θ coefficients consistent with the data

1115   following the method described by *Gelman and Hill, 2006*, and standard errors were propagated

1116   across sessions. The absolute value of each predictor was summed and divided by the total

1117   number of predictors for that input to show the contribution of the input to the model (*Figure*

1118   *5—figure supplement 1G*). To simulate the modeled session's photometry signal for each nest *j*,

1119   Yfit was calculated as $\Theta X^{(j)}$ and binned by the time of first-lick relative to the cue. The error in

1120   the simulation was shown by calculating $\text{Yfit}_{sim} = \Theta_{sim} X^{(j)}$ for 300 simulated sets of Θ_{sim}.

1121

1122   **Principal component analysis (PCA)**

1123   Unsmoothed ramping intervals for photometry timeseries were fit with PCA and reconstructed

1124   with the first three principal components (PCs). To derive a PCA fit matrix with ramping

1125   intervals of the same number of samples, the length of each trial was scaled up by interpolation

1126   to the maximum ramping interval duration:

1127   7 s – 0.7 s cue buffer – 0.6 s first-lick buffer = 5.7 s: 5,700 sample ramping interval

1128 Following PC-fitting, datasets were down-sampled to produce a fit of the correct time duration.

1129 Trials where the ramping interval was <0.1 s were excluded to exclude noise from down-

1130 sampling.

1131

1132 **First-lick time decoding model**

1133 A nested, generalized linear model was derived to predict the first-lick time on each trial in a

1134 session and quantify the contribution of previous reward history and photometry signals to the

1135 prediction. The model was of the form:

1136 $$\log(y) = bx$$

1137 where y is the first-lick time, b is a vector of fit coefficients and x is a vector of predictors. The

1138 nested model was constructed such that predictors occurring further back in time (such as reward

1139 history) and confounding variables (such as tdt photometry signals) were added first to determine

1140 the additional variance explained by predictors occurring closer to the time of first-lick, which

1141 might otherwise obscure the impact of these other variables. The predictors, in order of nesting,

1142 were:

1143   Nest 0: b_0 (Null model, average log-first-lick time)

1144   Nest 1: $b_1 = b_0$ + first-lick time on previous trial (trial "n-1")

1145   Nest 2-5: $b_2 = b_1$ + previous trial outcome (1,0)*

1146   Nest 6: $b_3 = b_2$ + median photometry signal in 10s window before lamp-off ("ITI")

1147   Nest 7: $b_4 = b_3$ + median photometry signal from lamp-off to cue ("lamp-off interval")

1148   Nest 8: $b_5 = b_4$ + tdt threshold crossing time**

1149   Nest 9: $b_6 = b_5$ + GCaMP6f threshold crossing time**

1150

1151   where all predictors were normalized to be in the interval (0,1).

1152

1153   * Outcomes included (in order of nest): Reaction (first-lick before 0.5 s), Early (0.5-3.333 s),

1154   Reward (3.333-7 s), ITI (7-17 s). No-lick was implied by all four outcomes encoded as zeros.

1155   ** Details on threshold-crossing time and alternative models included in *Methods: Derivation of*

1156   *threshold and alternative decoding models*.

1157

1158   To exclude the sensory- and motor-related transients locked to the cue and the first-lick events in

1159   the threshold-crossing nests, the ramping interval was conservatively defined as 0.7 s post-cue up

1160   until 0.6 s before first-lick, and the minimum ramping interval for fitting was 0.1 s. Thus, for a

1161   trial to be included in the model, the first lick occurred between 1.4 s to 17 s (end of trial).

1162

1163   Initial model goodness of fit was assessed by R^2, mean-squared loss and BIC. Models were 5-

1164   fold cross-validated with ridge regression at each nest to derive the final models, as described

1165   above. 95% confidence intervals on model coefficients were calculated by 2-sided t-test with

1166   standard errors propagated across sessions.

1167

1168   **Derivation of threshold and alternative decoding models**

1169   *Derivation of threshold models*

1170   As a metric of the predictive power of ramping DAN signals on first-lick time, we derived a

1171   threshold-crossing model. A threshold-crossing event was defined as the first time after the cue

1172   when the photometry signal exceeded and remained above a threshold level up until the time of

1173   first-lick on each trial. Importantly, while the analysis approach is reminiscent of pacemaker-

1174 accumulator models for timing, we make <u>no</u> claims that the analysis is evidence for pacemaker-

1175 accumulator models. Rather threshold-crossing times provided a convenient metric to compare

1176 the rate of increase in signals between trials.

1177

1178 Photometry timeseries for GCaMP6f and tdt were de-noised by smoothing with a 100 ms

1179 Gaussian kernel (kernel was optimized by grid screen of kernels ranging between 0-200 ms to

1180 minimize noise without signal distortion). To completely exclude the sensory- and motor-related

1181 transients locked to the cue and the first-lick events, the ramping interval was conservatively

1182 defined as 0.7 s post-cue up until 0.6 s before the first-lick. To eliminate chance crossings due to

1183 noise, we imposed a stiff, debounced threshold condition: to be considered a threshold crossing

1184 event, the photometry signal had to cross the threshold from low-to-high and remain above this

1185 level until the end of the ramping interval.

1186

1187 To derive an unbiased threshold for each session, we tested 100 evenly-spaced candidate

1188 threshold levels spanning the minimum-to-maximum photometry signal during the ramping

1189 interval for each session. Depending on threshold level, some trials never crossed, *i.e.*, signal

1190 always remained below threshold or started and ended above threshold. Thus, the lowest

1191 candidate threshold for which there was a maximum number of trials crossing during the timing

1192 interval was selected as the "mid-level" threshold-crossing point. This threshold was specific to

1193 each photometry signal tested on each session. Threshold-crossing time was included in the

1194 decoding model as the normalized time on the ramping interval (0,1). If a trial never crossed

1195 threshold, it was encoded as a zero. If no trials ever crossed threshold, the threshold predictor

1196   was encoded as a vector of ones, thus penalizing the model for an additional predictor but

1197   providing no new information.

1198

1199   *Multi-threshold Model*

1200   An alternative model employed 3 unbiased thresholds: 1) the lowest threshold with ≥50 trials

1201   crossing ("min"); 2) the lowest threshold with the most crossings ("mid," described above); and

1202   3) the highest threshold with ≥50 trials crossing ("max"). For tdt datasets, trials rarely met the

1203   monotonic threshold constraint (usually the signals oscillated above and below the threshold

1204   throughout the ramping interval, failing to meet the debouncing constraint). Thus, to include tdt

1205   signals as conservatively as possible, we relaxed the 50-trial minimum constraint, taking the

1206   threshold with the most trials crossing, which was usually around 10 or fewer. The addition of

1207   more thresholds did not substantially improve the cross-validated model compared to the single-

1208   threshold model (*Figure 6—figure supplement 1*).

1209

1210   *Principal component analysis (PCA) threshold-crossing models*

1211   In another version of the decoding model, the threshold-crossing procedures were applied to

1212   ramping intervals fit with the first three PCs (as described in *Methods: Principal Component*

1213   *Analysis (PCA)*) to derive a PCA version of the single-threshold and multi-threshold models.

1214   PCA analysis on tdt datasets showed no consistent PCs, and thus these PCs were not included in

1215   the decoding model. Instead, the actual tdt data was employed in the threshold model as in the

1216   other models described.

1217

1218   **Hierarchical Bayesian Modeling of Single-trial Dynamics.**

1219 The probability of each single-trial SNc GCaMP6f signal belonging to a ramp vs. step Model

1220 Class was determined via Hierarchical Bayesian Model fitting with probabilistic programs

1221 written in the novel probabilistic programming language, Gen.jl, which is embedded in the Julia

1222 Programming Language (*Cusumano-Towner et al., 2019*). The top of the model hierarchy was

1223 the model class (linear ramp *vs*. step function) and the lower level was the respective

1224 parameterization of the two model classes (described below).

1225

1226 The probability of the step *vs*. ramp model class was inferred with data-driven inference. The

1227 best fit (step or ramp and parameterization) for each trial was calculated across 20 iterations

1228 (Gen *Traces*) of hierarchical modeling with 50 rounds of probabilistic refinement (computation

1229 *via* Gen Importance Resampling) per iteration (in model testing, models typically converged to

1230 their steady-state probability of model class within only 30 rounds of refinement, but 50 rounds

1231 were used conservatively to reduce the likelihood of suboptimal classifications).

1232

1233 *Data-driven inference procedure:* Each iteration of model fitting began at the top level of the

1234 hierarchy with a coin toss: with 50% probability, the probabilistic program would initialize with

1235 a model of either the Ramp or Step class. For data-driven inference, a Gen *Proposal* for the

1236 parameterization for this model class was then probabilistically generated. Data-driven proposals

1237 were designed to improve fitting efficiency and reduce computation time, and this allowed for

1238 faster convergence and better model fits as determined by the fit log-likelihood. The proposal

1239 heuristics were as follows:

1240   Ramp model: A data-driven proposal was generated by dynamic noise random sample

1241 consensus (RANSAC; *Cusumano-Towner and Mansinghka, 2018*) with additional data-driven

1242 constraints (see function `ransac_assisted_model_selection_proposal` in the Gen Github

1243 files):

1244 1. SLOPE, *a*. The maximum data-supported slope was used to set the variance of slope

1245 sampling:

1246 $$a \sim Gaussian(\text{RANSAC-sampled slope}, maxslope/2).$$

1247 where *maxslope* was defined as the difference of the maximum and minimum signal within

1248 the trial dataset divided by the total duration of the trial (by definition, the largest slope

1249 supported by the data).

1250 2. INTERCEPT, *b*. The initial search for the intercept ("*b-max*") was calculated as the

1251 intercept for the calculated *maxslope* parameter), and this was used to set the noise level

1252 on sampling of the intercept parameter:

1253 $$b \sim Gaussian(\text{RANSAC-sampled intercept}, b\text{-}max/2)$$

1254 3. NOISE, σ. Parametrized noise level was sampled as:

1255 $$\sigma \sim Beta(\alpha, \beta)$$

1256 where α,β are the parameters of the beta distribution with mode=std(signal).

1257

1258 <u>Step model</u>**:** The data-driven proposal included two constraints/heuristics:

1259 1. STEPTIME. *Derivative constraint*: To avoid sampling all unlikely step-times, *steptimes*

1260 were sampled uniformly from the timepoints where the derivative of the signal was in

1261 the highest 5% of the signal's derivative across the trial dataset:

1262 $$steptime \sim uniform(\text{indices of } 95^{th} \text{ percentile of derivative of the signal})$$

1263   2. LEFT and RIGHT SEGMENTS. Once a *steptime* was sampled, likely *left* and *right*

1264   segment amplitudes were sampled near the mean of the signal on either side of the step,

1265   *e.g.*:

1266   *left ~ Gaussian*(mean(signal left of *steptime*), std(signal left of *steptime*))

1267   *right ~ Gaussian*(mean(signal right of *steptime*), std(signal right of *steptime*))

1268   3. NOISE, σ. The noise level was sampled as in the ramp model,

1269 $\sigma \sim Beta(\alpha, \beta)$

1270   except α,β were the parameterization of a Beta distribution with mode equal to the

1271   standard deviation of the signal left of *steptime*.

1272

1273   After model initialization for each Trace, 50 rounds of Importance Resampling of the

1274   hierarchical model were then conducted, each time randomly generating ramp or step hypotheses

1275   from the proposal heuristics. On each round, the best fitting hypothesis was retained, such that

1276   each of the 20 Trace iterations of model classification returned one optimized model from the 50

1277   rounds of Importance Resampling.

1278

1279   The probability of the model class for each single-trial was then defined as the proportion of the

1280   20 Trace iterations that found the optimal model to be derived from that model class (*e.g.,* if the

1281   model returned 15 step-fits and 5 ramp-fits, the p(ramp) was 0.25). Examples of the 20 Trace

1282   iterations for two sample trials are shown in *Figure 6—figure supplement 2B.*

1283

1284   To determine whether the step model detected step-functions in the GCaMP6f dataset, the step

1285   model was inferred alone to find step-fits for every trial, and single-trial signals were realigned to

1286   the optimal *steptime* (GCaMP6f, tdTomato, EMG, *Figure 6—figure supplement 4A-B*).

1287

1288   **Single-trial dynamics analysis with geometric modeling ("Multiple threshold modeling").**

1289   The multi-threshold procedure described above was also employed to determine whether single-

1290   trial ramping dynamics were more consistent with a continuous ramp *vs*. discrete step dynamic

1291   on single-trials. The threshold-crossing time for each trial was regressed against its first-lick time,

1292   and the slope of this relationship was reported, as well as the variance explained.

1293

1294   **Single-trial variance analysis for discrete step dynamics.**

1295   For discrete step single trial dynamics to produce ramping on average, the time of the step across

1296   trials must be distributed throughout the trial interval (importantly, a peri-motor spike occurring

1297   consistently just before first-lick *cannot* give rise to continuous ramping dynamics on average).

1298   As such, the variance in the GCaMP6f signals across trials for similar first-lick times should be

1299   minimal near the time of the cue (when few trials have stepped) *and* near the time of the first-lick

1300   (when all of the trials have stepped). This predicts an inverted-U shaped relationship of signal

1301   variance across trials *vs*. position in the timing interval.

1302

1303   To compare variance across trials equitably, trials were first aligned to the cue and pooled by

1304   first-lick time in pools of 1s each (1-2 s, 2-3 s, *etc*.) truncating at the earliest first-lick time within

1305   the pool. The variance in GCaMP6f signals across trials within a pool was quantified in 10%

1306   percent increments of time from the cue up to the earliest first-lick time in the pool (*i.e*, 1-2 s

1307   pool truncated at 1 s, divided into 100 ms increments). Measuring variance by percent of elapsed

1308   time within pool allowed pooling of trials across the entire session. The shape of the variance *vs*.

1309   percent of timed interval elapsed was compared to the inverted-U shape prediction to assess for

1310   discrete step dynamics.

1311

1312   **Optogenetics—determining the physiological range for activation experiments**

1313   To test whether optogenetic manipulations during the self-timing task were in the physiological

1314   range, we assessed the magnitude of the effect of activation on dopamine release in the DLS by

1315   simultaneous photometry recordings with optical activation (*Figure 7—figure supplement 2*). In

1316   two DAT-cre mice, we expressed ChrimsonR bilaterally in SNc DANs and the fluorescent

1317   dopamine indicator dLight1.1 bilaterally in DLS neurons. SNc cell bodies were illuminated

1318   bilaterally (ChrimsonR 550 nm lime or 660 nm crimson, 0.5-5 mW) on 30% of trials (10 Hz, 10

1319   or 20 ms up-time starting at cue onset and terminating at first-lick). dLight1.1 was recorded with

1320   35 µW 475 nm blue LED light at DLS. To avoid crosstalk between the stimulation LED and the

1321   photometry recording site, the brief stimulation up-times were omitted from the photometry

1322   signal and the missing points filled by interpolation between the adjacent timepoints.

1323

1324   In a few preliminary sessions, we also explored whether we could evoke short-latency licking

1325   (*i.e*., within a few hundred ms of the stimulation) if light levels were increased above the

1326   physiological range for DAN signals. Rather than eliciting immediate licking, higher light levels

1327   produced bouts of rapid, nonpurposive limb and trunk movements throughout stimulation, and

1328   task execution was disrupted. The animals appeared to have difficulty coordinating the extension

1329   of the tongue to touch the lick spout. Simultaneous DLS dopamine detection showed large,

1330   sustained surges in dopamine release throughout the period of stimulation, with an average

1331   amplitude comparable to that of the reward transient (*Figure 7—figure supplement 2, right*).

1332   This extent of dopamine release was never observed during unstimulated trials. Consequently, to

1333   avoid overstimulation in activation experiments, we kept light levels well below those that

1334   generated limb and trunk movements.

1335

1336   **Optogenetics—naïve/expert control sessions.**

1337   To determine whether optogenetic stimulation directly elicited or prevented licking, licking

1338   behavior was first tested outside the context of the self-timed movement task on separate

1339   sessions in the same head-fixed arena but with no cues or behavioral task. Opsin-expressing mice

1340   were tested before any exposure to the self-timed movement task ("Naïve") as well as after the

1341   last day of behavioral recording ("Expert"). In ChR2 control sessions, stimulation (5 mW 425

1342   nm light, 3 s duration, 10 Hz, 20% duty cycle) was applied randomly at the same pace as in the

1343   self-timed movement task. stGtACR2 control sessions were conducted similarly (12 mW 425

1344   mW light, 3 s duration, constant illumination); but to examine if inhibition could block ongoing

1345   licking, we increased the baseline lick-rate by delivering juice rewards randomly (5% probability

1346   checked once every 5 s).

1347

1348   **Optogenetics—self-timed movement task.**

1349   SNc DANs were optogenetically manipulated in the context of the 3.3 s self-timed movement

1350   task. To avoid overstimulation, light levels were adjusted to be subthreshold for eliciting overt

1351   movements as described above, and mice were not stimulated on consecutive days.

1352

1353 <u>Activation</u>: SNc cell bodies were illuminated bilaterally (ChR2: 0.5-5 mW 425 nm blue LED

1354   light; ChrimsonR 550 nm lime or 660 nm crimson) on 30% of trials (10 Hz, 10 or 20% duty

1355   cycle starting at cue onset and terminating at first-lick). DAN terminals in DLS were

1356   stimulated bilaterally via tapered fiber optics on separate sessions.

1357 <u>Inactivation</u>: SNc cell bodies were illuminated bilaterally (stGtACR2: 12 mW 425 nm blue light)

1358   on 30% of trials (constant illumination starting at cue onset and terminating at first lick).

1359

1360 **Quantification of optogenetic effects.**

1361 The difference in the distribution of trial outcomes between stimulated and unstimulated trials on

1362 *each session* was quantified in four ways.

1363   1. <u>2-Sample Unsigned Kolmogorov-Smirnov Test.</u>

1364   2. <u>Difference in empirical continuous probability distribution function (cdf)</u>. The difference

1365   in the integral of the stimulated and unstimulated cdf (dAUC) was calculated for each

1366   session from 0.7-7 s. Effect size was quantified by permutation test, wherein the identity

1367   of each trial (stimulated or unstimulated) was shuffled, and the distribution of dAUCs for

1368   the permuted cdfs was calculated 10,000x. Results were reported for all sessions.

1369   3. <u>Difference in mean movement time</u>. Movement times on stimulated and unstimulated

1370   trials were pooled and the distribution of movement time differences was determined by

1371   non-parametric bootstrap, in which a random stimulated and unstimulated trial were

1372   drawn from their respective pools 1,000,000x and the difference taken. The mean of each

1373   session's bootstrapped distribution was compared across sessions by the 1,000,000x

1374   bootstrapped difference of the mean between sessions of different categories.

1375   4. <u>Difference in median movement time</u>. Same as above but with median.

1376

1377 **Single-trial probabilistic movement state decoding model.**

1378 The probability of transitioning to a movement state, $s_t=1$, at time=t was decoded with a logistic

1379 generalized linear model of the form:

1380
$$p(s_t=1) = \text{logit}(bX_t)$$

1381 where X_t is a vector of predictors for the timepoint, t, and b is the vector of fit coefficients. The

1382 vector of predictors was comprised of the GCaMP6f signal at every timepoint (the current time, t)

1383 as well as the signal history, represented as 200 ms-wide signal averages moving back in time

1384 from t. Previous trial history (n-1th and n-2th first-lick times and reward/no-reward outcomes) did

1385 not contribute significantly to the model during model selection and were thus omitted (see

1386 Model Selection, below).

1387

1388 Movement state, s_t, was defined as a binary variable, where state=0 represented all timepoints

1389 between the cue up until 160 ms before the first-lick detection (to exclude any potential peri-

1390 movement responses), and state=1 represented the timepoint 150 ms before the first-lick.

1391 Because there were many more state=0 than state=1 samples in a session, state=0 points were

1392 randomly down-sampled such that states were represented equally in the fit. To avoid randomly

1393 sampling a particular model fit by chance, each dataset was fit on 100 randomly down-sampled

1394 (bootstrapped) sets, and the average fit across these 100 sets was taken as the model fit for the

1395 session.

1396

1397   GCaMP6f signals were smoothed with a 100 ms gaussian kernel and down-sampled to 100 Hz.

1398   The GCaMP6f predictors were then nested into the model starting with those furthest in time

1399   from the current timepoint, t:

1400

1401   Nest 0:   b0 (Null model)

1402   Nest 1:   b1 = b0 + mean GCaMP6f 1.8:2.0 s before current time=t

1403   Nest 2:   b2 = b1 + mean GCaMP6f 1.6:1.79 s before current time=t

1404   Nest 3:   b3 = b2 + mean GCaMP6f 1.4:1.59 s before current time=t

1405   Nest 4:   b4 = b3 + mean GCaMP6f 1.2:1.39 s before current time=t

1406   Nest 5:   b5 = b4 + mean GCaMP6f 1.0:1.19 s before current time=t

1407   Nest 6:   b6 = b5 + mean GCaMP6f 0.8:0.99 s before current time=t

1408   Nest 7:   b7 = b6 + mean GCaMP6f 0.6:0.79 s before current time=t

1409   Nest 8:   b8 = b7 + mean GCaMP6f 0.4:0.59 s before current time=t

1410   Nest 9:   b9 = b8 + mean GCaMP6f 0.2:0.39 s before current time=t

1411   Nest 10:   b10 = b9 + GCaMP6f signal at current time=t

1412

1413   Nesting the predictors from most distant in time to most recent permitted observation of the

1414   ability of more proximal signal levels to absorb the variance contributed by more distant signal

1415   history.

1416

1417   The fitted hazard function was then found as the average probability of being in the movement

1418   state across all trials in the session as calculated from the average model fit. Because s_t=0 states

1419   were significantly downsampled during fitting, this rescaled the fit hazard. Thus, to return the fit

1420   hazard to the scale of the hazard function calculated from the behavioral distribution, both the fit

1421   hazard and true hazard function were normalized on the interval (0,1), and the goodness of fit

1422   was assessed by R^2 comparison of the fit and true hazard functions. This metric was similar

1423   between individual session fits as well as the grand-average fit across all animals and sessions.

1424

1425   To guard against overfitting, this procedure was repeated on the same datasets, except the

1426   datasets were shuffled before fitting to erase any non-chance correlations between the predictors

1427   and the predicted probability of being in the movement state.

1428

1429   *Model selection*

1430   To evaluate the contribution of task performance history to the probability of being in the

1431   movement state at time=*t*, we could not observe every timepoint in the GCaMP6f trial period

1432   timeseries as we did in the final model because the trial history for a given timepoint was the

1433   same for all other points in the trial; hence this created bias because the movement state=1 was

1434   represented for all trials, but the likelihood of the a trial's 0 state being represented after down-

1435   sampling was dependent on the duration of the trial (*i.e.,* first-lick time). Consequently, model

1436   selection was executed on a modified version of the model that ensured that each trial would

1437   only be represented *one time at most* in the fit. Because this greatly reduced the power of the

1438   model, model selection was conducted on sessions from the two animals with the highest S:N

1439   ratio and most trials to ensure the best chance of detecting effects of each predictor (*Figure 8—*

1440   *figure supplement 1*).

1441

1442 The set of permutations of GCaMP6f signal and task history were fit separately, and the best

1443 model selected by BIC (though notably AIC and AICc were in agreement with the BIC

1444 selection). Each model was fit in "time-slices"—windows of 500 ms from the time of the cue up

1445 until the first-lick. Only one point for each trial was fit within this window to ensure the

1446 movement state within the window was uniquely represented. For each time-slice model, the

1447 GCaMP6f signal for each trial was thus averaged within the time-slice window, and the

1448 movement state was 1 only if the movement state occurred sometime within the window. The

1449 model fit for a session was taken as the average model fit across each of the time-slices. Notably,

1450 a time-slice required a sufficient number of trials to be present (either in the $s_t=0$ or terminating

1451 in the movement state $s_t=1$) for the fit to converge; once the first-lick occurred for a trial, it did

1452 not contribute data to later time-slices. The source data files for *Figure 8—figure supplement 1*

1453 contain plots of all time-slice coefficient fits, including for models with insufficient numbers of

1454 trials to converge.

1455

1461

1473

1474 **COMPETING INTERESTS**

1475 J.A.A. is a co-founder of OptogeniX, which produces the tapered optical fibers used in some

1476 experiments.
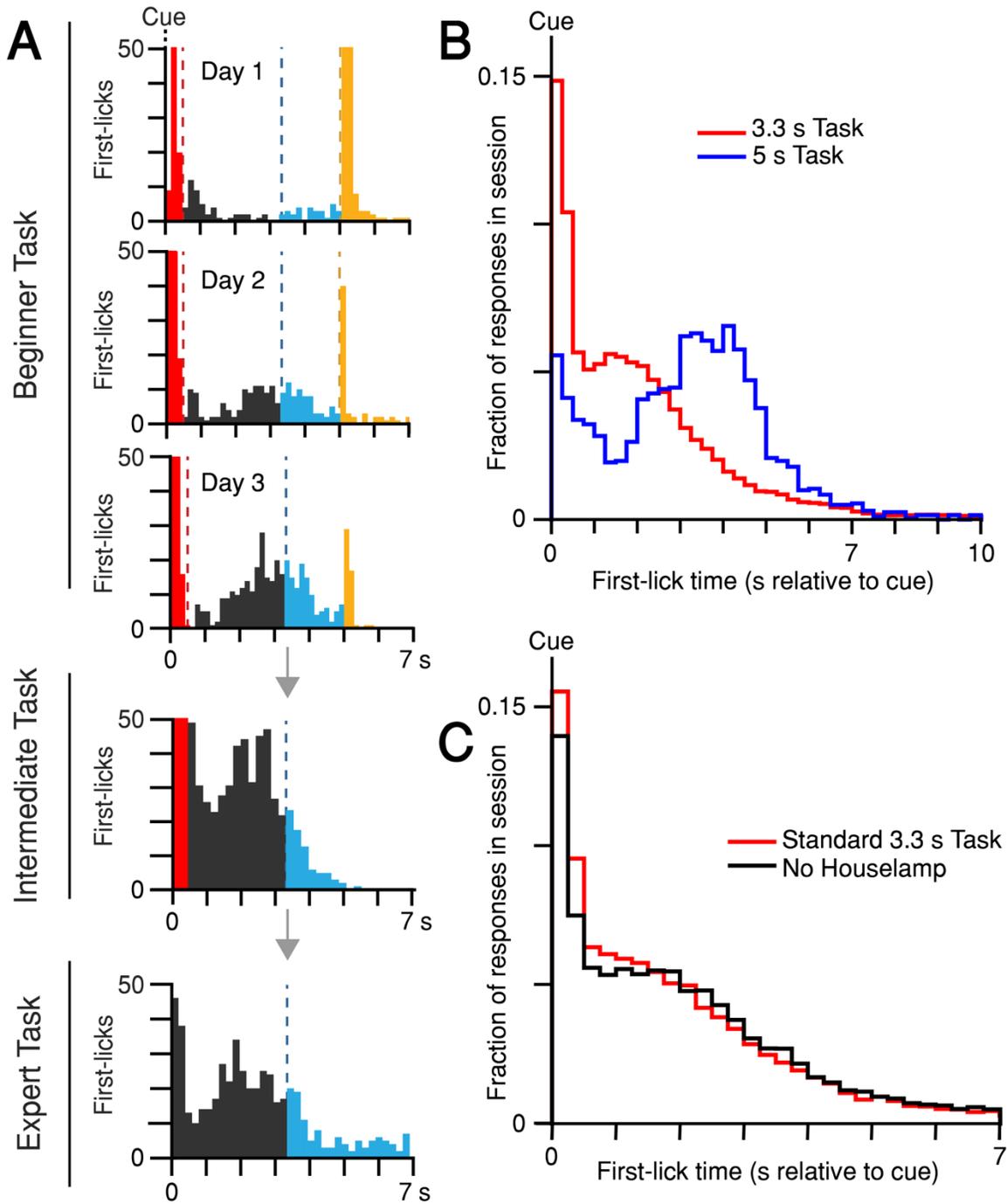
1478 **Figure 1—figure supplement 1.** Self-timed movement task learning and variations. (**A**) Task learning.
1479 Histogram of first-lick times from single sessions at different stages of training (red: reaction, grey: early,
1480 blue: operant-rewarded, yellow: Pavlovian-rewarded). Bars >50 first-licks truncated for clarity. (**B**) Mice
1481 adjust behavior to the timing-contingencies of the task. First-lick time distributions from tasks with
1482 different target timing intervals. Red: 3.3 s reward-boundary. Blue: 5 s reward-boundary (all sessions, all
1483 mice). (**C**) Mice time their first-licks relative to the start cue, not the houselamp. First-lick time
1484 distributions during behavior with (red) and without (black) houselamp events (4 mice, 4-5
1485 sessions/mouse on each version of the task). Source data: *Figure 1—source data*.
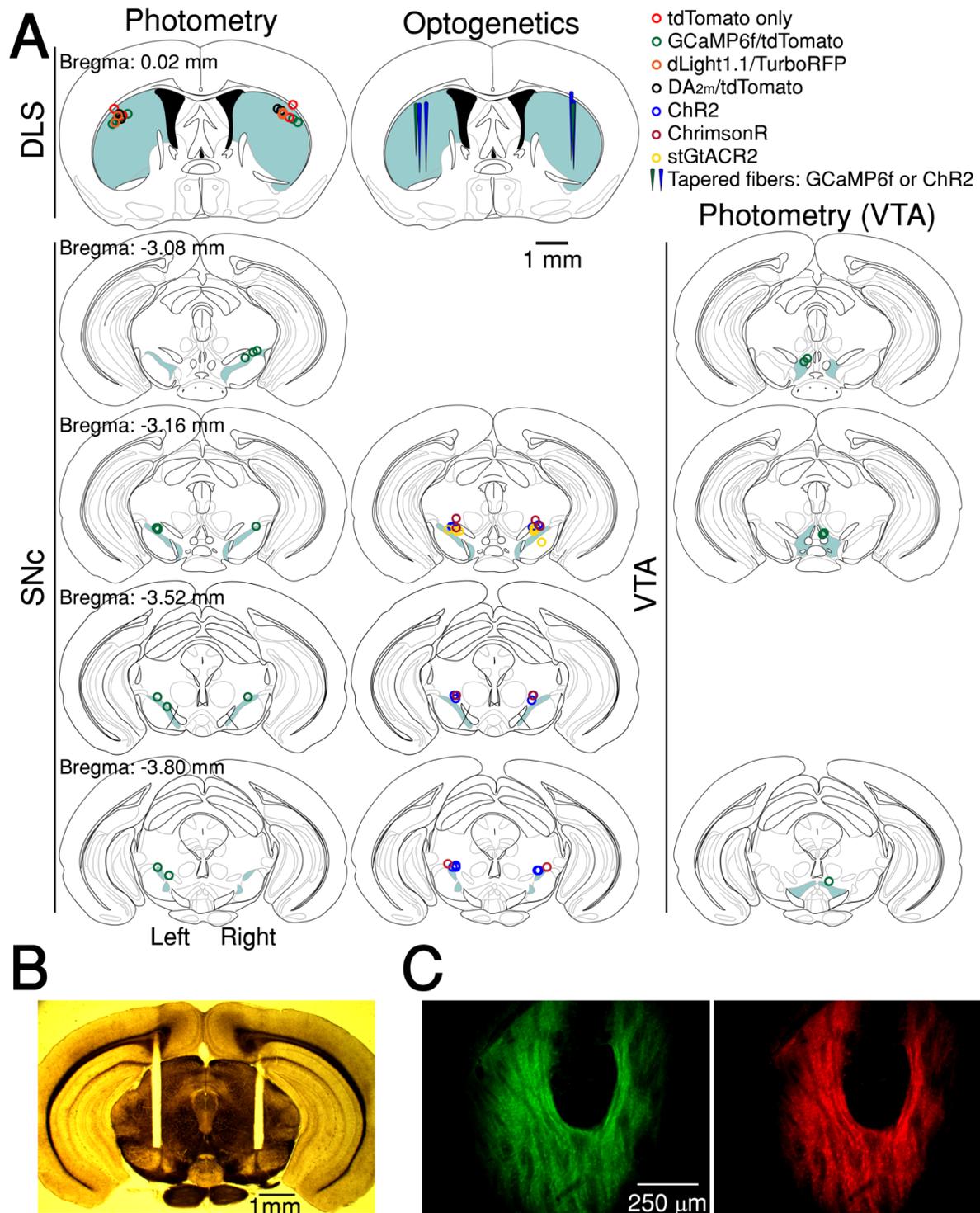
Figure 1—figure supplement 2. Fiber optic placement and histology**. (A)** Approximate fiber positions for all mice. **(B)** Brightfield microscopy with polarized filter on a freshly cut brain slice showing bilateral fiber placement at SNc (from stGtACR2 experiment). **(C)** Example of co-expression of green (DA$_{2m}$) and red (tdTomato) fluorophores relative to fiber optic tip.
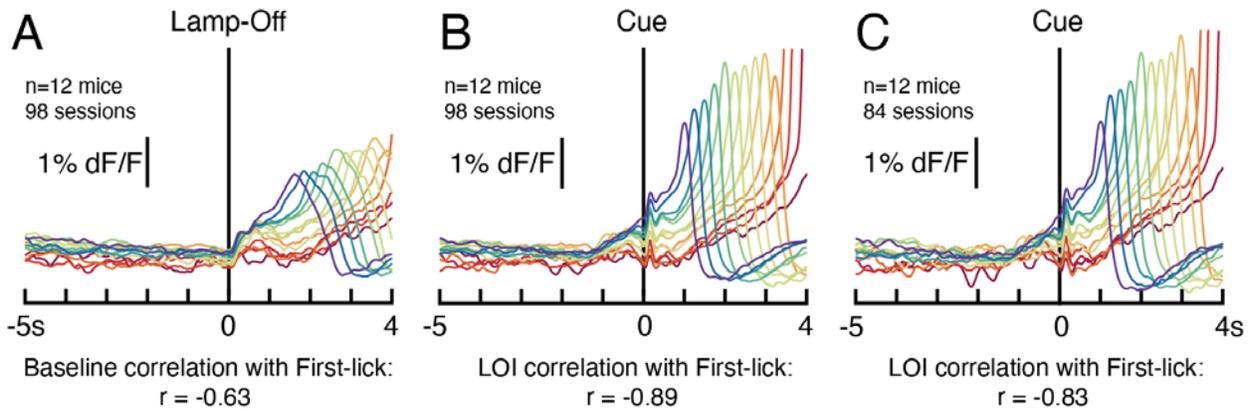
Figure 2—figure supplement 1. Baseline correlation of dopaminergic signal with first-lick time is not dependent on the duration of the lamp-off interval. **(A)** SNc GCaMP6f dopaminergic signals aligned to the lamp-off event (n=12 mice, all 98 sessions). "Baseline:" 2 s interval before lamp-off event. **(B)** SNc GCaMP6f dopaminergic signals aligned to the cue, all sessions. "LOI:" Lamp-Off-Interval between lamp-off and cue. **(C)** 14/98 sessions showed a small relationship between LOI duration and first-lick time ($R^2 < 0.04$ for 13/14 sessions, sign of correlation inconsistent among sessions). Omitting these 14 sessions did not eliminate the Baseline or Lamp-Off Interval correlation between dopaminergic signal amplitude and first-lick timing. Source data: *Figure 2—source data 2*.
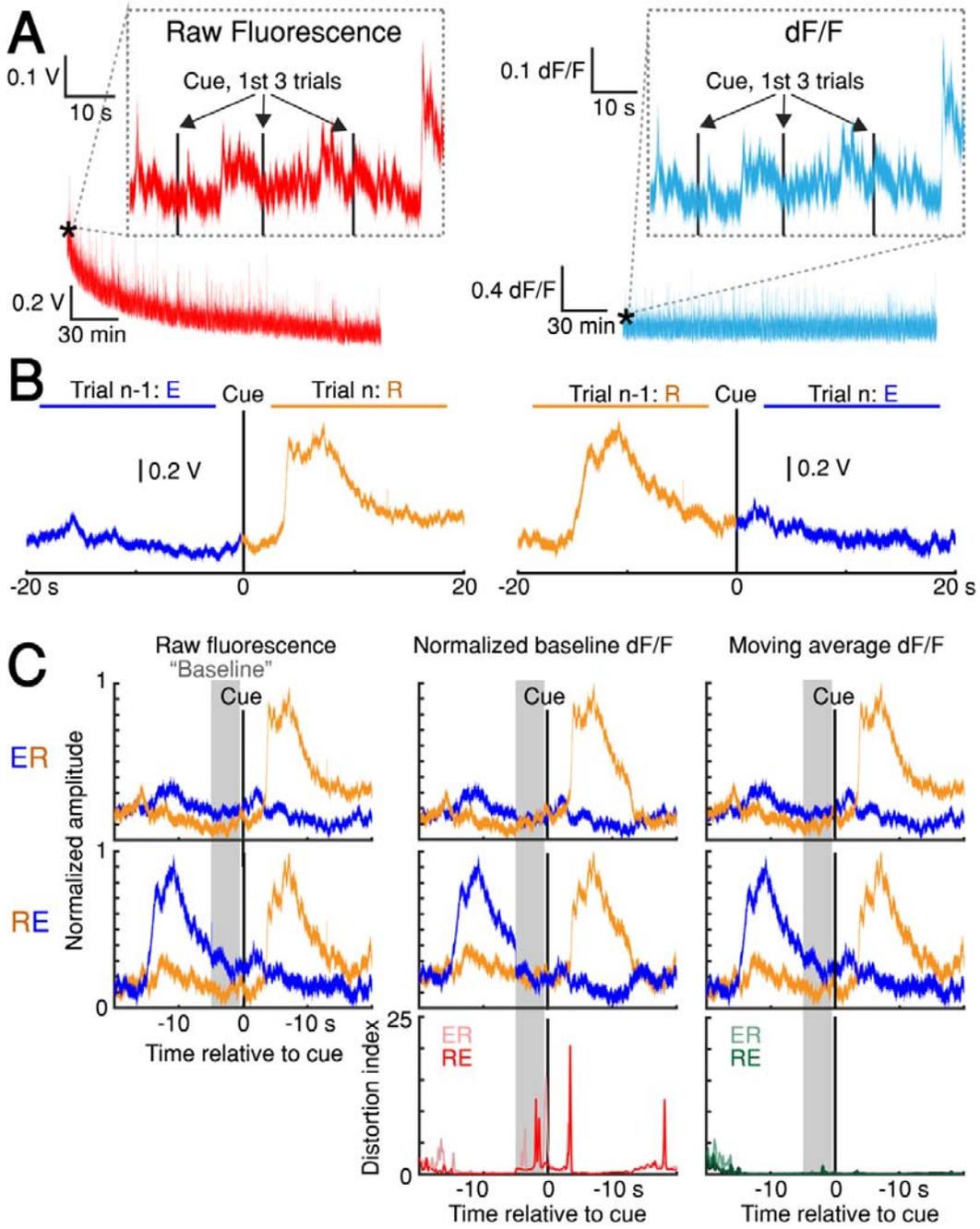
Figure 2—figure supplement 2. dF/F method validation. **(A)** Left: slow, raw fluorescence bleaching across one session. Left inset: Minimal bleaching occurs across the first 3 trials (~1 min). Right: dF/F removes slow bleaching dynamics. Right inset: The same 3-trial window shown for dF/F signal. **(B)** Average raw fluorescence on paired, consecutive trials from one session aligned to cue on the nth trial. Left: $n-1^{th}$ trial was early, n^{th} trial was rewarded ("ER" condition). Right: "RE" condition (See *Methods: dF/F method characterization and validation*). **(C)** Comparison of baseline GCaMP6f signals on paired, consecutive trials aligned to cue. Columns: three different versions of the signal (Raw fluorescence, Normalized baseline dF/F method, Moving average dF/F method). Top row: ER condition; middle row: RE condition; bottom row: distortion index. Red distortion index plot shows only Normalized baseline method. Green distortion index plot shows overlay of Moving Average, Low-Pass Filter, and Multiple Baseline dF/F Methods because the difference in signal distortion between these methods was indistinguis... : *Figure 2— source data*
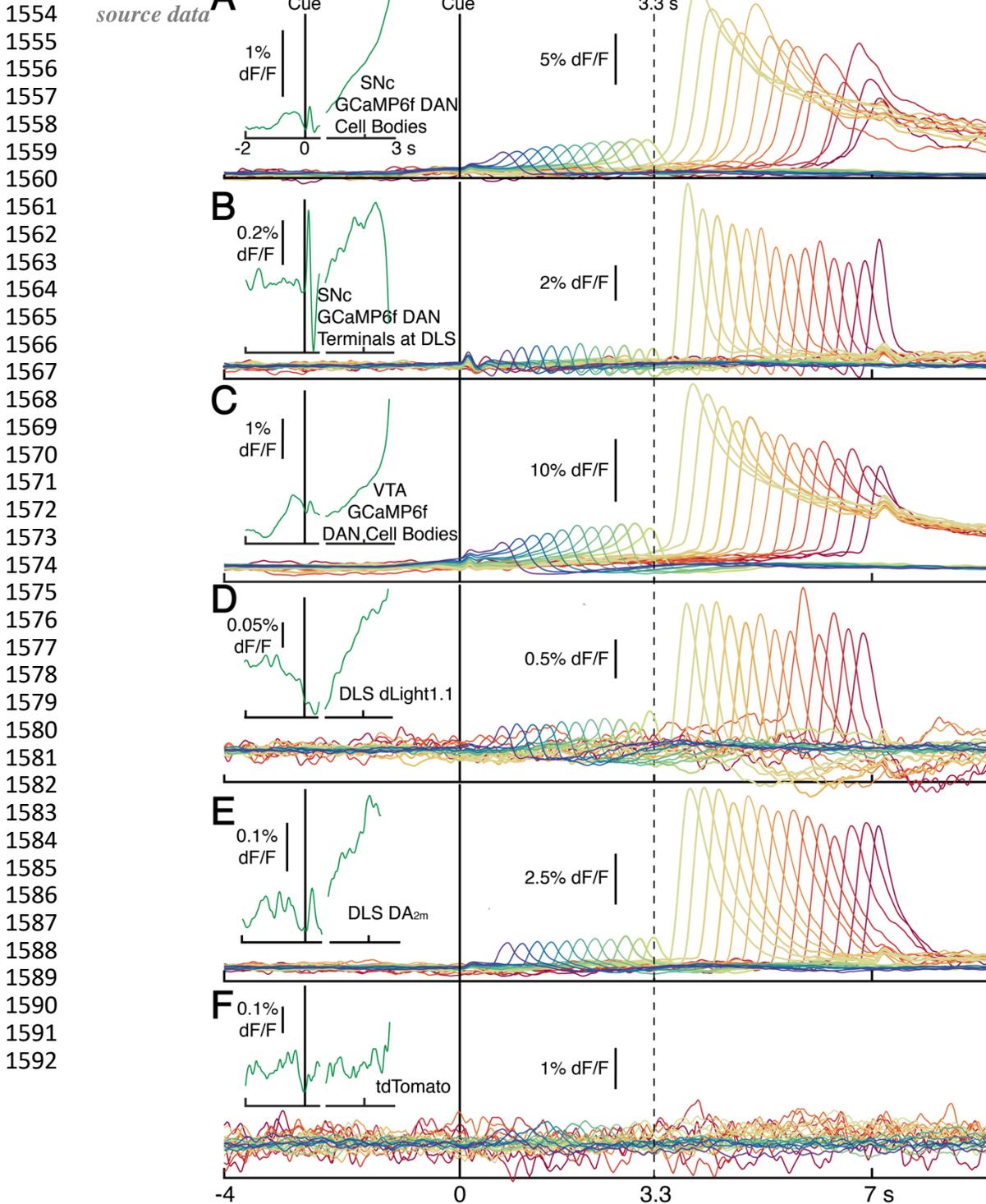
1593
1594
1595
1596
1597
1598   **Figure 2—figure supplement 3.** Average photometry signals, pooled every 250 ms by first-lick time,
1599   spanning 0.5 s (purple) to 7 s (red). Signals in main panels aligned only to cue, not first-lick. **(A)** Average
1600   DAN GCaMP6f signals at SNc cell bodies (12 mice). **(B)** DAN GCaMP6f signals at axon terminals in
1601   DLS (10 mice). **(C)** Striatal dopamine detection with dLight1.1 at DLS (5 mice). **(D)** Striatal DA_{2m}
1602   signals at DLS (4 mice). **(E)** DAN GCaMP6f signals at VTA cell bodies (4 mice). **(F)** tdTomato signals.
1603   **Insets (left)**: Cue *and* lick-aligned average signals for a single time bin before first-lick to show pre-lick
1604   ramping present in all dopaminergic signals. Left of axis break: aligned to cue. Right of axis break:
1605   aligned to first-lick. Traces plotted up until 150 ms before first-lick. Source data: *Figure 2—source data*
1606   *4*.
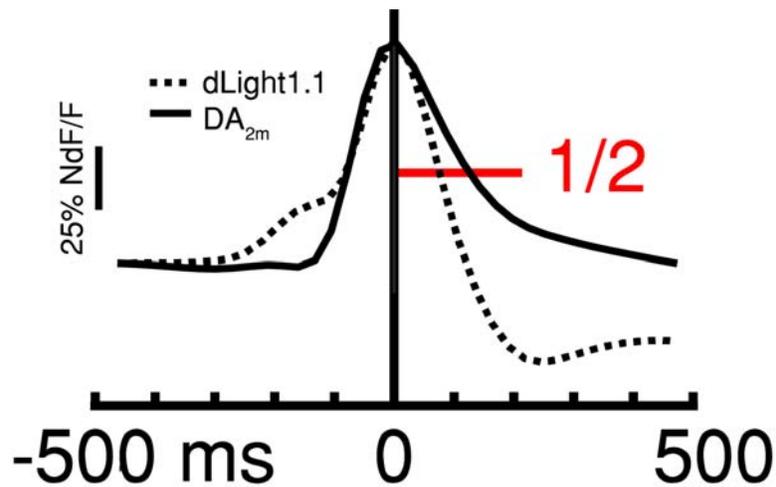
1607
1608
1609
1610   **Figure 3—figure supplement 1.** Comparison of dLight1.1 (dashed) and DA_{2m} (solid) kinetics
1611   surrounding peak of unrewarded transient (first-lick: 0.5-3.3 s). Red line: ½ baseline-to-peak amplitude
1612   for measuring decay $t_{1/2}$ (see *Methods*). Source data: *Figure 3—source data*.
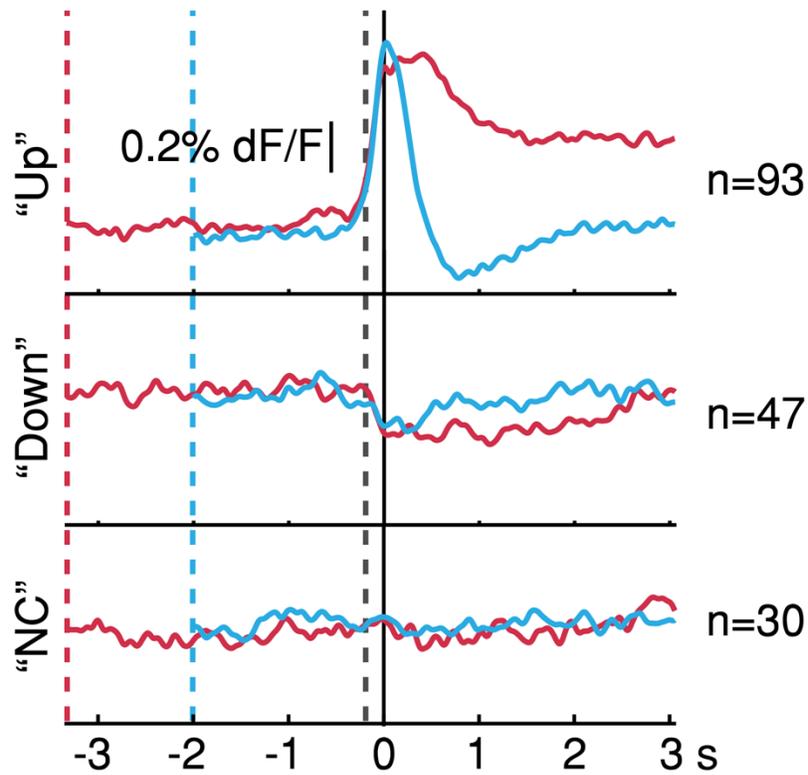
77
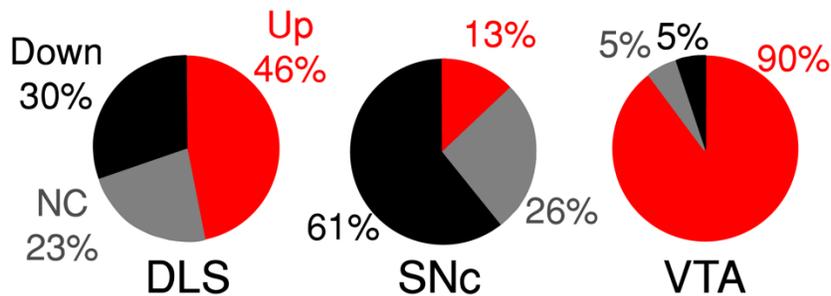
Figure 4—figure supplement 1. Average tdTomato optical artifacts (aligned to first-lick time) showed inconsistent directions even within the same session. Averages for all three types of artifact (consistently up, "Up"; consistently down, "Down"; and not consistent "NC") shown for all sessions. Pie plots: Breakdown of average tdt artifact direction by session at each recording site. Source data: *Figure 4—source data*.
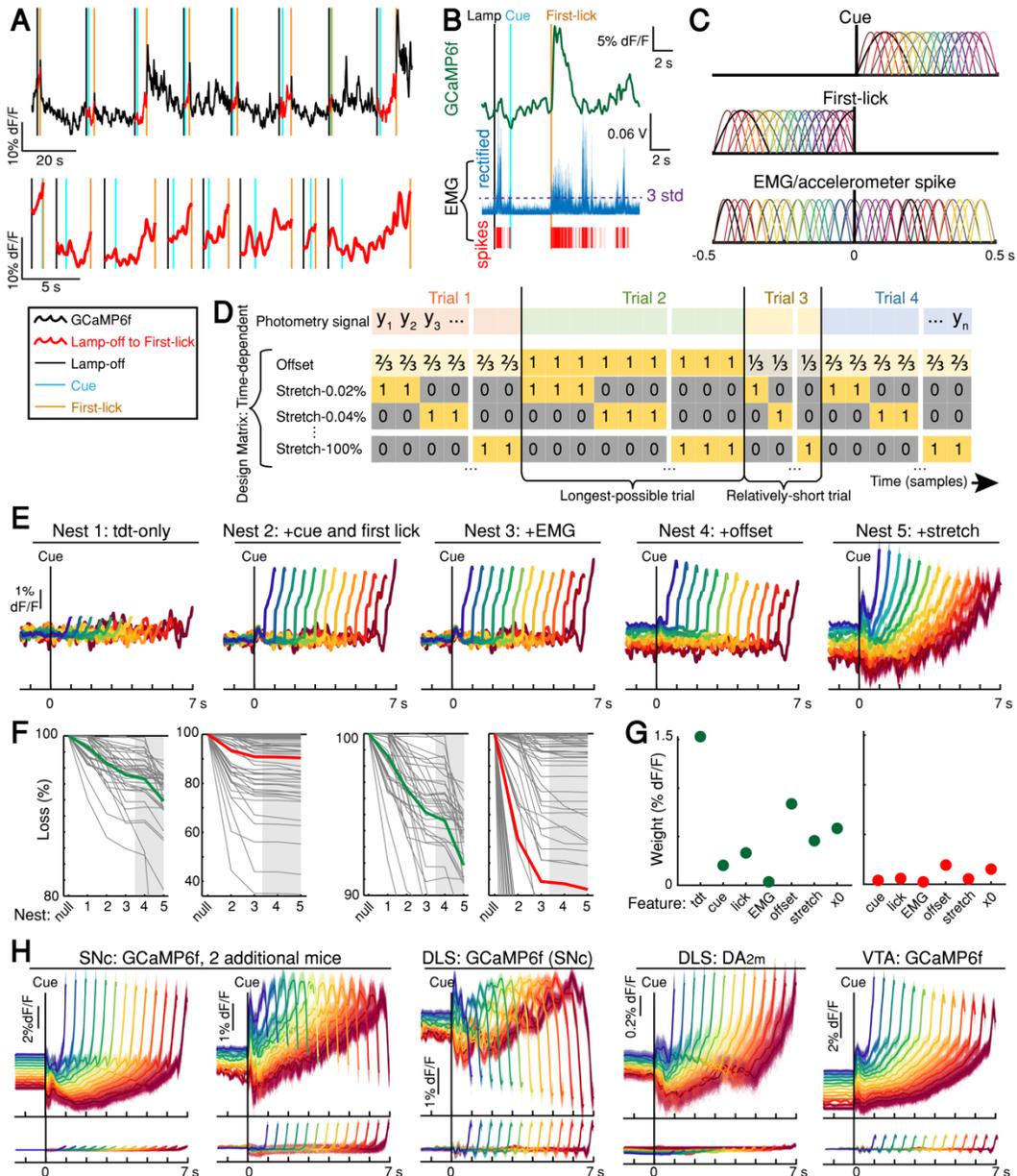
Figure 5—figure supplement 1. DAN signal encoding model parameterization and model selection. **(A)** Schematic of photometry timeseries fit by encoding model. The lamp-off to first-lick interval was excised from each trial in a session (top) and concatenated to produce the timeseries fit by the model (bottom). **(B)** EMG spikes derivation: thresholding rectified EMG at 3 standard deviations (example trial). **(C)** Optimized basis kernels to produce timing-independent features. **(D)** Schematic of Design Matrix for timing-dependent features. **(E)** GCaMP6f model fits by nest iteration for example session. Shading: model error simulated 300x. **(F)** Model loss by nest iteration. Green: mean loss for SNc GCaMP6f; red: mean loss for tdTomato (tdt); grey lines: individual sessions; grey shading: timing-dependent nests. Left: full-scale view of all datasets. Right: mean GCaMP6f and tdt loss compared on same scale. **(G)** Summary of feature weights across SNc GCaMP6f (left) and tdt (right) models. Coefficient weights were rectified, summed, and divided by the number of predictors per feature. 2x standard error bars (too small to see). All features were significant in both GCaMP6f and tdt models. **(H)** Top: examples of the full timing-dependent model (nest 5) from additional mice for all recorded dopaminergic signals. Bottom: tdt control channel fit. Model errors simulated 300x. Some mice show downward-going movement-related spikes at SNc cell bodies (second panel). All mice showed downward-going movement-related spikes from SNc terminals in DLS (middle panel). Source data: *Figure 5—source data*.
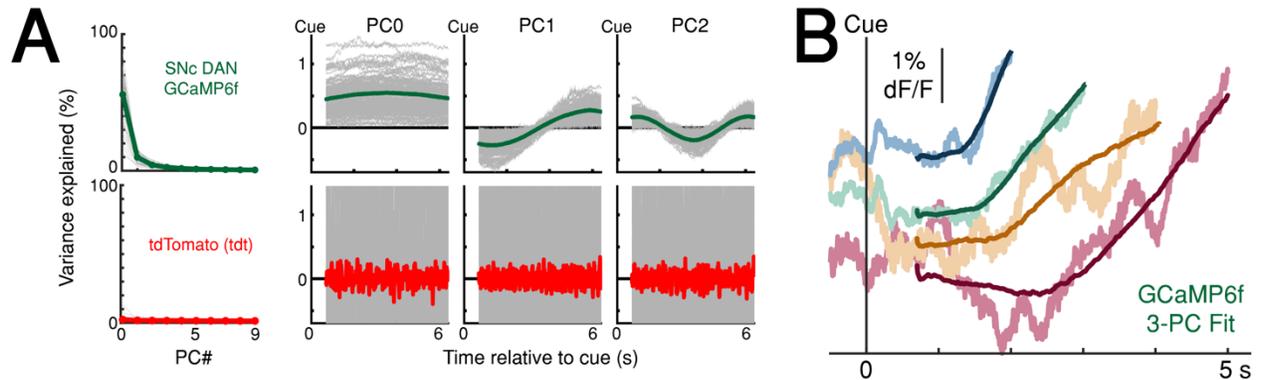
Figure 5—figure supplement 2. Principal component analysis (PCA) of the ramping interval (0.7 s up to first-lick relative to cue). **(A)** Left: Variance explained by first 10 principal components (PC). Right: first three principal components. Green line: mean PC, GCaMP6f recorded at SNc; Red line: mean PC, tdTomato (tdt) recorded at SNc and VTA; Grey lines: single-session data. X-axis shown for longest-possible interpolated trial duration; trials of shorter duration were interpolated to have the same number of samples for PCA. **(B)** Example session data simulated with first 3 PCs. Noisy traces: actual averaged GCaMP6f signals truncated at first-lick onset; Smooth traces: PC fits of the same trials. Source data: *Figure 5—source data*.
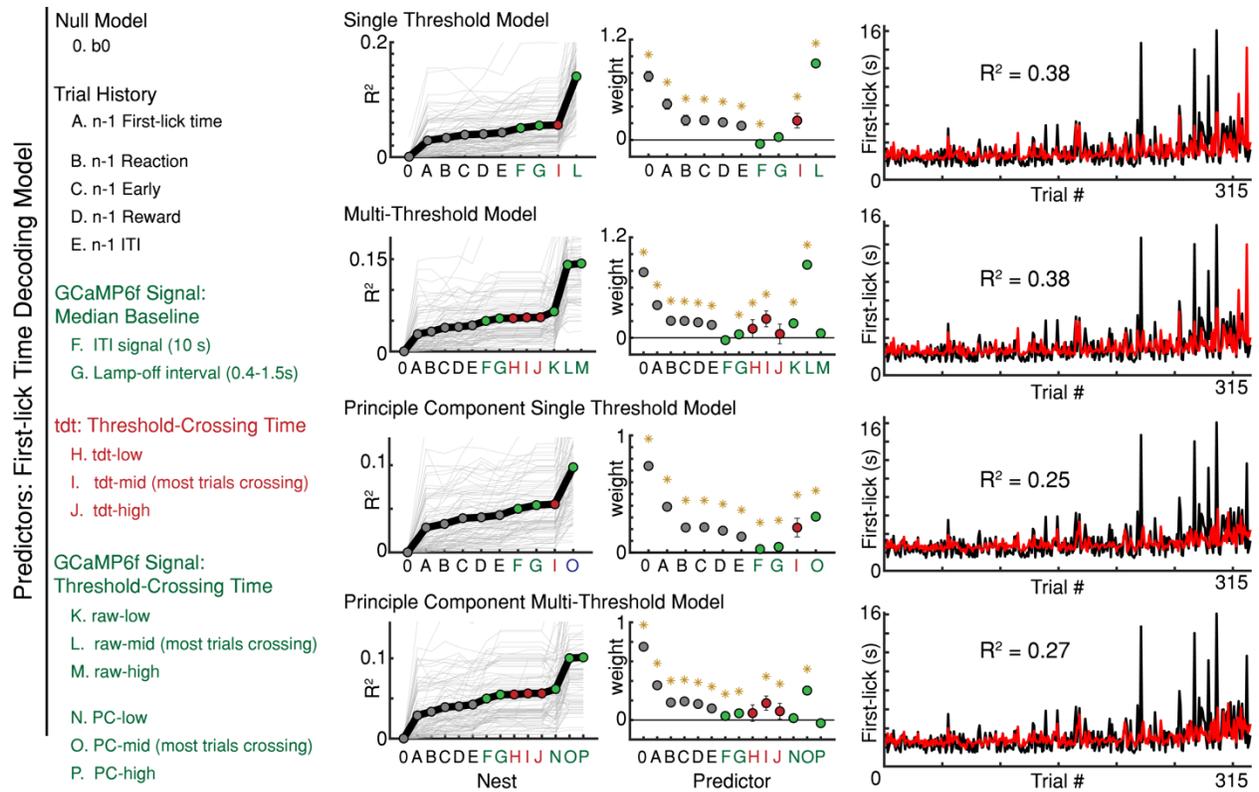
Figure 6—figure supplement 1. Variations of the first-lick time decoding model. *: p<0.05, error bars: 95% confidence intervals. GCaMP6f threshold crossing time dominated every version of the model; n-1th trial first-lick time was consistently the second-best predictor. Source data: *Figure 6—source data*.
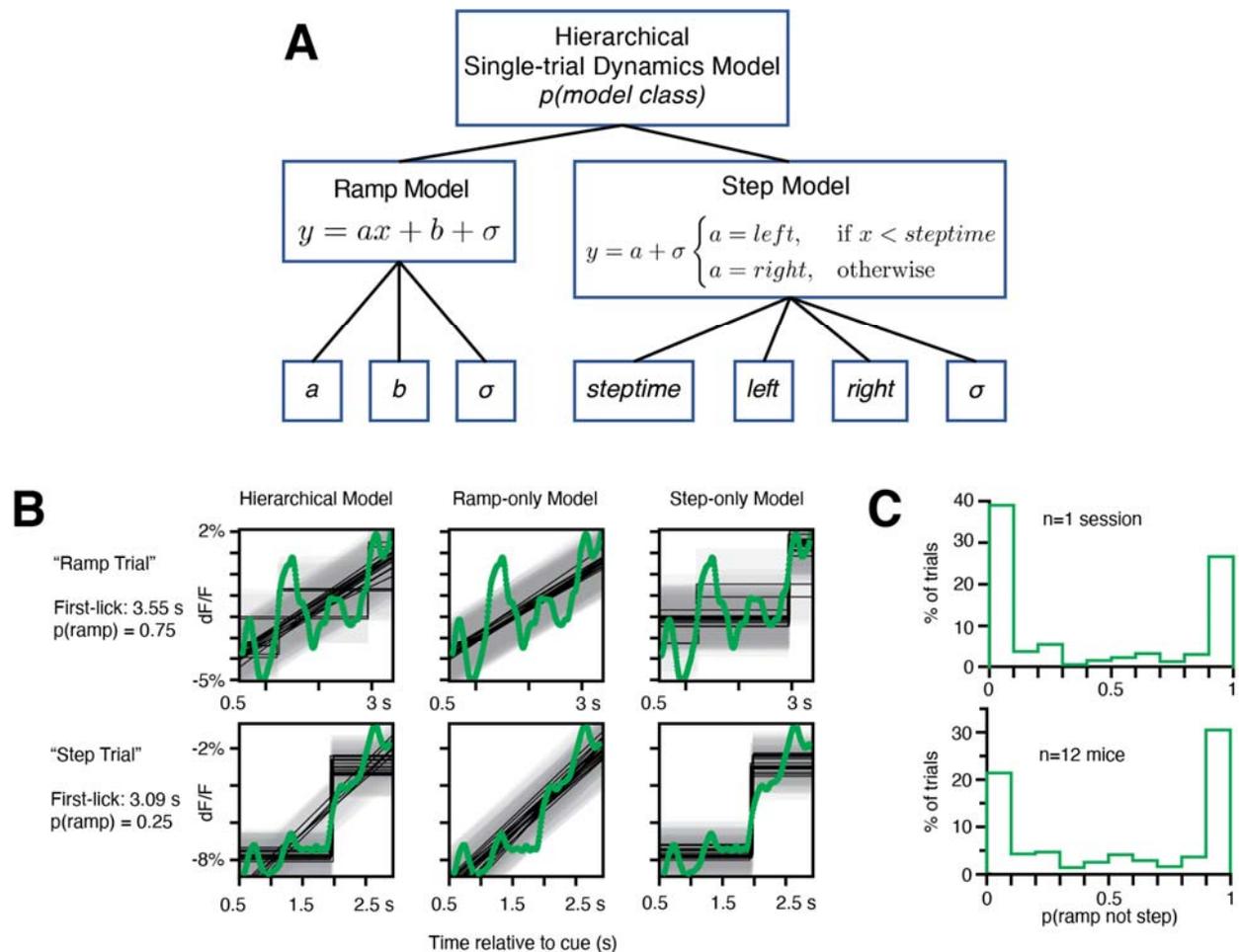
A

Hierarchical
Single-trial Dynamics Model
p(model class)

Ramp Model
$y = ax + b + \sigma$

Step Model
$y = a + \sigma \begin{cases} a = left, & \text{if } x < steptime \\ a = right, & \text{otherwise} \end{cases}$

a *b* σ *steptime* *left* *right* σ

B

Hierarchical Model   Ramp-only Model   Step-only Model

"Ramp Trial"

First-lick: 3.55 s
p(ramp) = 0.75

"Step Trial"

First-lick: 3.09 s
p(ramp) = 0.25

Time relative to cue (s)

C

n=1 session

n=12 mice

p(ramp not step)

Figure 6—figure supplement 2. Analysis of single-trial dynamics: Hierarchical Bayesian Ramp *vs.* Step Modeling. **(A)** Schematic (see *Methods: Hierarchical Bayesian Modeling of Single-trial Dynamics*). **(B)** Example fits from hierarchical model on 2 example single trials from the same epoch in a single session. Green: SNc GCaMP6f single-trial signal, light grey shading: noise band, dark grey lines: model fits. Note that the top trial is more frequently classified as a ramp, and the lower trial is more frequently classified as a step. However, both the ramp and step models return intuitive and reasonable fits to both single-trial signals. **(C)**. Probability of model class across all trials. X axis: 0 indicates all probabilistic fits for a given trial returned step-class models; 1 indicates all ramp-class models. Single sessions across mice showed considerable uncertainty in model classification. Source data: *Figure 6—source data*.
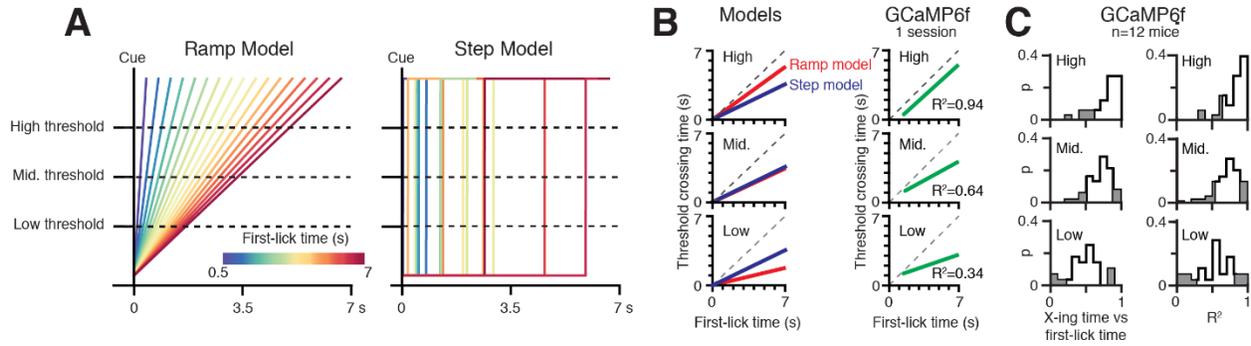
Figure 6—figure supplement 3. Geometric analysis of single-trial dynamics with Multiple Threshold Modeling. **(A)** Left: linear ramp model, Right: discrete step model. Step positions drawn from uniform distribution over the cue-to-first-lick interval. Low-, Mid- and High- level thresholds shown. **(B)** Threshold-crossing time *vs.* first-lick time ("X-ing time *vs.* first-lick time") for (from top to bottom) High-, Mid- and Low-level thresholds. Left: simulation predictions for ramp and step models. Right: X-ing time *vs.* first-lick time regression fit on single trials from 1 session (data from *Figure 6A*). The step model predicts X-ing time *vs.* first-lick time does not change across threshold levels, whereas ramp model predicts the slope of this relationship increases as threshold is raised. Single-trial GCaMP6f data exhibits increasing X-ing time *vs.* first-lick time slope with increasing threshold level, consistent with the ramp model but inconsistent with the step model. **(C)**. X-ing time *vs.* first-lick time across all mice. Left column: frequency of slope relationship across sessions, right column: variance explained. Source data: *Figure 6—source data*.
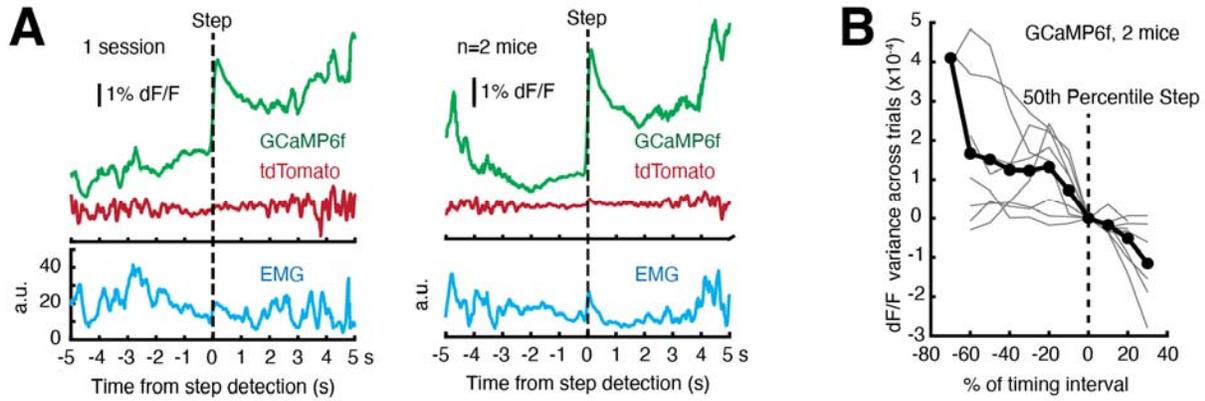
Figure 6—figure supplement 4. Assessing single-trial dynamics. **(A)** Single-trial signals aligned to discrete step position as found by Bayesian step model do not exhibit discrete step dynamics. To best estimate step times, the two animals with the highest GCaMP6f S:N were examined (Mouse B5 and B6). Left: 1 session, Right: average of signals from both mice. **(B)** Variance of GCaMP6f signals across trials. Step times were computed by Bayesian step model. An ideal step model predicts maximal variance at the 50th percentile step, but variance declined monotonically on average. Grey lines: single sessions; black line: average. For detailed explanation, see *Methods: Single-trial variance analysis for discrete step dynamics*. Source data: *Figure 6—source data*.
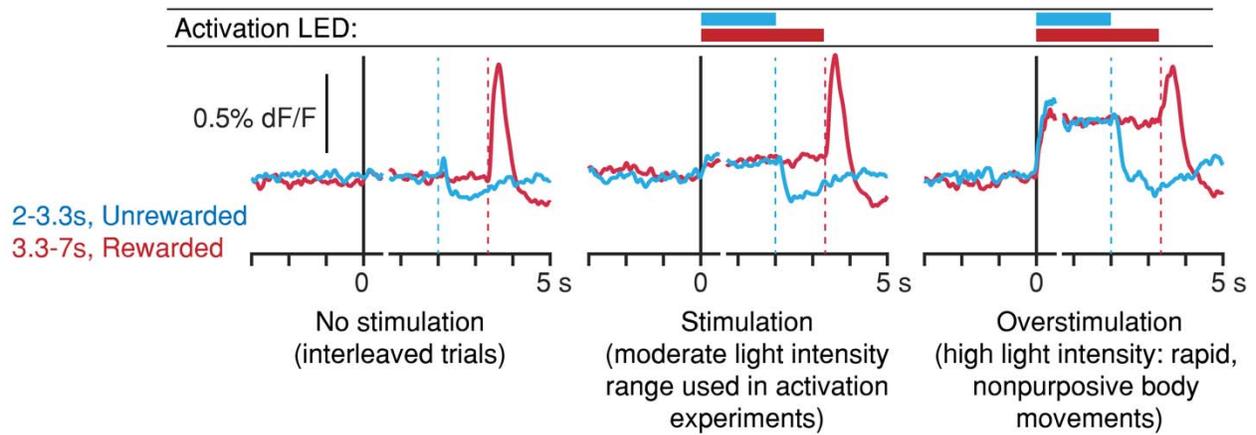
Figure 7—figure supplement 1. Variations on measurements of optogenetic effects. **(A)** Strategy for optogenetic targeting of DANs. **(B)** Comparison of four complementary metrics for addressing optogenetic effects. Left: unsigned Kolmogorov-Smirnov Distance (KS-D) analysis of differences in first-lick time distribution. Center: signed, bootstrapped comparison of difference in area under the cdf curves (dAUC). Right: mean and median bootstrapped difference in first-lick time. Source data: *Figure 7— source data*.

Figure 7—figure supplement 2. Light-power calibration for optogenetic activation of DANs. In preliminary experiments, DLS dopamine levels were monitored during the self-timed movement task, in which SNc DANs were activated randomly on 30% of interleaved trials. Dashed vertical lines: first-lick time. **Left**: interleaved, unstimulated trials (2 mice, 8 sessions). **Middle**: stimulated trials at the range of light levels used in the activation experiments show slightly elevated DLS dopamine signals compared to interleaved, unstimulated trials. First-lick timing was generally early-shifted in these sessions. **Right**: in a subset of preliminary calibration sessions, stimulation light levels were increased to the point where rapid, nonpurposive limb/trunk movements were observed throughout stimulation (1 mouse, 3 sessions). DLS dopamine signals show much higher, sustained increases throughout stimulation. Ongoing body movements disrupted task participation. Source data: *Figure 7—source data*.
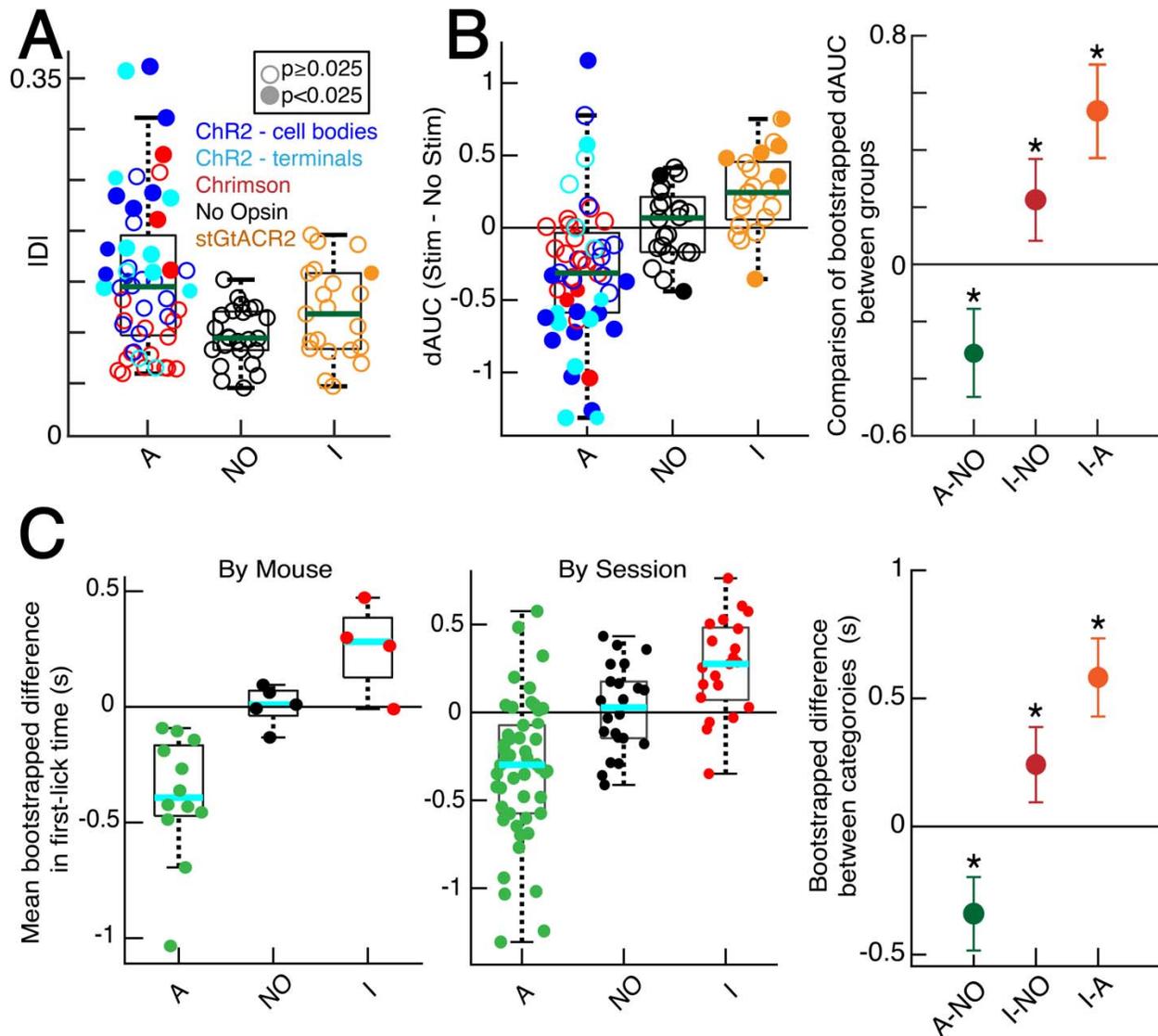
Figure 7—figure supplement 3. Quantification of optogenetic effects with additional metrics. **(A)** KS-D analysis: all sessions. "A": activation sessions; "NO": no opsin sessions; "I": inhibition sessions. Filled circles indicate significant difference between stimulated/unstimulated trials on single session ($p < 0.025$, 2-sided, 2-sample KS test). Standard box plot, line: median, box: upper/lower quartiles; whiskers: 1.5x IQR. **(B)** Left: bootstrapped dAUC Assay: all sessions, standard box plot as in *(A)*. Filled circles: significant difference on single session ($p < 0.025$, 2-sided bootstrapped dAUC test, see *Methods*). Right: comparison of dAUC in first-lick distributions across all sessions between groups. Error bars denote bootstrapped 95% confidence interval (*: $p < 0.05$). **(C)** Mean bootstrapped difference in first-lick time, stimulated-minus-unstimulated trials, standard box plot as in *(A)*. Left: single mice; Middle: single sessions. Right: Comparison of mean difference in first-lick time across all sessions. Error bars denote bootstrapped 95% confidence interval (*: $p < 0.05$). Source data: *Figure 7—source data*.
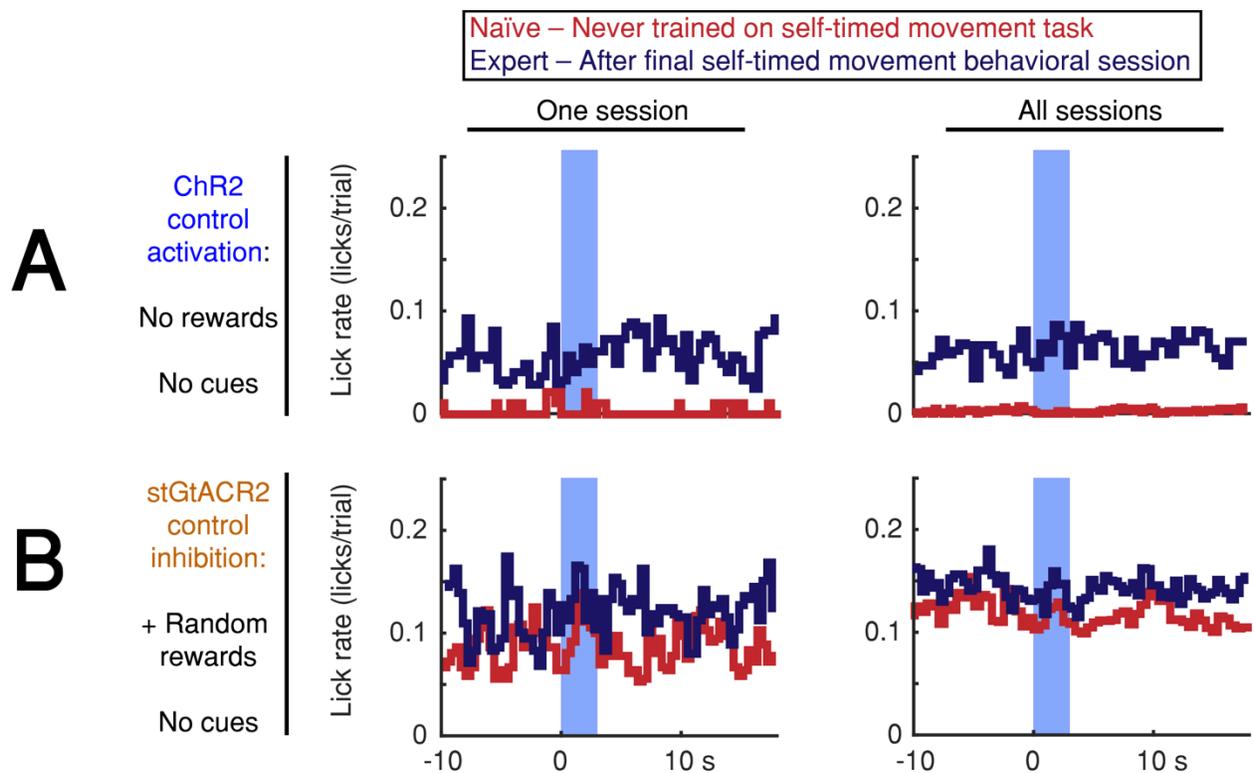
Figure 7—figure supplement 4. Optogenetic DAN stimulation does not cause or prevent licking. **(A,B)** Stimulation-aligned lick-rate during control sessions. Animals were tested in 1-3 control sessions both before exposure to the self-timed movement task (red) and in 1-2 control sessions after the end of behavioral training (navy). Blue bar indicates stimulation period (3 s). Left: one session, Right: all sessions. **(A)** Activation control sessions (no cues or rewards). Animals were head-fixed on the behavioral platform and stimulated randomly at the same pace as the standard 3.3 s self-timed movement task. Activation did not elicit immediate licking in any session. **(B)** Inhibition-control sessions (no cues, + random rewards). Animals were head-fixed on the behavioral platform while receiving juice rewards at random times. Inhibition did not prevent licking in any session. Source data: *Figure 7—source data*.
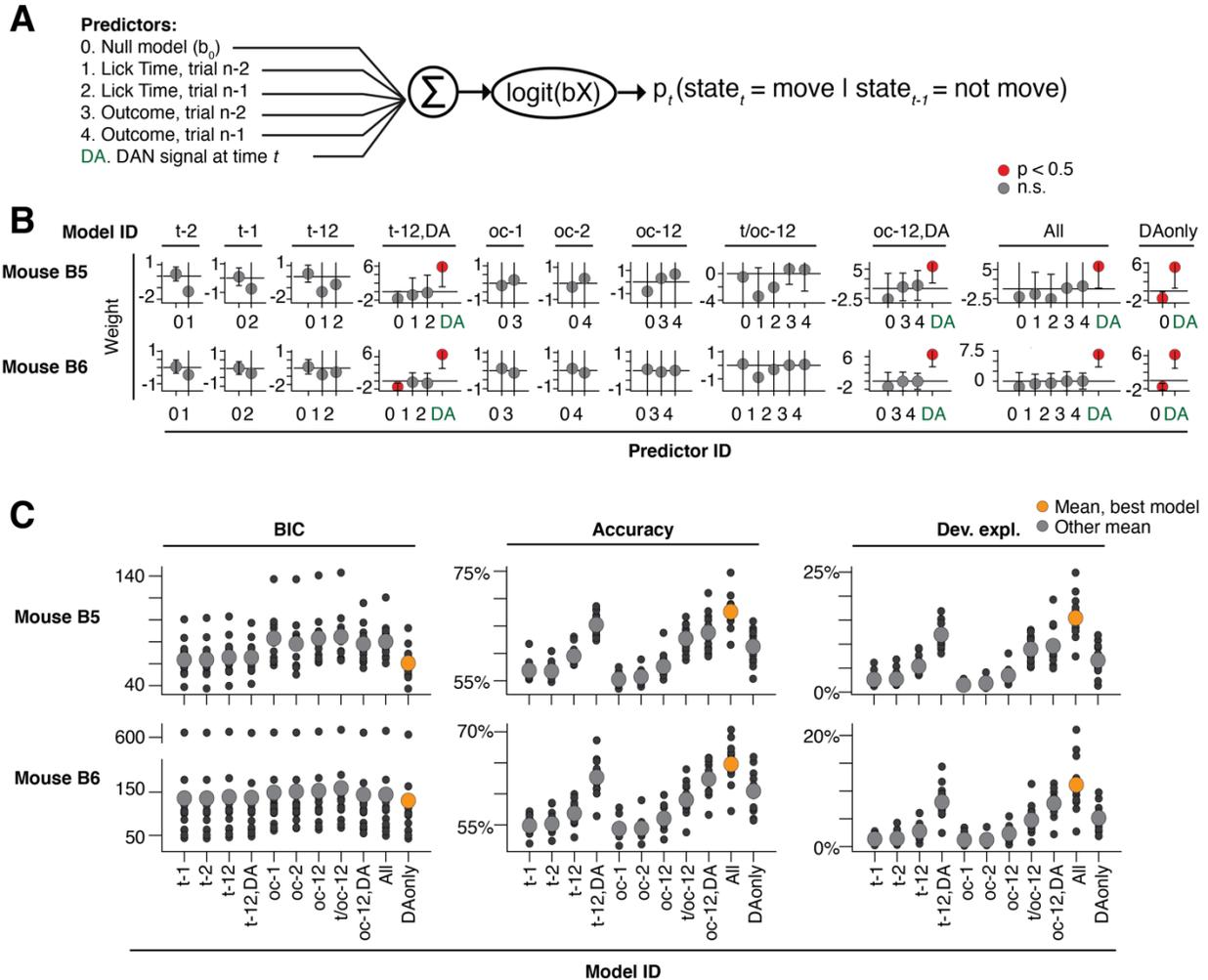
Figure 8—figure supplement 1. Probabilistic movement time decoding model: model selection. **(A)**
Model schematic. To assess previous trial history on the same footing as dopaminergic signals, time t
during model selection was limited to a 500 ms "time-slice," with each time-slice fit separately by the
model. Dopaminergic signals were averaged within each time-slice, such that each trial provided one and
only one dopaminergic measurement, one set of trial history terms, and one movement state per time slice
(see *Methods: Single-trial probabilistic movement state decoding model, model selection*). **(B)** Model fit
weights. Model ID: corresponds to the predictors included from the schematic. x-axis labels: the predictor
ID from the schematic. Predictor weights averaged across time-slices. **(C)** Model selection criteria. The
model omitting the previous trial history predictors (predictors #1-4) was consistently the best model as
selected by BIC, AIC and AICc (results similar across metrics, BIC shown alone for clarity). Source data:
Figure 8—source data.
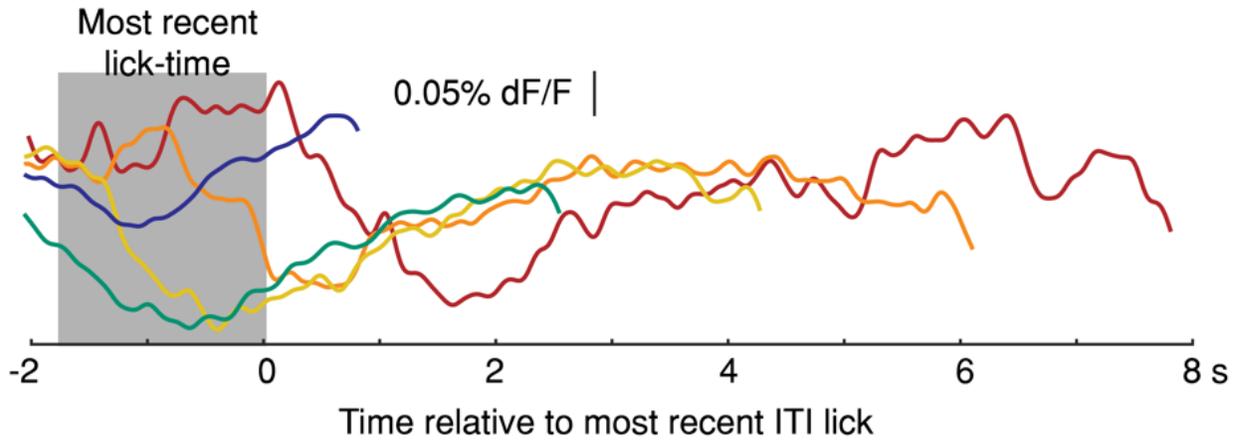
Figure 8—figure supplement 2. Average Intertrial Interval (ITI) GCaMP6f signals aligned to most recent previous lick-time. Signals plotted up to onset of next spontaneous, self-initiated lick during the ITI. (1 mouse, 5 sessions, truncated 150 ms before lick). Source data: *Figure 8—source data*.

REFERENCES

Albin, R.L., Young, A.B., and Penney, J.B. (1989). The functional anatomy of basal ganglia disorders. Trends Neurosci *12*, 366-375.

Anger, D. (1956). The dependence of interresponse times upon the relative reinforcement of different interresponse times. J Exp Psychol *52*, 145-161.

Backman, C.M., Malik, N., Zhang, Y., Shan, L., Grinberg, A., Hoffer, B.J., Westphal, H., and Tomac, A.C. (2006). Characterization of a mouse strain expressing Cre recombinase from the 3' untranslated region of the dopamine transporter locus. Genesis *44*, 383-390.

Barter, J.W., Li, S., Lu, D., Bartholomew, R.A., Rossi, M.A., Shoemaker, C.T., Salas-Meza, D., Gaidis, E., and Yin, H.H. (2015). Beyond reward prediction errors: the role of dopamine in movement kinematics. Front Integr Neurosci *9*, 39.

Barthel, C., Nonnekes, J., van Helvert, M., Haan, R., Janssen, A., Delval, A., Weerdesteyn, V., Debu, B., van Wezel, R., Bloem, B.R.*, et al.* (2018). The laser shoes: A new ambulatory device to alleviate freezing of gait in Parkinson disease. Neurology *90*, e164-e171.

Bartholomew, R.A., Li, H., Gaidis, E.J., Stackmann, M., Shoemaker, C.T., Rossi, M.A., and Yin, H.H. (2016). Striatonigral control of movement velocity in mice. Eur J Neurosci *43*, 1097-1110.

Bloxham, C.A., Mindel, T.A., and Frith, C.D. (1984). Initiation and execution of predictable and unpredictable movements in Parkinson's disease. Brain *107 (Pt 2)*, 371-384.

Chandrasekaran, C., Soldado-Magraner, J., Peixoto, D., Newsome, W.T., Shenoy, K.V., and Sahani, M. (2018). Brittleness in model selection analysis of single neuron firing rates. bioRxiv, 430710.

Coddington, L.T., and Dudman, J.T. (2018). The timing of action determines reward prediction signals in identified midbrain dopamine neurons. Nat Neurosci *21*, 1563-1573.

Coddington, L.T., and Dudman, J.T. (2019). Learning from Action: Reconsidering Movement Signaling in Midbrain Dopamine Neuron Activity. Neuron *104*, 63-77.

Coddington, L.T., and Dudman, J.T. (2021). In Vivo Optogenetics with Stimulus Calibration. Methods Mol Biol *2188*, 273-283.

Cusumano-Towner, M.F., and Mansinghka, V.K. (2018). Using probabilistic programs as proposals. ArXiv *abs/1801.03612*.

Cusumano-Towner, M.F., Saad, F.A., Lew, A.K., and Mansinghka, V.K. (2019). Gen: A General-Purpose Probabilistic Programming System with Programmable Inference. Proceedings of the 40th Acm Sigplan Conference on Programming Language Design and Implementation (Pldi '19), 221-236.

da Silva, J.A., Tecuapetla, F., Paixão, V., and Costa, R.M. (2018). Dopamine neuron activity before action initiation gates and invigorates future movements. Nature *554*, 244-248.

1798   Deecke, L. (1996). Planning, preparation, execution, and imagery of volitional action. Brain Res Cogn
1799   Brain Res *3*, 59-64.

1800   DeLong, M.R. (1990). Primate models of movement disorders of basal ganglia origin. Trends Neurosci *13*,
1801   281-285.

1802   Dews, P.B., and Morse, W.H. (1958). Some observations on an operant in human subjects and its
1803   modification by dextro amphetamine. J Exp Anal Behav *1*, 359-364.

1804   Dodson, P.D., Dreyer, J.K., Jennings, K.A., Syed, E.C., Wade-Martins, R., Cragg, S.J., Bolam, J.P., and
1805   Magill, P.J. (2016). Representation of spontaneous movement by dopaminergic neurons is cell-type
1806   selective and disrupted in parkinsonism. Proc Natl Acad Sci U S A *113*, E2180-2188.

1807   Dudman, J.T., and Krakauer, J.W. (2016). The basal ganglia: from motor commands to the control of
1808   vigor. Curr Opin Neurobiol *37*, 158-166.

1809   Eckard, M.L., and Kyonka, E.G.E. (2018). Differential reinforcement of low rates differentially decreased
1810   timing precision. Behav Processes *151*, 111-118.

1811   Emmons, E.B., De Corte, B.J., Kim, Y., Parker, K.L., Matell, M.S., and Narayanan, N.S. (2017). Rodent
1812   Medial Frontal Control of Temporal Processing in the Dorsomedial Striatum. J Neurosci *37*, 8718-8733.

1813   Engelhard, B., Finkelstein, J., Cox, J., Fleming, W., Jang, H.J., Ornelas, S., Koay, S.A., Thiberge, S.Y.,
1814   Daw, N.D., Tank, D.W.*, et al.* (2019). Specialized coding of sensory, motor and cognitive variables in VTA
1815   dopamine neurons. Nature *570*, 509-513.

1816   Fahn, S. (2011). Classification of movement disorders. Mov Disord *26*, 947-957.

1817   Fiorillo, C.D. (2011). Transient activation of midbrain dopamine neurons by reward risk. Neuroscience *197*,
1818   162-171.

1819   Fiorillo, C.D., Tobler, P.N., and Schultz, W. (2003). Discrete coding of reward probability and uncertainty
1820   by dopamine neurons. Science *299*, 1898-1902.

1821   Freeze, B.S., Kravitz, A.V., Hammack, N., Berke, J.D., and Kreitzer, A.C. (2013). Control of basal ganglia
1822   output by direct and indirect pathway projection neurons. J Neurosci *33*, 18531-18539.

1823   Gallistel, C.R., and Gibbon, J. (2000). Time, rate, and conditioning. Psychol Rev *107*, 289-344.

1824   Gelman, A., and Hill, J. (2006). Data analysis using regression and multilevel/hierarchical models
1825   (Cambridge ; New York: Cambridge University Press).

1826   Grillner, S., and Robertson, B. (2016). The Basal Ganglia Over 500 Million Years. Curr Biol *26*, R1088-
1827   R1100.

1828 Guo, Z.V., Hires, S.A., Li, N., O'Connor, D.H., Komiyama, T., Ophir, E., Huber, D., Bonardi, C., Morandell,
1829 K., Gutnisky, D., *et al.* (2014). Procedures for behavioral experiments in head-fixed mice. PLoS One *9*,
1830 e88678.

1831 Guru, A., Seo, C., Post, R.J., Kullakanda, D.S., Schaffer, J.A., and Warden, M.R. (2020). Ramping activity
1832 in midbrain dopamine neurons signifies the use of a cognitive map. bioRxiv.

1833 Hallett, M. (2007). Volitional control of movement: the physiology of free will. Clin Neurophysiol *118*, 1179-
1834 1192.

1835 Hallett, M., and Khoshbin, S. (1980). A physiological mechanism of bradykinesia. Brain *103*, 301-314.

1836 Hamid, A.A., Pettibone, J.R., Mabrouk, O.S., Hetrick, V.L., Schmidt, R., Vander Weele, C.M., Kennedy,
1837 R.T., Aragona, B.J., and Berke, J.D. (2016). Mesolimbic dopamine signals the value of work. Nat
1838 Neurosci *19*, 117-126.

1839 Hamilos, A.E., and Assad, J.A. (2020). Application of a unifying reward-prediction error (RPE)-based
1840 framework to explain underlying dynamic dopaminergic activity in timing tasks. bioRxiv,
1841 2020.2006.2003.128272.

1842 Hart, A.S., Clark, J.J., and Phillips, P.E.M. (2015). Dynamic shaping of dopamine signals during
1843 probabilistic Pavlovian conditioning. Neurobiol Learn Mem *117*, 84-92.

1844 Helassa, N., Podor, B., Fine, A., and Torok, K. (2016). Design and mechanistic insight into ultrafast
1845 calcium indicators for monitoring intracellular calcium dynamics. Sci Rep *6*, 38276.

1846 Howard, C.D., Li, H., Geddes, C.E., and Jin, X. (2017). Dynamic Nigrostriatal Dopamine Biases Action
1847 Selection. Neuron *93*, 1436-1450 e1438.

1848 Howe, M.W., and Dombeck, D.A. (2016). Rapid signalling in distinct dopaminergic axons during
1849 locomotion and reward. Nature *535*, 505-510.

1850 Howe, M.W., Tierney, P.L., Sandberg, S.G., Phillips, P.E., and Graybiel, A.M. (2013). Prolonged
1851 dopamine signalling in striatum signals proximity and value of distant rewards. Nature *500*, 575-579.

1852 Hughes, L.E., Altena, E., Barker, R.A., and Rowe, J.B. (2013). Perseveration and choice in Parkinson's
1853 disease: the impact of progressive frontostriatal dysfunction on action decisions. Cereb Cortex *23*, 1572-
1854 1581.

1855 Jaldow, E.J., Oakley, D.A., and Davey, G.C. (1990). Performance on two fixed-interval schedules in the
1856 absence of neocortex in rats. Behav Neurosci *104*, 763-777.

1857 Kim, H.R., Malik, A.N., Mikhael, J.G., Bech, P., Tsutsui-Kimura, I., Sun, F., Zhang, Y., Li, Y., Watabe-
1858 Uchida, M., Gershman, S.J., *et al.* (2019). A unified framework for dopamine signals across timescales.
1859 bioRxiv, 803437.

1860 Kirshenbaum, A.P., Brown, S.J., Hughes, D.M., and Doughty, A.H. (2008). Differential-reinforcement-of-
1861 low-rate-schedule performance and nicotine administration: a systematic investigation of dose, dose-
1862 regimen, and schedule requirement. Behav Pharmacol *19*, 683-697.

1863 Latimer, K.W., Yates, J.L., Meister, M.L., Huk, A.C., and Pillow, J.W. (2015). NEURONAL MODELING.
1864 Single-trial spike trains in parietal cortex reveal discrete steps during decision-making. Science *349*, 184-
1865 187.

1866 Latimer, K.W., Yates, J.L., Meister, M.L., Huk, A.C., and Pillow, J.W. (2016). Response to Comment on
1867 "Single-trial spike trains in parietal cortex reveal discrete steps during decision-making". Science *351*,
1868 1406.

1869 Lee, I.H., and Assad, J.A. (2003). Putaminal activity for simple reactions or self-timed movements. J
1870 Neurophysiol *89*, 2528-2537.

1871 Lee, K., Claar, L.D., Hachisuka, A., Bakhurin, K.I., Nguyen, J., Trott, J.M., Gill, J.L., and Masmanidis, S.C.
1872 (2020). Temporally restricted dopaminergic control of reward-conditioned movements. Nat Neurosci *23*,
1873 209-216.

1874 Libet, B., Gleason, C.A., Wright, E.W., and Pearl, D.K. (1983). Time of conscious intention to act in
1875 relation to onset of cerebral activity (readiness-potential). The unconscious initiation of a freely voluntary
1876 act. Brain *106 (Pt 3)*, 623-642.

1877 Lustig, C., and Meck, W.H. (2005). Chronic treatment with haloperidol induces deficits in working memory
1878 and feedback effects of interval timing. Brain Cogn *58*, 9-16.

1879 Maes, E.J.P., Sharpe, M.J., Usypchuk, A.A., Lozzi, M., Chang, C.Y., Gardner, M.P.H., Schoenbaum, G.,
1880 and Iordanova, M.D. (2020). Causal evidence supporting the proposal that dopamine transients function
1881 as temporal difference prediction errors. Nat Neurosci *23*, 176-178.

1882 Maimon, G., and Assad, J.A. (2006). A cognitive signal for the proactive timing of action in macaque LIP.
1883 Nat Neurosci *9*, 948-955.

1884 Malapani, C., Rakitin, B., Levy, R., Meck, W.H., Deweer, B., Dubois, B., and Gibbon, J. (1998). Coupled
1885 temporal memories in Parkinson's disease: a dopamine-related dysfunction. J Cogn Neurosci *10*, 316-
1886 331.

1887 Matsumoto, M., and Hikosaka, O. (2009). Two types of dopamine neuron distinctly convey positive and
1888 negative motivational signals. Nature *459*, 837-841.

1889 Mazzoni, P., Hristova, A., and Krakauer, J.W. (2007). Why don't we move faster? Parkinson's disease,
1890 movement vigor, and implicit motivation. J Neurosci *27*, 7105-7116.

1891 Meck, W.H. (1986). Affinity for the dopamine D2 receptor predicts neuroleptic potency in decreasing the
1892 speed of an internal clock. Pharmacol Biochem Behav *25*, 1185-1189.

1893 Meck, W.H. (2006). Neuroanatomical localization of an internal clock: a functional link between
1894 mesolimbic, nigrostriatal, and mesocortical dopaminergic systems. Brain Res *1109*, 93-107.

1895   Mello, G.B., Soares, S., and Paton, J.J. (2015). A scalable population code for time in the striatum. Curr
1896   Biol *25*, 1113-1122.

1897   Menegas, W., Bergan, J.F., Ogawa, S.K., Isogai, Y., Umadevi Venkataraju, K., Osten, P., Uchida, N., and
1898   Watabe-Uchida, M. (2015). Dopamine neurons projecting to the posterior striatum form an anatomically
1899   distinct subclass. Elife *4*, e10032.

1900   Merchant, H., Harrington, D.L., and Meck, W.H. (2013). Neural basis of the perception and estimation of
1901   time. Annu Rev Neurosci *36*, 313-336.

1902   Mikhael, J.G., and Gershman, S.J. (2019). Adapting the flow of time with dopamine. J Neurophysiol *121*,
1903   1748-1760.

1904   Mikhael, J.G., Kim, H.R., Uchida, N., and Gershman, S.J. (2019). Ramping and State Uncertainty in the
1905   Dopamine Signal. bioRxiv, 805366.

1906   Mita, A., Mushiake, H., Shima, K., Matsuzaka, Y., and Tanji, J. (2009). Interval time coding by neurons in
1907   the presupplementary and supplementary motor areas. Nat Neurosci *12*, 502-507.

1908   Mohebi, A., Pettibone, J.R., Hamid, A.A., Wong, J.T., Vinson, L.T., Patriarchi, T., Tian, L., Kennedy, R.T.,
1909   and Berke, J.D. (2019). Dissociable dopamine dynamics for learning and motivation. Nature *570*, 65-70.

1910   Morrens, J., Aydin, C., Janse van Rensburg, A., Esquivelzeta Rabell, J., and Haesler, S. (2020). Cue-
1911   Evoked Dopamine Promotes Conditioned Responding during Learning. Neuron *106*, 142-153 e147.

1912   Pan, W.X., Coddington, L.T., and Dudman, J.T. (2021). Dissociable contributions of phasic dopamine
1913   activity to reward and prediction. Cell Rep *36*, 109684.

1914   Panigrahi, B., Martin, K.A., Li, Y., Graves, A.R., Vollmer, A., Olson, L., Mensh, B.D., Karpova, A.Y., and
1915   Dudman, J.T. (2015). Dopamine Is Required for the Neural Representation and Control of Movement
1916   Vigor. Cell *162*, 1418-1430.

1917   Park, I.M., Meister, M.L., Huk, A.C., and Pillow, J.W. (2014). Encoding and decoding in parietal cortex
1918   during sensorimotor decision-making. Nat Neurosci *17*, 1395-1403.

1919   Parker, N.F., Cameron, C.M., Taliaferro, J.P., Lee, J., Choi, J.Y., Davidson, T.J., Daw, N.D., and Witten,
1920   I.B. (2016). Reward and choice encoding in terminals of midbrain dopamine neurons depends on striatal
1921   target. Nat Neurosci *19*, 845-854.

1922   Patriarchi, T., Cho, J.R., Merten, K., Howe, M.W., Marley, A., Xiong, W.H., Folk, R.W., Broussard, G.J.,
1923   Liang, R., Jang, M.J.*, et al.* (2018). Ultrafast neuronal imaging of dopamine dynamics with designed
1924   genetically encoded sensors. Science *360*.

1925   Phillips, P.E., Stuber, G.D., Heien, M.L., Wightman, R.M., and Carelli, R.M. (2003). Subsecond dopamine
1926   release promotes cocaine seeking. Nature *422*, 614-618.

1927   Rakitin, B.C., Gibbon, J., Penney, T.B., Malapani, C., Hinton, S.C., and Meck, W.H. (1998). Scalar
1928   expectancy theory and peak-interval timing in humans. J Exp Psychol Anim Behav Process *24*, 15-33.

1929 Remington, E.D., Narain, D., Hosseini, E.A., and Jazayeri, M. (2018). Flexible Sensorimotor
1930 Computations through Rapid Reconfiguration of Cortical Dynamics. Neuron *98*, 1005-1019 e1005.

1931 Romo, R., Scarnati, E., and Schultz, W. (1992). Role of primate basal ganglia and frontal cortex in the
1932 internal generation of movements. II. Movement-related activity in the anterior striatum. Exp Brain Res *91*,
1933 385-395.

1934 Runyan, C.A., Piasini, E., Panzeri, S., and Harvey, C.D. (2017). Distinct timescales of population coding
1935 across cortex. Nature *548*, 92-96.

1936 Saunders, B.T., Richard, J.M., Margolis, E.B., and Janak, P.H. (2018). Dopamine neurons create
1937 Pavlovian conditioned stimuli with circuit-defined motivational properties. Nat Neurosci *21*, 1072-1083.

1938 Schultz, W., Dayan, P., and Montague, P.R. (1997). A neural substrate of prediction and reward. Science
1939 *275*, 1593-1599.

1940 Schuster, C.R., and Zimmerman, J. (1961). Timing behavior during prolonged treatment with dl-
1941 amphetamine. J Exp Anal Behav *4*, 327-330.

1942 Shadlen, M.N., Kiani, R., Newsome, W.T., Gold, J.I., Wolpert, D.M., Zylberberg, A., Ditterich, J., de
1943 Lafuente, V., Yang, T., and Roitman, J. (2016). Comment on "Single-trial spike trains in parietal cortex
1944 reveal discrete steps during decision-making". Science *351*, 1406.

1945 Shenoy, K.V., Sahani, M., and Churchland, M.M. (2013). Cortical control of arm movements: a dynamical
1946 systems perspective. Annu Rev Neurosci *36*, 337-359.

1947 Sippy, T., Lapray, D., Crochet, S., and Petersen, C.C. (2015). Cell-Type-Specific Sensorimotor
1948 Processing in Striatal Projection Neurons during Goal-Directed Behavior. Neuron *88*, 298-305.

1949 Soares, S., Atallah, B.V., and Paton, J.J. (2016). Midbrain dopamine neurons control judgment of time.
1950 Science *354*, 1273-1277.

1951 Sohn, H., Narain, D., Meirhaeghe, N., and Jazayeri, M. (2019). Bayesian Computation through Cortical
1952 Latent Dynamics. Neuron *103*, 934-947 e935.

1953 Starkweather, C.K., Babayan, B.M., Uchida, N., and Gershman, S.J. (2017). Dopamine reward prediction
1954 errors reflect hidden-state inference across time. Nat Neurosci *20*, 581-589.

1955 Sun, F., Zhou, J., Dai, B., Qian, T., Zeng, J., Li, X., Zhuo, Y., Zhang, Y., Tan, K., Feng, J.*, et al.* (2020).
1956 New and improved GRAB fluorescent sensors for monitoring dopaminergic activity *in vivo*.
1957 bioRxiv, 2020.2003.2028.013722.

1958 Tian, J., Huang, R., Cohen, J.Y., Osakada, F., Kobak, D., Machens, C.K., Callaway, E.M., Uchida, N.,
1959 and Watabe-Uchida, M. (2016). Distributed and Mixed Information in Monosynaptic Inputs to Dopamine
1960 Neurons. Neuron *91*, 1374-1389.

1961 Tian, J., and Uchida, N. (2015). Habenula Lesions Reveal that Multiple Mechanisms Underlie Dopamine
1962 Prediction Errors. Neuron *87*, 1304-1316.

1963   Tobler, P.N., Fiorillo, C.D., and Schultz, W. (2005). Adaptive coding of reward value by dopamine neurons.
1964   Science *307*, 1642-1645.

1965   Turner, R.S., and Desmurget, M. (2010). Basal ganglia contributions to motor control: a vigorous tutor.
1966   Curr Opin Neurobiol *20*, 704-716.

1967   Wang, D.V., and Tsien, J.Z. (2011). Conjunctive processing of locomotor signals by the ventral tegmental
1968   area neuronal population. PLoS One *6*, e16528.

1969   Wang, J., Narain, D., Hosseini, E.A., and Jazayeri, M. (2018). Flexible timing by temporal scaling of
1970   cortical responses. Nat Neurosci *21*, 102-110.

1971   Xu, M., Zhang, S.Y., Dan, Y., and Poo, M.M. (2014). Representation of interval timing by temporally
1972   scalable firing patterns in rat prefrontal cortex. Proc Natl Acad Sci U S A *111*, 480-485.

1973   Yttri, E.A., and Dudman, J.T. (2016). Opponent and bidirectional control of movement velocity in the basal
1974   ganglia. Nature *533*, 402-406.

1975   Zoltowski, D.M., Latimer, K.W., Yates, J.L., Huk, A.C., and Pillow, J.W. (2019). Discrete Stepping and
1976   Nonlinear Ramping Dynamics Underlie Spiking Responses of LIP Neurons during Decision-Making.
1977   Neuron *102*, 1249-1258 e1210.

1978   Zylberberg, A., and Shadlen, M.N. (2016). Cause for pause before leaping to conclusions about stepping.
1979   bioRxiv, 085886.
1980