

25 **SUMMARY**

26 The ventral pallidum (VP) is critical for invigorating reward seeking and is also involved in
27 punishment avoidance, but how it contributes to such opposing behavioural actions remains
28 unclear. Here we show that GABAergic and glutamatergic VP neurons selectively control
29 behaviour in opposing motivational contexts. *In vivo* recording combined with optogenetics in
30 mice revealed that these two populations oppositely encode positive and negative motivational
31 value, are differentially modulated by animal's internal state and determine the behavioural
32 response during motivational conflict. Furthermore, GABAergic VP neurons are essential for
33 movements towards reward in a positive motivational context, but suppress movements in an
34 aversive context. In contrast, glutamatergic VP neurons are essential for movements to avoid a
35 threat but suppress movements in an appetitive context. Our results indicate that GABAergic and
36 glutamatergic VP neurons encode the drive for approach and avoidance, respectively, with the
37 balance between their activities determining the type of motivational behaviour.

38

39 **INTRODUCTION**

40 The decision to approach or avoid depends on the situation in the environment and the internal
41 state of the animal. For example, thirst may encourage animals to seek a water source, but a sated
42 animal may not find this goal worth the energy expenditure or risk (Ydenberg, 1986). Equally, an
43 extremely thirsty animal may approach a water source despite the known risk of predators.

44

45 A key region involved in goal-directed motivation is the ventral pallidum (VP). The VP is the
46 major output structure of the ventral basal ganglia (Heimer et al., 1997). It receives projections
47 from areas including the nucleus accumbens (NAc), prefrontal cortex and basolateral amygdala,

48 and transmits information to multiple brain regions involved in motor control and motivation, such
49 as the ventral tegmental area (VTA), lateral habenula (LHb), mediodorsal thalamus and
50 pedunclopontine tegmental nucleus (Haber and Knutson, 2010). This connectivity places the VP
51 in an ideal location to transform information about the expected value of stimuli into motivation
52 (Mogenson et al., 1980). In particular, the VP sends a major projection to the LHb (Haber and
53 Knutson, 2010; Humphries and Prescott, 2010), a structure that has been shown to regulate both
54 positive and negative reinforcement learning (Lammel et al., 2012; Lecca et al., 2017; Stamatakis
55 et al., 2013; Stamatakis and Stuber, 2012; Stephenson-Jones et al., 2016; Stopper and Floresco,
56 2014; Tian and Uchida, 2015). Thus, some of the behavioral roles of the VP neurons could be
57 mediated by the VP→LHb pathway.

58

59 A large body of work, comprehensively reviewed by others (Humphries and Prescott, 2010; Root
60 et al., 2015; Smith et al., 2009; Stephenson-Jones, 2019), has identified the VP as a crucial driver
61 of reward-seeking behaviour. For example, the VP is important for the normal hedonic reactions
62 to sucrose (Farrar et al., 2008), and lesions to the VP decrease an animal's willingness to work for
63 reward (Farrar et al., 2008; Richard et al., 2016). Conversely, rats will work to electrically self-
64 stimulate their VP (Panagis et al., 1995; Panagis et al., 1997), and pharmacological activation and
65 disinhibition of the VP can both trigger feeding in sated animals (Stratford et al., 1999).

66

67 The VP is also implicated in avoidance behaviours (Stephenson-Jones, 2019; Wulff et al., 2018),
68 as intra-VP mu-opioid activity is sufficient to drive conditioned place aversion (Skoubis and
69 Maidment, 2003), and activating mu opioid receptors in the VP can impair conditioned taste
70 avoidance (Inui and Shimura, 2017). In a similar manner, disinhibiting the VP through injections

71 of the GABAergic antagonist bicuculline induces anxiety-related behaviours and increases
72 avoidance in an approach/avoidance task in primates (Saga et al., 2017; Smith and Berridge, 2005).
73 These findings suggest that the VP plays a role in the motivation to both seek reward and avoid
74 punishment.

75

76 While the VP appears critical for motivating behaviour in appetitive and aversive contexts, how
77 this structure contributes to these opposing motivational drives remains unclear. One possibility is
78 that separate populations of VP neurons drive opposing motivated behaviours. In line with this
79 idea, *in vivo* recordings in the VP have identified two main types of neurons that are activated by
80 the prediction of either reward or punishment (Richard et al., 2016; Saga et al., 2017; Tachibana
81 and Hikosaka, 2012). These neurons encode information related to the expected or incentive value
82 of stimuli (Richard et al., 2016; Tian et al., 2016; Tindell et al., 2006) and their responses are
83 modulated by the internal state of the animal (Tindell et al., 2006). These different populations of
84 VP neurons may be molecularly distinct, as recent optogenetic activation experiments revealed
85 that activation of GABAergic or glutamatergic VP neurons is reinforcing or aversive, respectively
86 (Faget et al., 2018; Knowland et al., 2017; Tooley et al., 2018b).

87

88 However, despite these findings, how the GABAergic or glutamatergic VP neurons normally
89 participate in naturalistic and motivational behaviours, and how these neurons are related to the
90 functionally distinct VP neurons observed in *in vivo* recordings, are unknown. Here our aim was
91 to test the hypothesis that distinct populations of VP neurons play selective roles in appetitive or
92 aversive motivation, and determine if GABAergic and glutamatergic VP neurons play opposing
93 roles in the motivation to approach reward or avoid punishment. Our aim was to determine how

94 these populations encode appetitive and aversive motivational value, examine what roles they play
95 in appetitively or aversively motivated behaviour, and determine how their interactions influence
96 the overall decision to approach or avoid.

97

98 **RESULTS**

99 **Mapping functional classes onto neurochemical identities in the VP**

100 In order to determine what GABAergic and glutamatergic VP neurons are engaged by different
101 motivated behaviours, we trained head-fixed mice on a Pavlovian conditioning task. Each trial
102 started with illumination of a house light and proceeded with presentation of an auditory
103 conditioned stimulus (CS) announcing the delivery of one of the following unconditioned stimuli
104 (USs): 0, 1 or 5 μ l of water in reward blocks, and 0, 100 or 500 ms of air puff blowing to the face
105 in punishment blocks (Figure S1A, Figure 1A, and Methods). As training progressed, mice began
106 licking in response to the reward predicting cues. The lick rate was significantly higher for cues
107 that predicted large rewards than for cues predicting small rewards (Figure S1B) indicating that
108 mice had learnt the CS–US associations.

109

110 We recorded the activity of VP neurons ($n = 331$ neurons / 6 mice; 55 ± 15 per mouse) in *Vglut2-*
111 *Cre;Ai40D* ($n = 2$) or *Gad2-Cre;Ai40D* ($n = 4$) mice, in which glutamatergic or GABAergic VP
112 neurons could be optogenetically tagged due to the expression of the light-sensitive proton
113 pump archaerhodopsin (ArchT) (see Methods) (Figure S1C-S1K). Hierarchical clustering
114 revealed that there were four functional classes of VP neurons (Figure 1B, 1C). All identified
115 glutamatergic neurons belonged to one functional cluster (type II). These neurons were activated
116 by punishment-predictive CSs and punishments, and inhibited by reward-predictive CSs and

117 rewards; we will refer to these as negative valence neurons (NVNs) (Figure 1B-1D). Two other
118 clusters (type III, IV) contained identified GABAergic neurons (Figure 1B-1D). Type IV neurons
119 were activated by reward-predictive CSs and rewards, and inhibited by punishment-predictive CSs
120 and punishments; we'll refer to these as positive valence neurons (PVNs) (Figure 1B-1D). As the
121 vast majority of VP neurons are GABAergic (Faget et al., 2018) (and see Figure S1L, M), we
122 expected that type IV neurons (PVNs) should outnumber type II neurons (NVNs). Indeed, we
123 recorded a total of 216 PVNs (of which 30 were confirmed to be GABAergic with the tagging)
124 and 34 NVNs (of which 12 were confirmed to be glutamatergic with the tagging). For both the
125 NVNs and PVNs, the magnitude of the CS responses was graded, reflecting the expected
126 magnitude of reward or punishment (Figure 1E). We conclude that these neurons bi-
127 directionally and oppositely encode the positive or negative valences of expected and actual
128 outcomes.

129

130 In contrast to these two valence encoding populations, the two other populations appear to encode
131 stimulus saliency as they were excited (type I) or inhibited (type III) by both rewards and
132 punishments or the cues that predict these USs (Figure 1B-1E). Type I neurons were never
133 identified as either glutamatergic or GABAergic and resemble cholinergic basal forebrain neurons
134 that have been described (Hangya et al., 2015). Type III neurons that were inhibited by salient
135 stimuli were identified as GABAergic (Figure 1B-1D).

136

137 Over the course of learning, the responses of both PVNs and NVNs to CS increased while their
138 response to positive and negative US decreased, respectively (Figure 2). This reduction in US
139 response in VP neurons is reminiscent of the temporal backpropagation seen in reward

140 prediction error coding dopamine neurons (Cohen et al., 2012; Pan et al., 2005). To examine if
141 these two populations encode reward and punishment prediction errors, we omitted an expected
142 reward in 10% of trials. PVNs displayed a decrease in firing relative to baseline (Figure 3A, B),
143 consistent with prediction error encoding. We also analysed the responses of PVNs to the
144 neutral CS (which predicts nothing) following the house light. As the reward trials and
145 punishment trials were arranged in a block-wise manner in this task, in principle mice should
146 quickly learn if they were likely to receive reward or punishment in a given block. In a reward
147 block, the onset of house light was followed by a reward CS in the majority ($\sim 2/3$) of trials, but
148 by a neutral CS in a small fraction ($\sim 1/3$) of trials. As the neutral CS represented the worst
149 situation, it might be perceived as being worse than expected. By contrast, in a punishment
150 block the neutral CS represented the best situation and could be perceived as being better than
151 expected. Consistent with this notion, PVNs displayed a decrease in firing relative to baseline
152 when the neutral cue was presented in a reward block (Figure 3C-E). In the punishment block
153 PVNs did not display an increase in firing when the neutral cue was delivered (Figure 3C-E).
154 This suggests that PVNs selectively respond to reward but not punishment prediction errors.
155 NVNs were not significantly modulated when a neutral cue was presented in a punishment
156 block, nor were they modulated when the neutral cue was presented in the reward block (Figure
157 3F, G), indicating that these neurons do not respond when an outcome is better or worse than
158 expected. The properties of the PVNs are similar to those of the reward prediction error-coding
159 dopamine neurons (Cohen et al., 2012; Pan et al., 2005; Schultz et al., 1997).

160

161 While prediction error-coding dopamine neurons respond phasically to reward predicting cues, the
162 duration of the CS response in VP PVNs was variable (Figure 3H-J). Indeed, a number of PVNs

163 had a sustained CS response that lasted till the onset of the US (Figure 3H-K). Notably, these
164 neurons also appeared to encode prediction error, as in the reward block the activity of these
165 neurons was suppressed by the neutral CS (Figure 3L). The sustained responsive neurons are
166 similar to the VP neurons that have been reported to encode state values (Tachibana and Hikosaka,
167 2012). However, when sorted for the duration of the CS responses there did not appear to be two
168 populations of PVNs (phasic and sustained) (Figure 3H), nor did our hierarchical clustering
169 identify two sub-clusters within the population; rather, the population appeared to represent a
170 continuum as has been previously reported (Richard et al., 2016; Richard et al., 2018).

171

172 As the VP has been linked to motivation, we examined if the value coding depended on the
173 motivational state of the mice (Figure 4A-D; S2A-D). In the reward sessions of the Pavlovian task,
174 mice showed vigorous licking response starting from CS onset and lasting until well beyond the
175 delivery of water in the early trials, a period when mice were thirsty, but dramatically reduced
176 licking towards the end of a session (Figure 4B; S2B, D). This decrease in licking presumably
177 reflects a reduction in the motivation to pursue water as the mice were sated. Remarkably, the
178 excitatory CS responses of the PVNs, which were prominent in ‘thirsty trials’, completely
179 disappeared and inverted in ‘sated trials’ (Figure 4A-4D, S2A-S2D). By contrast, the NVNs were
180 differently modulated by animals’ thirsty state. While in thirsty trials these neurons were inhibited
181 by the CS predicting water delivery, in sated trials they were excited by the same CS (Figure 4A-
182 4D). Notably, thirst also strongly modulated the baseline activities of PVNs and NVNs, such that
183 both populations showed markedly lower baseline activities in sated trials than in thirsty trials
184 (Figure 4C-4D).

185

186 Thus, both changes in the predicted value of the environment and the animal's internal state
187 differentially modulates the activities in PVNs and NVNs, as reflected in both the transient cue-
188 induced responses and the tonic baseline activities in these neurons. These results point to the
189 possibility that PVNs and NVNs differentially and potentially opposingly contribute to the
190 generation of incentive and aversive motivation.

191

192 **PVNs and NVNs opposingly and cooperatively regulate motivation**

193 To investigate how PVNs and NVNs might influence the motivation to approach or avoid, we
194 virally expressed channelrhodopsin (ChR2) or ArchT in GABAergic or glutamatergic VP neurons
195 (referred to as PVNs or NVNs, respectively, hereafter for simplicity). We found that optogenetic
196 activation of PVNs or NVNs induced real-time place preference (RTPP) or aversion (RTPA),
197 respectively (Figure S3A-S3D). Furthermore, optogenetic activation of PVNs also supported self-
198 stimulation (Figure S3C). In contrast, optogenetic inhibition of these neurons induced neither
199 RTPP nor RTPA (Figure S3E, F). These results, which are largely consistent with previous
200 findings (Faget et al., 2018; Knowland et al., 2017; Tooley et al., 2018a), suggest that activation
201 of PVNs or NVNs are innately appetitive or aversive, respectively.

202

203 To understand how the PVNs and NVNs contribute to motivated behaviour, we first designed a
204 reward-and-punishment conflict task (or 'conflict task'), in which incentive value can be
205 modulated by either changing reward size or by introducing punishment (Figure 5A). Specifically,
206 in the reward block of this task, mice needed to lick during a choice window following a CS in
207 order to obtain a water reward, whereas in the conflict block, licking during the choice window
208 led to simultaneous delivery of a water reward and an air-puff in each trial (Figure 5A). In both

209 blocks, different CSs predicted rewards of different sizes. We found that the licking probability
210 during the choice window increased as reward size increased, as would be expected from the
211 associated increase in incentive value. In contrast, the probability decreased when the reward was
212 paired with an air-puff, as would be expected from a decrease in incentive value (Figure 5B).

213

214 To test how the different classes of VP neurons influence choice, we optogenetically activated or
215 inhibited these neurons, as described above, during the time window between CS onset and US
216 onset (which covered the choice window) in 20% of randomly selected trials in the conflict task.
217 Notably, activation of PVNs increased the probability that mice would lick in the choice window,
218 in particular when the CS predicted a small reward (i.e., when the motivation to lick was low)
219 (Figure 5C); inhibition of these neurons decreased the probability that mice would lick on a given
220 trial, even if the CS predicted a large reward (i.e., when the incentive to lick was high) (Figure
221 5D). By contrast, activation of NVNs decreased the probability that mice would lick on a given
222 trial (Figure 5C). Although inhibition of NVNs had no effect on the behavior in reward blocks
223 (Figure 5D), inhibition of these neurons in the conflict blocks increased the probability that mice
224 would lick in a trial (Figure 5E). Importantly, inhibition or activation of PVNs did not influence
225 licking per se, either in thirsty or sated mice (Figure S3G, H), indicating that these neurons do not
226 directly control movements. In addition, in control experiments (Figure S4A-H) we verified that
227 light illumination alone in the VP had no effect on animal behavior in the conflict task (Figure
228 S4A, B). These results suggest that PVNs play an essential role in generating or regulating
229 incentive value. By contrast, NVNs are less critical for reward seeking but are essential for
230 constraining reward seeking when there is a potential adverse outcome associated with the action.

231

232 To determine how the endogenous activities of these VP neurons might be modulated by the
233 ‘conflict’, we recorded VP neurons in mice performing a modified version of the conflict task (see
234 Methods). We found that, in the conflict trials as compared to the reward-only trials, the excitatory
235 CS responses of PVNs were reduced, whereas the activity of the NVNs was higher in the conflict
236 trials (Figure 5F, G). These results indicate that perceived risk associated with licking in the
237 conflict task suppresses and increases (or disinhibits), respectively, the responses of PVNs and
238 NVNs to reward cues. These results, together with those from optogenetic manipulations, suggest
239 that the balance between PVN’s and NVN’s activity sets the motivation to seek the reward when
240 there is a motivational conflict.

241

242 While the NVNs do not play a direct role in driving reward-seeking behavior, we hypothesized
243 that they may drive behavior in an aversive context. To determine how PVNs and NVNs contribute
244 to behaviour in an appetitive or aversive context, we designed a “run-for-water” (RFW) task and
245 a “run-to-avoid-air-puff” (RTAA) task, respectively (Figure 6A, B). In the RFW task, mice needed
246 to run in response to a CS in order to obtain a water reward, whereas in the RTAA task mice had
247 to run in response to a CS in order to avoid an air-puff (Figure 6C, D). Once mice learned the tasks,
248 we optogenetically activated or inhibited PVNs or NVNs, as described above, during the time
249 window between CS onset and US onset in 20% of randomly selected trials (Figure 6E-J; Figure
250 S5). We found that, in the RFW task, activating and inhibiting PVNs promoted and abolished
251 running, respectively (Figure 6E, F, H, I; Figure S4C, D). In contrast, activating NVNs decreased
252 the velocity in the RFW task, while inhibiting these neurons did not affect running (Figure 6E, F,
253 H, I; Figure S4C, D). In stark contrast to the RFW task, in the RTAA task, activating PVNs
254 completely abolished running responses, whereas inhibiting these neurons had no effect (Figure

255 6E, G, H, J; Figure S4C, D); moreover, activating and inhibiting NVNs promoted and abolished
256 running, respectively (Figure 6E, G, H, J; Figure S4C, D). Thus, while PVNs promote running in
257 a reward-seeking task, activating them suppresses running in a punishment-avoidance task. In the
258 opposite manner, NVNs are essential for promoting running to avoid punishment, but suppress the
259 same behavior when it is being performed to obtain a reward. Of note, males and females behaved
260 similarly in these tasks (Figure S6; Table S1). These results suggest that the motivational context
261 switches the role that each population plays in behavior; in a reward-seeking context PVNs drive
262 the behaviour while NVNs constrain the behaviour. In an aversive context these roles reverse, with
263 the NVNs driving avoidance behavior and the PVNs actively constraining the avoidance.

264

265 **PVNs and NVNs act via the VP–LHb pathway**

266 In order to explore how downstream structures integrate the activity of PVNs and NVNs, we first
267 looked at the projections of these neurons. Consistent with previous findings, both PVNs and
268 NVNs receive input from areas such as the nucleus accumbens, prefrontal cortex and basolateral
269 amygdala (Figure S7A-S7G) and project to qualitatively the same structures, including the ventral
270 tegmental area, lateral hypothalamus, lateral habenula (LHb) and rostromedial tegmental nucleus
271 (Root et al., 2015; Tooley et al., 2018b). Each of the two functionally distinct populations of VP
272 neurons may synapse onto different cell types within these areas or individual neurons may receive
273 opposing inputs from both of these populations. To explore this, we choose one projection target,
274 the LHb. The VP projects to the medial portion of the LHb (Figure S8) (Herkenham and Nauta,
275 1977; Root et al., 2015), which in turn projects to the dorsal raphe (DR) (Herkenham and Nauta,
276 1979; Quina et al., 2015). The DR also has neurons that encode the motivational value (Cohen et
277 al., 2015) and the behavioural effects of serotonin, like those reported here, also depend on the

278 state of the environment (Seo et al., 2019). Retrograde tracing combined with single molecule in
279 situ hybridization confirmed that both PVNs and NVNs project to the LHb (Figure S9A-C). In
280 addition, using optogenetics combined with electrophysiology in acute slices, we recorded from
281 LHb neurons that were retrogradely labeled from the DR. Every neuron that was recorded received
282 PVN (GABAergic) or NVN (glutamatergic) input, suggesting that individual DR-projecting LHb
283 neurons receive inputs from both VP populations (Figure S9D-G).

284

285 To determine the behavioural effect of the VP→LHb pathway, we optogenetically activated PVN
286 or NVN axon terminals in the LHb. Activation of the PVN (PVN^{VP→LHb}) or NVN (NVN^{VP→LHb})
287 inputs to the LHb induced RTPP or RTPA, respectively (Figure S10A, B; Figure S8). These results
288 confirmed that both PVNs and NVNs send substantial projections to the LHb that can differentially
289 influence animal behavior, consistent with previous findings (Faget et al., 2018; Knowland et al.,
290 2017; Tooley et al., 2018a).

291

292 Next, we assessed the roles of PVN^{VP→LHb} and NVN^{VP→LHb} in the behaviors driven by incentive or
293 aversive value, as described above. In the reward and conflict tasks, we found that optogenetically
294 inhibiting PVN^{VP→LHb} decreased choice in the reward-only block (Figure 7A, B). This
295 manipulation did not reduce choice in the conflict block, likely due to a “floor effect” as this group
296 of mice were highly sensitive to the punishment (Figure 7C). By contrast, inhibiting NVN^{VP→LHb}
297 did not affect choice in the reward-only block, but increased the probability of choice in the conflict
298 block (Figure 7B, C).

299

300 In the RFW and RTAA tasks (see Figure 6), inhibiting PVN^{VP→LHb} abolished running for water

301 (Figure 7D, E), but only slightly reduced running to avoid air-puff (Figure 7F). By contrast,
302 inhibiting NVN^{VP→LHb} did not affect running for water (Figure 7E), but abolished running to avoid
303 air-puff (Figure 7F). Of note, inhibition of either pathway did not induce any obvious effects in
304 the RTPP or RTPA task (Figure S10C, D). Furthermore, control experiments demonstrated that
305 the behavioral effects we observed were not induced by light illumination per se (Figure S4E-H).
306 Together, these results are consistent with the optogenetic manipulations of cells in the VP (Figure
307 5 & 6), and suggest that the distinct behavioral roles of the GABAergic and glutamatergic VP
308 neurons are, at least in part, mediated by their projections to the LHb.

309

310 **DISCUSSION**

311 The decision to approach or avoid depends on the situation in the environment and the internal
312 state of the animal. Here we show that separate populations of VP neurons drive behaviour in
313 different motivational contexts and their activity is differentially regulated by both the valence of
314 the environment and the internal state of the animal. The PVNs are necessary for driving reward-
315 seeking behaviour in positive motivational context, while the NVNs are needed for driving
316 avoidance behaviour in an aversive motivational context. Previous work had described two types
317 of VP neurons that represent the prediction of either reward or punishment (Richard et al., 2016;
318 Saga et al., 2017; Tachibana and Hikosaka, 2012), and had shown that the VP is important for
319 reward seeking and punishment avoidance (Root et al., 2015; Smith et al., 2009; Stephenson-Jones,
320 2019; Wulff et al., 2018). However, the identity of the VP cells representing either reward or
321 punishment prediction is unknown. It is also unclear if these distinct VP representations are
322 causally related to valence-specific behaviours and, if so, how they contribute to behaviours in
323 naturalistic settings.

324

325 We have now linked the NVNs and PVNs in the VP to glutamatergic and GABAergic neurons,
326 respectively, in *in vivo* recording. To our knowledge, this is the first time that the different
327 functional classes of VP neurons have been mapped onto distinct neurochemical identities.
328 Furthermore, we have designed novel behavioral tasks to probe the precise functions of VP
329 neurons. Previous studies showed that activation of glutamatergic or GABAergic VP neurons
330 induces aversive or preference responses, respectively. Our behavioral tasks, which were designed
331 to mimic naturalistic behaviours, allow us to demonstrate how in normal reward-seeking or
332 punishment-avoidance behavior, or in different behavioral contexts (rewarding vs. punishing), the
333 different classes of VP neurons make unique contributions to distinct, adaptive behavioral actions.
334 In particular, we uncover how these two populations interact in risk/reward decision making, and
335 identify a mechanism where the activity of these populations interact in a push/pull manner to
336 determine the overall behavioral strategy. Finally, we show for the first time that the contributions
337 of these distinct VP populations to behavior is state dependent.

338

339 In the natural world, reward seeking is associated with certain costs, such as the effort needed to
340 obtain a reward, the risk of punishment or the presence of threats (Ydenberg, 1986). Our results
341 show that NVNs are needed to constrain reward seeking when there is a risk associated with
342 seeking reward. In line with this a recent study showed that glutamatergic VP neurons are needed
343 to constrain reward seeking when effort is required to obtain a reward, or to limit reward seeking
344 when the reward has been devalued (Tooley et al., 2018b). In light of this finding our results
345 suggest that glutamatergic NVNs represent the risk of punishment associated with seeking-reward
346 and work to balance the incentive value represented by the GABAergic PVNs. In the opposing

347 manner the internal state of an animal, such as thirst may encourage animals to take a risk and
348 reduce the likelihood an animal would avoid a threatening environment. Our results now suggest
349 that PVNs can constrain avoidance. Together our results suggest that the balance between PVN
350 and NVN activity sets the motivation for approach and avoidance.

351

352 If the balance between these populations sets the behavioural response in different motivational
353 contexts, how then do downstream neurons read out this balance? Our results suggest one
354 mechanisms by which this is achieved, i.e., individual neurons downstream of the VP receive input
355 from both PVNs and NVNs. In this way individual neurons can integrate the excitatory and
356 inhibitory input from the VP, with the balance between these drives determining the activity of the
357 postsynaptic neuron. While other patterns of integration likely exist, individual neurons in the VTA
358 also appear to integrate the excitatory and inhibitory input from the VP, as more than half of VTA
359 neurons receive direct GABAergic or glutamatergic VP input (Faget et al., 2018). This suggests
360 that as predicted from their common projection pattern (Faget et al., 2018), the different VP
361 populations have an opposing influence on common downstream targets.

362

363 If downstream neurons integrate input from the different VP populations, they should be activated
364 or inhibited in opposing motivational contexts. In line with this a recent study has shown that
365 GABAergic and serotonergic neurons in the dorsal raphe are selectively activated and drive
366 movement in high threat environments, but are inhibited and suppress movement in a rewarding
367 environment (Seo et al., 2019). Interestingly, the VP innervates the medial portion of the LHb
368 which projects to the dorsal raphe (Herkenham and Nauta, 1977, 1979). As with neurons in the
369 VP, the firing rate of the serotonergic neurons is modulated on a short and long timescale by the

370 predicted value and utility of a stimulus (Cohen et al., 2015). This suggests the contextual
371 dependence of raphe activation may be driven by a switch in the balance between PVN and NVN
372 activity in the VP. On the other hand, the habenula-projecting globus pallidus (GPh) innervates
373 the lateral portion of the LHb with synapses that corelease GABA and glutamate (Shabel et al.,
374 2014). Neurons in this portion of the LHb project to the ventral tegmental area and participate in
375 outcome evaluation during reinforcement learning (Hong and Hikosaka, 2008; Stephenson-Jones
376 et al., 2016). Thus, there appears to be parallel LHb circuits for distinct aspects of motivational
377 behaviours.

378

379 Previous *in vivo* recording data has shown that VP neurons encode a number of variables related
380 to motivation such as the expected reward value (Tachibana and Hikosaka, 2012; Tian et al., 2016),
381 state and action values (Ito and Doya, 2009; Saga et al., 2017; Tachibana and Hikosaka, 2012) as
382 well as reward prediction error (Tian and Uchida, 2015). In line with these findings the activity of
383 a subset of our value coding PVNs could be characterized as encoding reward prediction error or
384 state value. Despite this, our data suggest that value coding PVNs represent one continuous
385 variable with different neurons responding to motivationally relevant cues with different durations.
386 An alternative variable has been suggested to account for the activity pattern of VP neurons, this
387 is incentive value, or the degree to which stimuli have the ability to activate motivational states
388 (Richard et al., 2016; Richard et al., 2018; Smith et al., 2009). This variable may better account
389 for the PVN activity that we describe here, as it relates both the external stimulus value as well as
390 the internal state of the animal (Berridge, 2012) . This is important as previous recordings, as does
391 our study, show that the activity of VP neurons depends on the internal motivational state of the
392 animal as well as the external environment (Smith et al., 2009; Tindell et al., 2006). In light of

393 these findings we propose that value coding PVNs and NVNs encode incentive and aversive value,
394 respectively.

395

396 In our recordings, the identified glutamatergic VP neurons seem to be homogeneous, only
397 belonging to NVNs. One caveat to this finding is that, because the glutamatergic neurons in the
398 VP are sparse, we were only able to tag a small number of these neurons. Future experiments
399 where different methods, such as calcium imaging, are used to simultaneously record a large
400 number of neurons may reveal functional subpopulations within the glutamatergic neurons.
401 Furthermore, we do not know how the expectation signals in VP neurons develop during learning.
402 The imaging or fiber photometry methods can be used to monitor VP neuronal activities
403 throughout learning and thus to address this outstanding question. In addition, future experiments
404 are required to monitor the changes in activity in NVNs while optogenetically manipulating PVNs,
405 and vice versa, and determine the contribution of such changes to behavior. Finally, how the
406 valence-specific neurons interact with the other functional populations, the type I and type III
407 neurons (Figure 1), in the VP during behavior also awaits further study. With regard to the last two
408 points, we found that, in all the neurons that were simultaneously recorded during the optogenetic
409 tagging, there was no change in the firing of the non-tagged neurons when the tagged neurons were
410 inhibited, suggesting that the distinct VP populations are unlikely to be highly connected.

411

412 The decision to approach or avoid depends on the situation in the environment and the internal
413 state of the animal. Here we show that two populations of VP neurons are critical for driving
414 approach and avoidance behaviour. Each of these populations is differentially regulated by both
415 the internal state and the predicted motivational value. These results indicate that the VP is a

416 critical area where information about the internal state and the environmental context is combined
417 to determine the overall behavioural strategy to either approach or avoid.

418

419 **ACKNOWLEDGEMENTS**

420 We thank Dr. Thomas Mrsic-Flogel for critically reading the manuscript, Dr. Z. Josh Huang for
421 providing mouse strains, Ga-Ram Hwang and Dylan Rebolini for technical assistance, and
422 members of the Li laboratory for helpful discussions. This work was supported by grants from
423 NARSAD (M.S., S.A., B.L.), the National Institutes of Health (NIH) (F32MH113316, C.B.R;
424 R01MH101214, R01MH108924, R01NS104944, 1R01DA050374, B.L.), Human Frontier
425 Science Program (RGP0015/2016, B.L.), the Stanley Family Foundation (B.L.), Simons
426 Foundation (344904, B.L.), Wodcroft Foundation (B.L.), the Cold Spring Harbor Laboratory and
427 Northwell Health Affiliation (B.L.) and Feil Family Neuroscience Endowment (B.L.).

428

429 **AUTHOR CONTRIBUTIONS**

430 M.S., C.B.R and B.L. designed the study. M.S. and C.B.R conducted the majority of the
431 experiments and analyzed data. S.A. performed the patch clamp recording experiments. A.F.
432 performed the RNAscope experiments. X.X. generated Matlab codes for controlling behavioral
433 devices and data analysis. C.F.H. performed the rabies tracing experiments. M.S. and B.L. wrote
434 the paper with inputs from all authors.

435

436 **DECLARATION OF INTERESTS**

437 The authors declare no competing interests.

438

439 **MAIN FIGURE TITLES AND LEGENDS**

440 **Figure 1. Separate VP populations opposingly encode motivational value and salience. (A)**

441 Illustrations of experimental design of the reward and punishment classical conditioning tasks.

442 The neutral CSs are the tones that predict nothing. **(B)** Z-score activity plots of the responses of

443 all neurons recorded in the reward and punishment tasks. Red, increase from baseline; blue,

444 decrease from baseline; each row represents one neuron. Green and purple dashes indicate

445 neurons that were optogenetically tagged as being glutamatergic and GABAergic, respectively.

446 **(C)** First three principle components (PC) and hierarchical clustering dendrogram showing the

447 relationship of each neuron within the four clusters. **(D)** Average firing rates of the four types of

448 neurons in the reward and punishment blocks, shown as spike density functions (n=331 from 6

449 mice). **(E)** Average CS response magnitude in the reward and punishment blocks for each of the

450 four types of neurons. All comparisons between the average CS responses were significant (at

451 least $P < 0.05$; Wilcoxon signed-rank test) except between the two neutral CSs (which predict no

452 reward and no punishment in the reward and punishment blocks, respectively, as indicated) in

453 Type I, III and IV neurons. There was also no significant difference between the CS response

454 predicting large and small punishments in Type III neurons. Data in D are presented as mean \pm

455 s.e.m.

456

457 **Figure 2. Development of responses in VP neurons during learning. (A)** The responses to

458 reward cue (CS) or reward (US) in an example PVN tracked over multiple sessions (S1-S9).

459 Responses are shown as spike density plots. **(B)** CS-US (reward) response index for all PVNs

460 across different stages of training ($r^2 = 0.69$, $P < 0.001$ by linear regression). **(C)** Z-score activity

461 plots of the responses of all PVNs during reward and punishment blocks. Each row represents

462 the activities of one neuron. Neurons are sorted according to their CS/US response ratio. **(D)** The
463 CS-US response index for all NVNs in the punishment block across different stages of training
464 ($r^2 = 0.41$, $P < 0.05$ by linear regression). **(E)** The CS and US responses of two example NVNs in
465 the punishment block at different stages of training. Responses are shown as spike density plots.
466 **(F)** Z-score activity plots of the responses of all NVNs during reward and punishment blocks.
467 Each row represents the activities of one neuron. Neurons are sorted according to their CS/US
468 response ratio.

469

470 **Figure 3. VP neuron responses are modulated by expectation.** **(A)** Average firing rates of
471 type IV neurons (PVNs) in response to reward omission, shown as spike density functions. **(B)**
472 auROC analysis of difference in firing rate between baseline and reward omission trials ($n = 15$
473 from 2 mice). Filled bars, $P < 0.05$, t -test **(C)** Graph showing the CS responses of all PVNs during
474 the neutral cue trials in the reward and punishment blocks (reward block, mean, -1.26 Hz;
475 punishment block, mean, -1.05 Hz). Data points in red come from PVN's with sustained
476 responses to the CS. **(D)** Average responses of PVNs during neutral trials in reward and
477 punishments blocks. Responses are shown as spike density plots. **(E)** auROC analysis of
478 difference in firing rate between baseline and neutral cue presentation in PVNs in the reward
479 block ($n = 221$ from 6 mice). Filled bars, $P < 0.05$, t -test. **(F)** Graph showing the CS responses of
480 all NVNs during the neutral cue trials in the reward and punishment blocks (reward block, mean,
481 -0.32 Hz; punishment block, mean, -1.96 Hz). **(G)** Average responses of NVNs during neutral
482 trials in reward and punishments blocks. Responses are shown as spike density plots. **(H)** Z-score
483 activity plots of the responses of all type IV neurons sorted for the duration of their CS response.
484 **(I, J)** Two example neurons showing phasic (G) and sustained (H) responses to reward

485 predicting cues. **(K)** Average responses of all PVNs and the PVNs with a sustained response to
486 CS on large reward trials. Responses are shown as spike density plots. **(L)** Average responses of
487 the ‘sustained’ PVNs, which showed sustained responses to the CS predicting large reward (see
488 **K**), to the CS during neutral trials in reward and punishments blocks. Responses are shown as
489 spike density plots.

490

491 **Figure 4. The response of VP neurons to reward predicting CS depends on the internal**
492 **motivational state. (A)** Top: raster plots showing the neural activity of a PVN (left) and a NVN
493 (right) during large reward trials. Bottom: spike density plots showing the average response of
494 the corresponding two neurons when the mouse was thirsty or sated. **(B)** Raster plot showing the
495 licking behavior in the same behavioral session. **(C)** Average spike density plots showing the
496 activity of PVNs ($n = 19$, from 2 mice) and NVNs ($n = 7$, from 2 mice) in thirsty and sated trials.
497 **(D)** Left: graphs showing the average CS response of PVNs (top) and NVNs (bottom) when mice
498 were thirsty or sated. Right: graphs showing the baseline firing rates of PVNs (top) and NVNs
499 (bottom) when mice were thirsty or sated (** $P < 0.001$, ** $P < 0.01$, * $P < 0.05$; paired t -test).
500 Data in C are presented as mean \pm s.e.m. (shaded areas).

501

502 **Figure 5. The balance of activity between PVNs and NVNs controls reward seeking during**
503 **motivational conflict. (A)** Schematics of the experimental design. **(B)** Motivational conflict
504 reduced reward seeking ($F_{(1,9)} = 35.19$, $p < 0.0001$). ** $p < 0.01$, * $p < 0.05$, two-way ANOVA
505 followed by Tukey’s test. **(C)** Top: a schematic of the approach. Middle: activation of PVNs
506 increased reward seeking ($F_{(1,15)} = 92.32$, $p < 0.001$). Bottom: activation of NVNs decreased
507 reward seeking ($F_{(1,7)} = 108.68$, $p < 0.001$). ** $P < 0.01$, two-way ANOVA followed by Tukey’s

508 test. **(D)** Top: a schematic of the approach. Middle: inhibition of PVNs decreased reward seeking
509 ($F_{(1, 9)} = 50.37, p < 0.0001$). Bottom: inhibition of NVNs had no effect on reward seeking ($F_{(1, 9)}$
510 $= 0.055, p = 0.82$). **P < 0.01, *P < 0.05, two-way ANOVA followed by Tukey's test. **(E)** Top:
511 inhibition of PVNs further decreased reward seeking in the conflict task ($F_{(1, 9)} = 27.60, p <$
512 0.0001). Bottom: inhibition of NVNs increased reward seeking to pre-conflict levels ($F_{(1, 9)} =$
513 $50.37, p < 0.0001$). **p < 0.01, *p < 0.05, two-way ANOVA followed by Tukey's test. **(F)**
514 Average response of PVNs (top; n = 41, from two mice) or NVNs (bottom; n = 14, from two
515 mice) in reward or conflict trials, shown as spike density plots. **(G)** auROC analysis of difference
516 in the CS response during reward and conflict trials. Top: PVNs (n = 41 from two mice); bottom:
517 NVNs (n = 14, from two mice). Filled bars, $P < 0.05$, *t*-test. Data are presented as mean \pm s.e.m.
518

519 **Figure 6. PVNs and NVNs switch roles in controlling actions when motivational context**
520 **changes. (A-D)** The running tasks. **(A)** A schematic of the experimental design. **(B)** Schematics
521 of the experimental procedure. **(C)** Behavioral performance of mice in the run-for-water task (n
522 = 7). **(D)** Behavioral performance of mice in the run-to-avoid-air-puff task (n = 4). **(E)**
523 Schematics of the approach. **(F)** Left: activation of PVNs increased running for water reward
524 ($F_{(1, 13)} = 7.90, p = 0.0055$). Right: activation of NVNs decreased running for water reward ($F_{(1,$
525 $9)} = 132.73, p < 0.001$). **P < 0.01, *P < 0.05, two-way ANOVA followed by Tukey's test. **(G)**
526 Left: activation of PVNs decreased running to avoid air puff ($F_{(1,7)} = 18.76, P < 0.0001$). Right:
527 activation of NVNs increased running to avoid air puff ($F_{(1,7)} = 11.61, P = 0.0010$). **P < 0.01,
528 *P < 0.05, two-way ANOVA followed by Tukey's test. **(H)** Schematics of the approach. **(I)** Left:
529 inhibition of PVNs decreased running for water reward ($F_{(1,9)} = 29.283, p < 0.0001$). Right:
530 inhibition of NVNs had no effect on running for water reward ($F_{(1, 7)} = 0.30, p = 0.59$). **P <

531 0.01, *P < 0.05, two-way ANOVA followed by Tukey's test. **(J)** Left: inhibition of PVNs had no
532 effect on running to avoid air puff ($F_{(1,9)} = 1.30$, $p = 0.26$). Right: inhibition of PVNs decreased
533 running to avoid air puff ($F_{(1,7)} = 22.06$, $p < 0.001$). **P < 0.01, *P < 0.05, two-way ANOVA
534 followed by Tukey's test. Data are presented as mean \pm s.e.m.

535

536 **Figure 7. PVNs and NVNs act via the VP-LHb pathway. (A)** A schematic of the approach.

537 **(B)** Left: inhibition of $PVN^{VP \rightarrow LHb}$ decreased reward seeking ($F_{(1,7)} = 9.55$, $p = 0.0043$). Right:

538 inhibition of $NVN^{VP \rightarrow LHb}$ had no effect on reward seeking ($F_{(1,9)} = 0.0041$, $p = 0.95$). **P < 0.01,

539 two-way ANOVA followed by Tukey's test. **(C)** Left: inhibition of $PVN^{VP \rightarrow LHb}$ did not further

540 decrease reward seeking in these mice in the conflict task ($F_{(1,7)} = 1.39$, $p = 0.25$). Right:

541 inhibition of $NVN^{VP \rightarrow LHb}$ increased reward seeking to pre-conflict levels ($F_{(1,7)} = 13.17$, $p =$

542 0.0010). **P < 0.01, *P < 0.05, two-way ANOVA followed by Tukey's test. **(D)** A schematic of

543 the approach. **(E)** Left: inhibition of $PVN^{VP \rightarrow LHb}$ decreased running for water reward ($F_{(1,11)} =$

544 8.56, $p = 0.004$). Right: inhibition of $NVN^{VP \rightarrow LHb}$ had no effect on running for water reward ($F_{(1,$

545 7) = 0.060, $p = 0.81$). *P < 0.05, two-way ANOVA followed by Tukey's test. **(F)** Left: inhibition

546 of $PVN^{VP \rightarrow LHb}$ had no effect on running to avoid air puff during the cue, although it induced an

547 earlier termination of the running response after the cue ($F_{(1,11)} = 5.57$, $p = 0.020$). Right:

548 inhibition of $NVN^{VP \rightarrow LHb}$ decreased running to avoid air puff ($F_{(1,9)} = 40.72$, $p < 0.0001$). **P <

549 0.01, *P < 0.05, two-way ANOVA followed by Tukey's test. Data are presented as mean \pm

550 s.e.m.

551

552 **STAR METHODS**

553

554 **KEY RESOURCES TABLE**

555

556 **LEAD CONTACT AND MATERIALS AVAILABILITY**

557 Further information and requests for resources and reagents should be directed to and will be
558 fulfilled by the Lead Contact Bo Li (bli@cshl.edu). This study did not generate new unique
559 reagents.

560

561 **EXPERIMENTAL MODEL AND SUBJECT DETAILS**

562 All procedures were approved by the Institutional Animal Care and Use Committee of Cold Spring
563 Harbor Laboratory (CSHL) and conducted in accordance to the United States' National Institutes
564 of Health guidelines. Mice were housed under a 12 h light-dark cycle (8 a.m. to 8 p.m. light). All
565 behavioural experiments were performed during the light cycle. All mice had free access to food,
566 but water was restricted for mice used in certain behavioural experiments. Free water was provided
567 on days with no experimental sessions. Male and female mice 2-4 months of age were used in all
568 experiments. No differences were observed in the behavior of male or female mice during the
569 behavioural tasks (see Figure S6, Table S1). All animals were randomly allocated to the different
570 experimental conditions used in this study. The *Vglut2-Cre* (*Slc17a6^{tm2(cre)Lowl}/J*, stock #016963
571 from Jackson Laboratory, Bar Harbor, Maine, USA), *GAD2-IRES-Cre* (from Dr. Z. Josh Huang,
572 CSHL, available from Jackson Laboratory, *Gad2^{tm2(cre)Zjh}/J*, stock # 010802), *Ai40D*
573 (*Gt(ROSA)_{26Sor}^{tm40.1(CAG-aop3/GFP)Hze/J}*), stock #021188 from Jackson Laboratory), *Rosa26-stop^{fllox-}*
574 *tTA* (stock #012266 from Jackson Laboratory) (Li et al., 2010; Penzo et al., 2015) mouse strains
575 have all been previously characterized. All mice were bred onto a C57BL/6J background.

576

577 **METHOD DETAILS**

578 **Viral vectors.** All adeno-associated viruses (AAV) were produced by the University of North
579 Carolina vector core facility (Chapel Hill, North Carolina, USA) or the University of Pennsylvania
580 vector core (Pennsylvania, USA) and have previously been described: AAV9-Efla-DIO-
581 hChr2(H134R)-eYFP, AAV9-CAG-FLEX-ArchT-GFP, AAV9-Efla-DIO-eYFP and AAV-
582 TRE-hGFP-TVA-G. The EnvA-pseudotyped, protein-G-deleted rabies-EnvA-SAD-ΔG-mCherry
583 virus (Miyamichi et al., 2011) was produced by the Viral Vector Core Facility at Salk Institute.
584 All viral vectors were aliquoted and stored at –80 °C until use.

585

586 **Stereotaxic surgery.** Mice were anesthetized with isoflurane inhalant gas (3%) in an induction
587 chamber and positioned in a stereotaxic frame (myNeuroLab, Leica Microsystems Inc., Buffalo
588 Grove, Illinois, USA). Isoflurane (1.5%) was be delivered through a facemask for anesthesia
589 maintenance. Lidocaine (20 μl) was injected subcutaneously into the head and neck area as a local
590 anaesthetic. For *in vivo* recordings, mice were implanted with a head-bar and a microdrive
591 containing the recording electrodes and an optical fibre. Viral injections were performed using
592 previously described procedures (Penzo et al., 2015) at the following stereotaxic coordinates: VP,
593 0.75 – 0.30 mm from bregma, 1.00 mm lateral from midline, and 5.10 mm ventral from cortical
594 surface; LHb, –1.82 mm from bregma, 0.40 mm lateral from midline, and 2.46 mm ventral from
595 cortical surface; and DR, -4.24 mm from bregma, 0 mm lateral from midline, and 3.00 mm ventral
596 from cortical surface. During the surgical procedure, mice were kept on a heating pad and were
597 brought back to their home-cage for post-surgery recovery and monitoring. Postoperative care
598 included intraperitoneal injection with 0.3-0.5 ml of Lactated Ringer’s solution and Metacam (1-
599 2 mg kg⁻¹ meloxicam; Boehringer Ingelheim Vetmedica, Inc., St. Joseph, Missouri, USA) for

600 analgesia and anti-inflammatory purposes. All AAVs were injected at a total volume of
601 approximately 0.2 μ l, and were allowed at least 4 weeks for maximal expression. For retrograde
602 tracing of projection cells in the VP, CTB-555 (0.2 μ l, 0.5% in phosphate-buffered saline (PBS);
603 Invitrogen, Thermo Fisher Scientific, Waltham, Massachusetts, USA) was injected into the LHb
604 or DR and allowed 5 days for sufficient retrograde transport.

605

606 **Immunohistochemistry.** Immunohistochemistry experiments were performed following standard
607 procedures. Briefly, mice were anesthetized with Euthazol (0.4 ml; Virbac, Fort Worth, Texas,
608 USA) and transcardially perfused with 40 ml of PBS, followed by 40 ml of 4% paraformaldehyde
609 in PBS. Coronal sections (50 μ m) were cut using a freezing microtome (Leica SM 2010R, Leica).
610 Sections were first washed in PBS (3 x 5 min), incubated in PBST (0.3% Triton X-100 in PBS)
611 for 30 min at room temperature (RT) and then washed with PBS (3 x 5 min). Next, sections were
612 blocked in 5% normal goat serum in PBST for 30 min at RT and then incubated with primary
613 antibodies overnight at 4 °C. Sections were washed with PBS (5 x 15 min) and incubated with
614 fluorescent secondary antibodies at RT for 1 h. After washing with PBS (5 x 15 min), sections
615 were mounted onto slides with Fluoromount-G (eBioscience, San Diego, California, USA). Images
616 were taken using a LSM 710 laser-scanning confocal microscope (Carl Zeiss, Oberkochen,
617 Germany). The primary antibodies used were: chicken anti-GFP (1:1000, Aves Labs, catalogue
618 number GFP1020, lot number GFP697986), rabbit anti-RFP (1:1000, Rockland, catalogue number
619 600-401-379, lot number 34135), and rabbit anti-Substance P (SP) (1:1000, ImmunoStar, catalog
620 number 20064, lot number 1531001). Primary antibodies were incubated with appropriate
621 fluorophore-conjugated secondary antibodies (1:1000, Life Technologies, Carlsbad, California,
622 USA) depending on the desired fluorescence colour.

623

624 **Fluorescent *in situ* hybridization.** Single molecule fluorescent *in situ* hybridization (ACDBio,
625 RNAscope) was used to detect expression of *Gad2* and *Slc17a6* in LHb-projecting VP neurons.
626 Alexa-Fluor 555 Conjugate Cholera Toxin Subunit B (CTB555, Thermo Fisher Cat. No. C22843)
627 was injected unilaterally into the LHb of wild-type adult mice. After 5 days, mice were decapitated
628 and their brain tissue was first embedded in cryomolds (Sakura Finetek, Cat. No. 25608-924) filled
629 with M-1 Embedding Matrix (Thermo Scientific, Cat. No. 1310) then quickly fresh-frozen on dry
630 ice. The tissue was stored at -80°C until it was sectioned. 12 µm cryostat-cut sections containing
631 VP were collected rostro-caudally in a series of four slides and stored at -80°C, until used for
632 hybridization. Briefly, the day of the experiment, frozen sections were post-fixed in 4% PFA in
633 RNA-free PBS (hereafter referred to as PBS) at room temperature (RT) for 15', then washed in
634 PBS, dehydrated using increasing concentrations of ethanol in water (50% once, 70% once, 100%
635 twice) for 5'. Sections were then dried at RT and incubated with Protease IV for 30' at RT. Sections
636 were washed in PBS three times for 5' at RT, then hybridized. Probes against *Gad2* (Cat. No.
637 #439371, dilution 1:50) and *Slc17a6* (Vglut2) (Cat. No. #319171-C3, dilution 1:50) were applied
638 to VP sections. Hybridization was carried out for 2h at 40°C. After that, sections were washed
639 twice in PBS at RT for 2', then incubated with three consecutive rounds of amplification reagents
640 (30', 15' and 30', respectively, at 40°C). After each amplification step, sections were washed in
641 PBS at RT for 2', twice. Finally, fluorescence detection was carried out for 15' at 40°C.
642 Fluorescent dyes used were Alexa-488 (for *Gad2* detection) and Atto-647 (for *Slc17a6* detection).
643 CTB555 signal was detected in the red channel. Sections were then washed twice in PBS,
644 incubated with DAPI (1:10000 in PBS) for 2', washed twice in PBS for 2', then mounted with

645 cover-slip using mounting medium. Images were acquired using an LSM780 confocal microscope
646 with 20x or 40x lenses, and visualized and processed using ImageJ and Adobe Illustrator.

647

648 For quantifications of fluorescent *in situ* hybridization samples, only CTB555⁺ cells with a clear
649 DAPI⁺ nucleus were counted. The percentage of CTB555⁺ cells expressing Gad2, Vglut2, or both
650 markers (i.e. Gad2⁺Vglut2⁺ cells) in total CTB555⁺ cells were calculated based on images of the
651 same sections from the CTB experiments. Briefly, the number of Gad2⁺, Vglut2⁺ and
652 Gad2⁺Vglut2⁺ neurons in a region of interest (ROI) within the VP were annotated and then
653 normalized to the area of the ROI.

654

655 **Monosynaptic tracing with pseudotyped rabies virus.** Retrograde tracing of monosynaptic
656 inputs onto genetically-defined cell populations of the VP was accomplished using a previously
657 described method (Callaway and Luo, 2015; Penzo et al., 2015). In brief, *Vglut2-Cre;Rosa26-*
658 *stop^{fllox}-tTA* or *GAD2-Cre;Rosa26-stop^{fllox}-tTA* mice that express tTA in glutamatergic or
659 GABAergic cells, respectively, were injected into the VP with AAV-TRE-hGFP-TVA-G (0.2–0.3
660 μl) that expresses the following components in a tTA-dependent manner: a fluorescent reporter
661 histone GFP (hGFP); TVA (which is a receptor for the avian virus envelope protein EnvA); and
662 the rabies envelope glycoprotein (G). Two weeks later, mice were injected in the same VP location
663 with the rabies-EnvA-SAD-DG-mCherry (500 nl), a rabies virus that is pseudotyped with EnvA,
664 lacks the envelope glycoprotein, and expresses mCherry. This method ensures that the rabies virus
665 exclusively infects cells expressing TVA. Furthermore, complementation of the modified rabies
666 virus with envelope glycoprotein in the TVA-expressing cells allows the generation of infectious
667 particles, which then can trans-synaptically infect presynaptic neurons.

668

669 **Classical conditioning task.** Four *GAD2-Cre;Ai40D* and two *Vglut2-Cre;Ai40D* mice were
670 trained on an auditory classical conditioning task. One week after surgery mice were water-
671 deprived in their home-cage. During training, mice were head restrained using custom-made
672 clamps and metal head-bars. Each mouse was habituated to head restraint for one day prior to
673 training. There were five possible outcomes (unconditioned stimuli, US), each associated with a
674 different auditory cue (conditioned stimulus, CS): a large water reward (5 μ l), a small water reward
675 (1 μ l), nothing, a small air puff (100 ms) or a large air puff (500 ms). The air puff was delivered
676 to the animal's face. Each trial began with a houselight turning on (the light stayed on for 5
677 seconds). A CS (1 second sound) was played one second after the houselight was turned on,
678 followed by a 0.5 second delay and then a US (the outcome). For the CSs, we used 75 dB tones
679 with 12, 10, 8, 4, 2 kHz, and 75 dB white noise.

680

681 In each session, reward and punishment trials were presented in two sequential blocks, with each
682 cue chosen pseudorandomly. Each block contained the neutral CS. We defined satiated trials as
683 trials where the mouse licked two or less times in the choice window.

684

685 For recording in the modified 'conflict task' (see below) the same auditory classical conditioning
686 paradigm was used, except that one CS predicted a water reward (12 μ l), another CS predicted the
687 simultaneous delivery of both the water reward and a 100-ms air puff blowing to the face (the trials
688 with different CSs were randomly interleaved).

689

690 **Reward-and-punishment conflict task.** Mice were first trained and tested in the reward-only
691 task, and subsequently trained and tested in the reward-and-punishment conflict task (or ‘conflict
692 task’). In each trial of the reward-only task, one of five distinct auditory cues (CS; 1-s duration,
693 presented pseudo-randomly) was presented, followed by a 500-ms delay and then an outcome
694 (US). Each CS uniquely predicted one of five sizes of water reward: 3, 5, 7, 10 or 12 μ l. Water
695 delivery was contingent on mice licking the waterspout during a 1-s choice window, which
696 spanned the last 500 ms of CS and the 500-ms delay after CS ended. Failure of licking during the
697 choice window led to omission of water reward. Mice were trained with one session per day (250
698 trials per session; inter-trial-interval, 8 s). In the optogenetic testing day, animals were subjected
699 to 250 trials, wherein laser stimulation occurred in 20% of randomly interleaved trials (laser
700 stimulation started from CS onset and lasted until the time of water delivery). Next, mice were
701 trained in the ‘conflict task’, which was similar to the reward-only task except that licking the
702 waterspout during the choice window triggered delivery of both the water reward and a 100-ms
703 air puff blowing to the face.

704

705 For the reward-only part of the ‘conflict task’, mice were trained for 4-8 weeks until they exhibited
706 an escalated choosing of trials during the choice window proportional to increasing reward for two
707 days in a row, and then were tested. For the conflict part of this task, we trained mice for 1-2 weeks
708 in the conflict situation following training in the reward-only situation. Mice were trained until
709 they showed a 60% or less probability in choosing the biggest reward (12 μ l) under the conflict
710 situation for two days in a row before testing.

711

712 **Run-for-water task.** Mice were water deprived for a day prior to training in the run-for-water
713 (RFW) task. After being habituated to the head-fixation apparatus and the running wheel, mice
714 were presented with a CS (12-kHz, 1-s tone) that predicted the conditional availability of water
715 (10 μ l) in a spout close to the mouth. Only if the mice reached a running velocity of 10 cm/s in a
716 response window spanning from tone onset to 500 ms after the tone ended. Failure to reach the
717 velocity threshold during the response window resulted in water omission. Animals were trained
718 in one session per day for 4-8 weeks (100 trials per session; average inter-trial-interval, 30 s), until
719 they exhibited a reliable running response to the tone (2 z-scores above baseline running during
720 any period within the CS) for two consecutive days before optogenetic testing.

721

722 For the optogenetic testing, animals received 100 trials, where 20 randomly interleaved trials
723 included laser delivery that covered the period from CS onset to the onset of water delivery. For
724 all optogenetic manipulations, the laser parameter used were: pulse duration, 5 ms (for ChR2) or
725 constant light (for ArchT); frequency, 30 Hz (for ChR2) or constant light (for ArchT); train
726 duration, 1.5 sec; laser intensity, 10 mW (measured at the tip of optic fibers). Throughout the
727 training and testing, animals were ensured to receive a total of 1 ml of water every day.

728

729 **Run-to-avoid-air-puff task.** Mice were habituated to the head-fixation apparatus and the running
730 wheel prior to running-to-avoid-air-puff task (RTAA) training. Mice were then presented with
731 auditory tones (white noise, 1 s) that predicted delivery of an air puff to the face (400 ms) if the
732 animal did not reach a running speed of 10 cm/s at some point from tone onset to 500 ms after tone
733 offset. Failure to reach the speed threshold resulted in punishment delivery. The intertrial interval
734 was 30 s on average. Animals were trained every day for one session of 100 trials. Animals were

735 trained from 4-8 weeks until they exhibited a reliable running response to the tone (2 z-scores
736 above baseline running) before optogenetic testing. For the optogenetic testing day, animals
737 received 100 trials, where 20 randomly interleaved trials included laser delivery that covered from
738 tone onset to the air puff delivery time.

739

740 **Real-time place preference or aversion task.** Freely moving mice were initially habituated to a
741 two-sided chamber, and were subsequently subjected to a 10-min session in which laser
742 stimulation (laser power, 10 mW (measured at fiber tip); laser frequency, 40 Hz for ChR2
743 experiments or the laser was constantly on for ArchT experiments) was turned on once mice
744 entered the left side of the chamber. This was followed by another 10-min session in which laser
745 stimulation was turned on once mice entered the right side of the chamber. We recorded the
746 percentage of time the mice spent on either side of the chamber during baseline and during the
747 laser stimulation.

748

749 **Self-stimulation.** Freely moving mice were placed in a chamber equipped with two nose-poke
750 ports. Nose-poking to one of the ports (the active port) triggered laser delivery (pulse duration, 5
751 ms, train duration, 2 s, intensity, 10 mW (measured at fiber tip), frequency, 40 Hz), while nose-
752 poking to the other port (the inactive port) did not trigger laser delivery. Mice were allowed to
753 freely poke the two ports and were tested in a single 1-hr session.

754

755 ***In vivo* electrophysiology.** Custom-built screw-driven microdrives with 4 implantable tetrodes
756 and a 50 μm fibre-optic were used to record simultaneously from multiple neurons. Each tetrode
757 was glued to the fibre-optic with epoxy, such that the end of each tetrode was 200-400 μm from

758 the end of the fibre-optic. Neural recordings and time stamps for behavioural variables were
759 acquired with a Tucker-Davis Technologies RZ recording system (with a 32 channel preamp PZ2-
760 32 and a RZ5D Bioamp processor; Alachua, Florida, USA).

761

762 Broadband signals from each wire were filtered between 0.2 and 8,500 Hz and recorded
763 continuously at 25 kHz. To extract the timing of spikes, signals were band-pass-filtered between
764 300-5,000 Hz. Data analyses were carried out using software in Matlab (The Mathworks, Inc.,
765 Natick, Massachusetts, USA). Spike waveforms were manually sorted offline based on amplitude
766 and waveform energy features using MClust-3.5 (from Dr. A. David Redish, University of
767 Minnesota, Minneapolis, Minnesota, USA). Individual neurons were only included in the dataset
768 if they were well isolated based on their isolation distance (>20) and L-ratio (<0.1) (Schmitzer-
769 Torbert et al., 2005). Prior to implantation, tetrodes were dipped in DiI to aid the post-hoc
770 visualization of the recording locations.

771

772 In order to convert raster plots of firing rate into continuous spike density functions, spike times
773 were first binned into 1 ms time windows and then convolved with a Gaussian kernel ($\sigma = 15$ ms).
774 To determine the response to the CS or US presentation, the average firing rates were calculated
775 using a 300 ms window defined as 180-480 ms following the stimulus. These time windows were
776 chosen to cover the time of the peak neuronal response. Average baseline firing was calculated
777 using a 300 ms window immediately preceding the delivery of the CS.

778

779 To identify putative GABAergic or glutamatergic VP neurons, we used ArchT-mediated optic
780 tagging (Courtin et al., 2014), whereby 200 ms light pulses ($\lambda = 532$ nm; OEM Laser Systems Inc.,

781 Bluffdale, Utah, USA) were delivered every 5 seconds for 100 trials following each behavioural
782 recording session. In early sessions we also used 500 ms ($n = 3$) or 1 second ($n = 1$) light pulses,
783 which tagged VP neurons in a similar way to that of the 200 ms light pulses. Units that showed
784 rapid suppression (with latencies < 10 ms) in response to the laser stimulation and had neural
785 activity suppressed to below 0.5Hz during laser stimulation were considered to be optogenetically
786 tagged.

787

788 In addition to their response to light, putative VP neurons were identified based on their firing
789 pattern through a previously described unsupervised clustering approach (Cohen et al., 2012).
790 Briefly, the first three principal components (PCs) of the Z scored firing rates of all neurons in the
791 reward and punishment blocks were calculated using principal component analysis (PCA), with
792 the singular value decomposition algorithm. Hierarchical clustering of the first three PCs was then
793 performed using a Euclidean distance metric and a complete agglomeration method.

794

795 Cross-correlations between spike waveforms across sessions were used to determine whether the
796 same unit was recorded over multiple sessions. The cross-correlations were calculated after
797 aligning the negative peak of each waveform, averaging separately, and aligning the peaks of the
798 averages. A conservative session-to-session cross-correlation coefficient of >0.95 was used to
799 positively classify two sets of waveforms as belonging to the same unit. The correlation was
800 calculated using the full duration of the spike in a window 10 ms prior to and 40 ms after the peak
801 negative response.

802

803 CS-US indices were calculated as $(CS - US)/(CS + US)$, where CS is the difference between the

804 peak firing rate (maximum value of the PSTH) in the 500 ms after CS onset and the baseline firing
805 rate, and US is the difference between the peak firing rate in the 500 ms after US onset and the
806 baseline firing rate. The baseline firing rate was calculated as the mean of the PSTH in the 0.5 s
807 before CS onset.

808

809 To calculate receiver-operating characteristic (ROC) curves, the distributions of firing rates were
810 compared between 1 second of activity prior to the CS presentation (baseline activity) and 1 second
811 of activity after the CS presentation, or between the distributions of firing rates following two
812 different cues.

813

814 ***In vitro* electrophysiology.** Patch clamp recording was performed as previously described (Penzo
815 et al., 2015). Briefly, mice were anesthetized with isoflurane before they were decapitated; their
816 brains were then dissected out and placed in ice chilled dissection buffer (110 mM choline chloride,
817 25 mM NaHCO₃, 1.25 mM NaH₂PO₄, 2.5 mM KCl, 0.5 mM CaCl₂, 7.0 mM MgCl₂, 25.0 mM
818 glucose, 11.6 mM ascorbic acid and 3.1 mM pyruvic acid, gassed with 95% O₂ and 5% CO₂). An
819 HM650 Vibrating-blade Microtome (Thermo Fisher Scientific) was then used to cut 300 μm thick
820 coronal sections that contained the LHb. These slices were subsequently transferred to a storage
821 chamber that contained oxygenated artificial cerebrospinal fluid (ACSF) (118 mM NaCl, 2.5 mM
822 KCl, 26.2 mM NaHCO₃, 1 mM NaH₂PO₄, 20 mM glucose, 2 mM MgCl₂ and 2 mM CaCl₂, at 34
823 °C, pH 7.4, gassed with 95% O₂ and 5% CO₂). Following 40 min of recovery time, slices were
824 transferred to RT (20-24 °C), where they were continuously bathed in the ACSF.

825

826 Whole-cell patch clamp recording from LHb neurons was obtained with Multiclamp 700B

827 amplifiers and pCLAMP 10 software (Molecular Devices, Sunnyvale, California, USA), and was
828 visually guided using an Olympus BX51 microscope equipped with both transmitted and
829 epifluorescence light sources (Olympus Corporation, Shinjuku, Tokyo, Japan). DR-projecting
830 LHb neurons retrogradely labeled with CTB555 (red fluorescent) were identified and patched. The
831 external solution was ACSF. The internal solution contained 115 mM caesium methanesulphonate,
832 20 mM CsCl, 10 mM HEPES, 2.5 mM MgCl₂, 4 mM Na₂ATP, 0.4 mM Na₃GTP, 10 mM sodium
833 phosphocreatine and 0.6 mM EGTA (pH 7.2).

834

835 As the acute slices were prepared from mice in which GABAergic or glutamatergic VP neurons
836 were infected with AAV expressing ChR2-YFP, to evoke VP synaptic transmission onto LHb
837 neurons, a blue light was used to stimulate ChR2-expressing axons originating from the VP. The
838 light source was a single-wavelength LED system ($\lambda = 470$ nm; <http://www.cooled.com/>)
839 connected to the epifluorescence port of the Olympus BX51 microscope. A light pulse of 1 ms,
840 triggered by a TTL signal from the Clampex software, was delivered every 10 seconds to drive
841 synaptic responses. Inhibitory or excitatory postsynaptic currents (IPSCs or EPSCs, respectively)
842 were low-pass filtered at 1 KHz and recorded. IPSCs were recorded at a holding potential of 0 mV,
843 with APV (100 μ M) and CNQX (5 μ M) added to the ACSF. EPSCs were recorded at holding
844 potentials of -70 mV (for AMPA-receptor-mediated responses) and $+40$ mV (for NMDA-
845 receptor-mediated responses), with picrotoxin (100 μ M) added to the ACSF. Synaptic responses
846 were analyzed using pCLAMP 10 software.

847

848 **QUANTIFICATION AND STATISTICAL ANALYSIS**

849 To determine whether parametric tests could be used, the Shapiro-Wilk Test was performed on all

850 data as a test for normality. The statistical test used for each comparison is indicated when used.
851 The sample sizes used in this study were based on estimations by a power analysis. Behavioural
852 tests and electrophysiological data acquisition were performed by investigators with knowledge of
853 the identities of experimental groups. All these experiments were controlled by computer systems,
854 with data collected and analysed in an automated and unbiased way. No data points were excluded.

855

856 **DATA AND CODE AVAILABILITY**

857 This study did not generate datasets/code.

858

859 **REFERENCES**

860

861 Berridge, K.C. (2012). From prediction error to incentive salience: mesolimbic computation of
862 reward motivation. *Eur J Neurosci* 35, 1124-1143.

863 Callaway, E.M., and Luo, L. (2015). Monosynaptic Circuit Tracing with Glycoprotein-Deleted
864 Rabies Viruses. *J Neurosci* 35, 8979-8985.

865 Cohen, J.Y., Amoroso, M.W., and Uchida, N. (2015). Serotonergic neurons signal reward and
866 punishment on multiple timescales. *Elife* 4.

867 Cohen, J.Y., Haesler, S., Vong, L., Lowell, B.B., and Uchida, N. (2012). Neuron-type-specific
868 signals for reward and punishment in the ventral tegmental area. *Nature* 482, 85-88.

869 Courtin, J., Chaudun, F., Rozeske, R.R., Karalis, N., Gonzalez-Campo, C., Wurtz, H., Abdi, A.,
870 Baufreton, J., Bienvenu, T.C., and Herry, C. (2014). Prefrontal parvalbumin interneurons shape
871 neuronal activity to drive fear expression. *Nature* 505, 92-96.

872 Faget, L., Zell, V., Souter, E., McPherson, A., Ressler, R., Gutierrez-Reed, N., Yoo, J.H., Dulcis,
873 D., and Hnasko, T.S. (2018). Opponent control of behavioral reinforcement by inhibitory and
874 excitatory projections from the ventral pallidum. *Nat Commun* 9, 849.

875 Farrar, A.M., Font, L., Pereira, M., Mingote, S., Bunce, J.G., Chrobak, J.J., and Salamone, J.D.
876 (2008). Forebrain circuitry involved in effort-related choice: Injections of the GABAA agonist
877 muscimol into ventral pallidum alter response allocation in food-seeking behavior. *Neuroscience*
878 152, 321-330.

879 Haber, S.N., and Knutson, B. (2010). The reward circuit: linking primate anatomy and human
880 imaging. *Neuropsychopharmacology* : official publication of the American College of
881 Neuropsychopharmacology 35, 4-26.

882 Hangya, B., Ranade, S.P., Lorenc, M., and Kepecs, A. (2015). Central Cholinergic Neurons Are
883 Rapidly Recruited by Reinforcement Feedback. *Cell* 162, 1155-1168.

884 Heimer, L., Harlan, R.E., Alheid, G.F., Garcia, M.M., and de Olmos, J. (1997). Substantia
885 innominata: a notion which impedes clinical-anatomical correlations in neuropsychiatric disorders.
886 *Neuroscience* 76, 957-1006.

887 Herkenham, M., and Nauta, W.J. (1977). Afferent connections of the habenular nuclei in the rat.
888 A horseradish peroxidase study, with a note on the fiber-of-passage problem. *J Comp Neurol* 173,
889 123-146.

890 Herkenham, M., and Nauta, W.J. (1979). Efferent connections of the habenular nuclei in the rat. *J*
891 *Comp Neurol* 187, 19-47.

892 Hong, S., and Hikosaka, O. (2008). The globus pallidus sends reward-related signals to the lateral
893 habenula. *Neuron* 60, 720-729.

894 Humphries, M.D., and Prescott, T.J. (2010). The ventral basal ganglia, a selection mechanism at
895 the crossroads of space, strategy, and reward. *Prog Neurobiol* 90, 385-417.

896 Inui, T., and Shimura, T. (2017). Activation of mu-opioid receptors in the ventral pallidum
897 decreases the negative hedonic evaluation of a conditioned aversive taste in rats. *Behav Brain Res*
898 320, 391-399.

899 Ito, M., and Doya, K. (2009). Validation of decision-making models and analysis of decision
900 variables in the rat basal ganglia. *J Neurosci* 29, 9861-9874.

901 Knowland, D., Lilascharoen, V., Pacia, C.P., Shin, S., Wang, E.H., and Lim, B.K. (2017). Distinct
902 Ventral Pallidal Neural Populations Mediate Separate Symptoms of Depression. *Cell* 170, 284-
903 297 e218.

904 Lammel, S., Lim, B.K., Ran, C., Huang, K.W., Betley, M.J., Tye, K.M., Deisseroth, K., and
905 Malenka, R.C. (2012). Input-specific control of reward and aversion in the ventral tegmental area.
906 *Nature* 491, 212-217.

907 Lecca, S., Meye, F.J., Trusel, M., Tchenio, A., Harris, J., Schwarz, M.K., Burdakov, D., Georges,
908 F., and Mameli, M. (2017). Aversive stimuli drive hypothalamus-to-habenula excitation to
909 promote escape behavior. *Elife* 6.

910 Li, L., Tasic, B., Micheva, K.D., Ivanov, V.M., Spletter, M.L., Smith, S.J., and Luo, L. (2010).
911 Visualizing the distribution of synapses from individual neurons in the mouse brain. *PloS one* 5,
912 e11503.

- 913 Mogenson, G.J., Jones, D.L., and Yim, C.Y. (1980). From motivation to action: functional
914 interface between the limbic system and the motor system. *Prog Neurobiol* *14*, 69-97.
- 915 Pan, W.X., Schmidt, R., Wickens, J.R., and Hyland, B.I. (2005). Dopamine cells respond to
916 predicted events during classical conditioning: evidence for eligibility traces in the reward-learning
917 network. *J Neurosci* *25*, 6235-6242.
- 918 Panagis, G., Miliareisis, E., Anagnostakis, Y., and Spyrali, C. (1995). Ventral pallidum self-
919 stimulation: a moveable electrode mapping study. *Behav Brain Res* *68*, 165-172.
- 920 Panagis, G., Nomikos, G.G., Miliareisis, E., Chergui, K., Kastellakis, A., Svensson, T.H., and
921 Spyrali, C. (1997). Ventral pallidum self-stimulation induces stimulus dependent increase in c-fos
922 expression in reward-related brain regions. *Neuroscience* *77*, 175-186.
- 923 Penzo, M.A., Robert, V., Tucciarone, J., De Bundel, D., Wang, M., Van Aelst, L., Darvas, M.,
924 Parada, L.F., Palmiter, R.D., He, M., *et al.* (2015). The paraventricular thalamus controls a central
925 amygdala fear circuit. *Nature* *519*, 455-459.
- 926 Quina, L.A., Tempest, L., Ng, L., Harris, J.A., Ferguson, S., Zhou, T.C., and Turner, E.E. (2015).
927 Efferent pathways of the mouse lateral habenula. *J Comp Neurol* *523*, 32-60.
- 928 Richard, J.M., Ambroggi, F., Janak, P.H., and Fields, H.L. (2016). Ventral Pallidum Neurons
929 Encode Incentive Value and Promote Cue-Elicited Instrumental Actions. *Neuron* *90*, 1165-1173.
- 930 Richard, J.M., Stout, N., Acs, D., and Janak, P.H. (2018). Ventral pallidal encoding of reward-
931 seeking behavior depends on the underlying associative structure. *Elife* *7*.
- 932 Root, D.H., Melendez, R.I., Zaborszky, L., and Napier, T.C. (2015). The ventral pallidum:
933 Subregion-specific functional anatomy and roles in motivated behaviors. *Prog Neurobiol* *130*, 29-
934 70.
- 935 Saga, Y., Richard, A., Sgambato-Faure, V., Hoshi, E., Tobler, P.N., and Tremblay, L. (2017).
936 Ventral Pallidum Encodes Contextual Information and Controls Aversive Behaviors. *Cereb Cortex*
937 *27*, 2528-2543.
- 938 Schmitzer-Torbert, N., Jackson, J., Henze, D., Harris, K., and Redish, A.D. (2005). Quantitative
939 measures of cluster quality for use in extracellular recordings. *Neuroscience* *131*, 1-11.
- 940 Schultz, W., Dayan, P., and Montague, P.R. (1997). A neural substrate of prediction and reward.
941 *Science* *275*, 1593-1599.
- 942 Seo, C., Guru, A., Jin, M., Ito, B., Sleezer, B.J., Ho, Y.Y., Wang, E., Boada, C., Krupa, N.A.,
943 Kullakanda, D.S., *et al.* (2019). Intense threat switches dorsal raphe serotonin neurons to a
944 paradoxical operational mode. *Science* *363*, 538-542.
- 945 Shabel, S.J., Proulx, C.D., Piriz, J., and Malinow, R. (2014). Mood regulation. GABA/glutamate
946 co-release controls habenula output and is modified by antidepressant treatment. *Science* *345*,
947 1494-1498.

- 948 Skoubis, P.D., and Maidment, N.T. (2003). Blockade of ventral pallidal opioid receptors induces
949 a conditioned place aversion and attenuates acquisition of cocaine place preference in the rat.
950 *Neuroscience 119*, 241-249.
- 951 Smith, K.S., and Berridge, K.C. (2005). The ventral pallidum and hedonic reward: neurochemical
952 maps of sucrose "liking" and food intake. *J Neurosci 25*, 8637-8649.
- 953 Smith, K.S., Tindell, A.J., Aldridge, J.W., and Berridge, K.C. (2009). Ventral pallidum roles in
954 reward and motivation. *Behav Brain Res 196*, 155-167.
- 955 Stamatakis, A.M., Jennings, J.H., Ung, R.L., Blair, G.A., Weinberg, R.J., Neve, R.L., Boyce, F.,
956 Mattis, J., Ramakrishnan, C., Deisseroth, K., *et al.* (2013). A unique population of ventral
957 tegmental area neurons inhibits the lateral habenula to promote reward. *Neuron 80*, 1039-1053.
- 958 Stamatakis, A.M., and Stuber, G.D. (2012). Activation of lateral habenula inputs to the ventral
959 midbrain promotes behavioral avoidance. *Nat Neurosci 15*, 1105-1107.
- 960 Stephenson-Jones, M. (2019). Pallidal circuits for aversive motivation and learning (*Current*
961 *Opinion in Behavioral Sciences*), pp. 82-89.
- 962 Stephenson-Jones, M., Yu, K., Ahrens, S., Tucciarone, J.M., van Huijstee, A.N., Mejia, L.A.,
963 Penzo, M.A., Tai, L.H., Wilbrecht, L., and Li, B. (2016). A basal ganglia circuit for evaluating
964 action outcomes. *Nature 539*, 289-293.
- 965 Stopper, C.M., and Floresco, S.B. (2014). What's better for me? Fundamental role for lateral
966 habenula in promoting subjective decision biases. *Nat Neurosci 17*, 33-35.
- 967 Stratford, T.R., Kelley, A.E., and Simansky, K.J. (1999). Blockade of GABAA receptors in the
968 medial ventral pallidum elicits feeding in satiated rats. *Brain research 825*, 199-203.
- 969 Tachibana, Y., and Hikosaka, O. (2012). The primate ventral pallidum encodes expected reward
970 value and regulates motor action. *Neuron 76*, 826-837.
- 971 Tian, J., Huang, R., Cohen, J.Y., Osakada, F., Kobak, D., Machens, C.K., Callaway, E.M., Uchida,
972 N., and Watabe-Uchida, M. (2016). Distributed and Mixed Information in Monosynaptic Inputs to
973 Dopamine Neurons. *Neuron 91*, 1374-1389.
- 974 Tian, J., and Uchida, N. (2015). Habenula Lesions Reveal that Multiple Mechanisms Underlie
975 Dopamine Prediction Errors. *Neuron 87*, 1304-1316.
- 976 Tindell, A.J., Smith, K.S., Pecina, S., Berridge, K.C., and Aldridge, J.W. (2006). Ventral pallidum
977 firing codes hedonic reward: when a bad taste turns good. *J Neurophysiol 96*, 2399-2409.
- 978 Tooley, J., Marconi, L., Alipio, J.B., Matikainen-Ankney, B., Georgiou, P., Kravitz, A.V., and
979 Creed, M.C. (2018a). Glutamatergic Ventral Pallidal Neurons Modulate Activity of the Habenula-
980 Tegmental Circuitry and Constrain Reward Seeking. *Biol Psychiatry 83*, 1012-1023.

- 981 Tooley, J., Marconi, L., Alipio, J.B., Matikainen-Ankney, B., Georgiou, P., Kravitz, A.V., and
982 Creed, M.C. (2018b). Glutamatergic Ventral Pallidal Neurons Modulate Activity of the Habenula-
983 Tegmental Circuitry and Constrain Reward Seeking. *Biol Psychiatry*.
- 984 Wulff, A.B., Tooley, J., Marconi, L.J., and Creed, M.C. (2018). Ventral pallidal modulation of
985 aversion processing. *Brain Res*.
- 986 Ydenberg, R.C. (1986). The economics of fleeing from predators, L.M. Dill, ed. (*Advances in the*
987 *Study of Behaviour: Elsevier*), pp. 229-249.
- 988

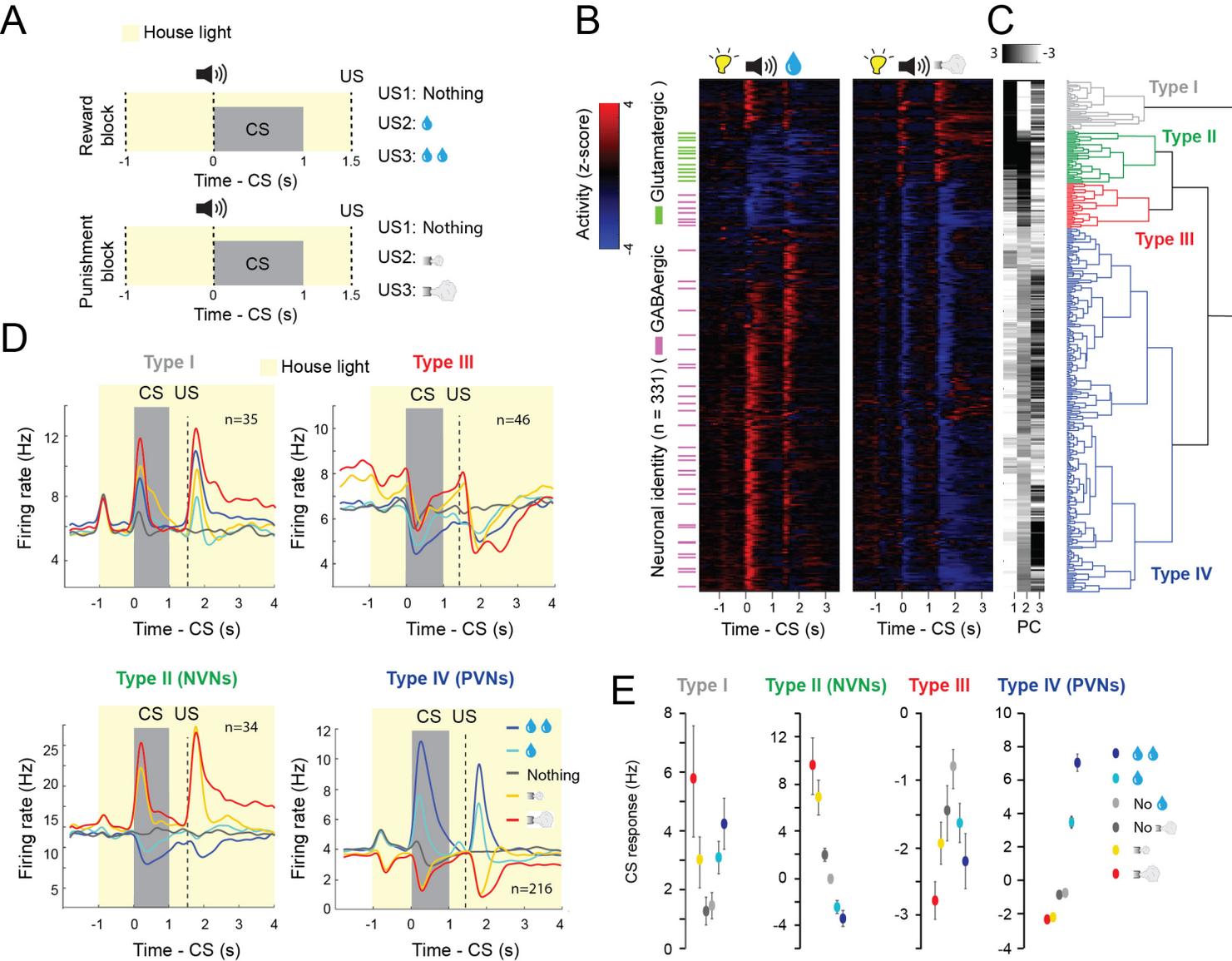


Figure 1

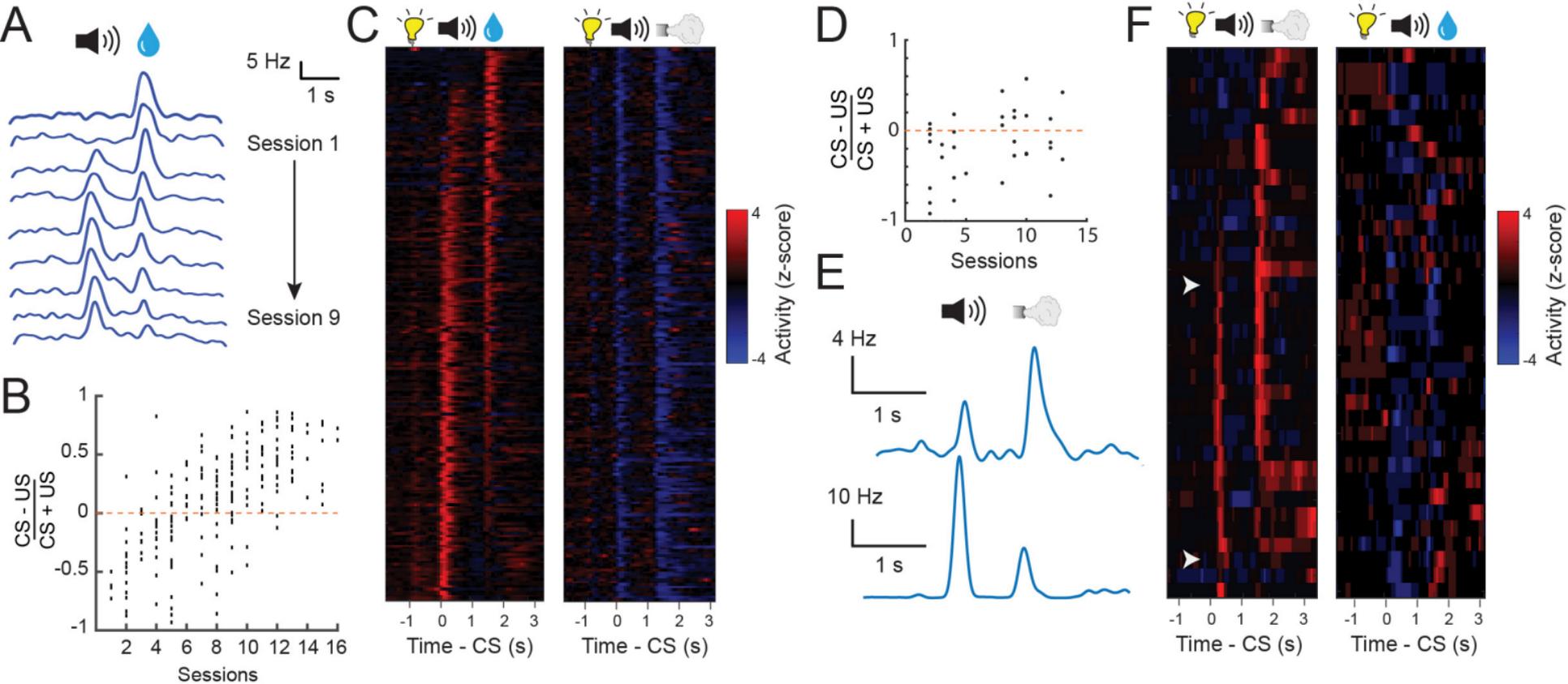


Figure 2

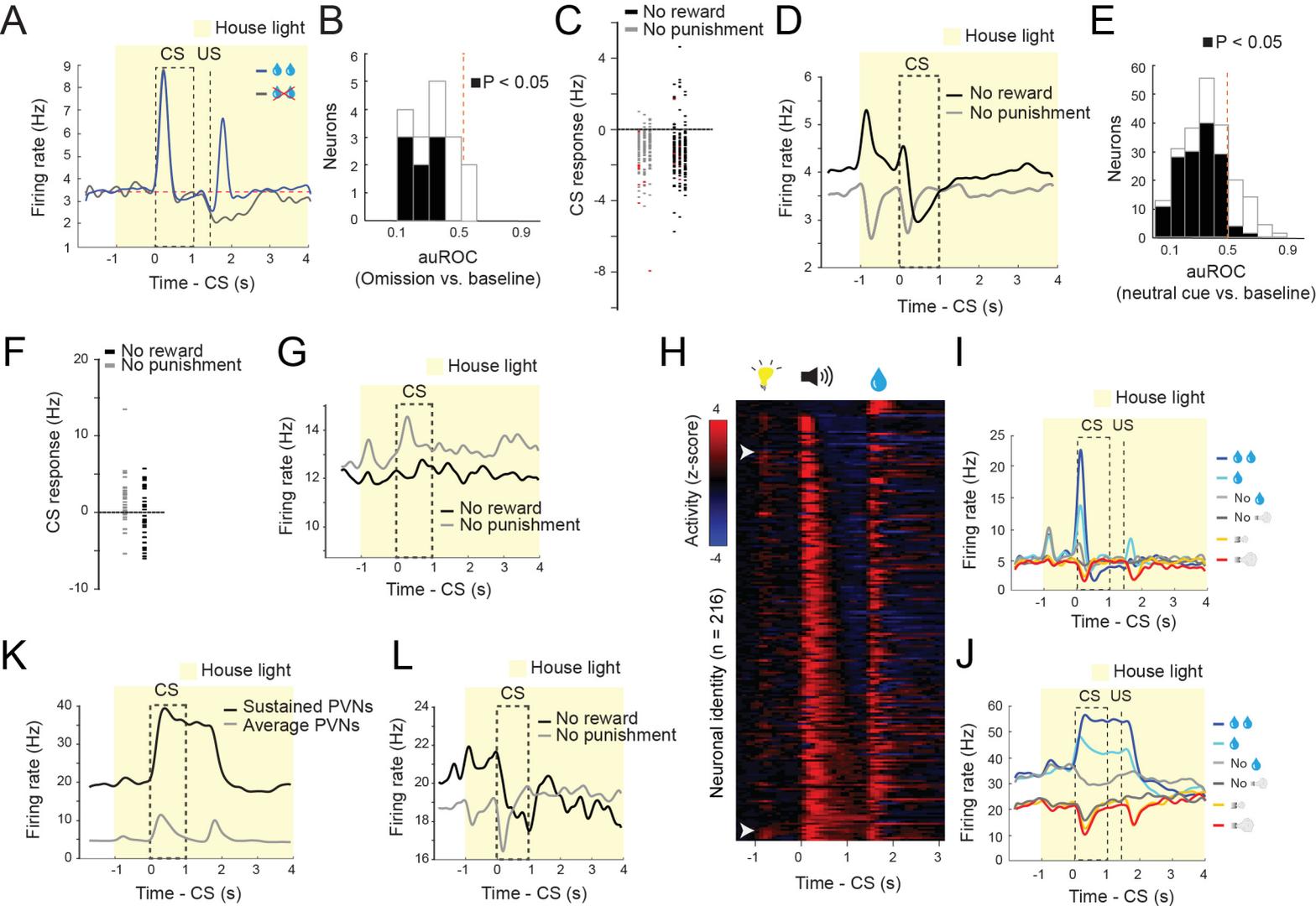


Figure 3

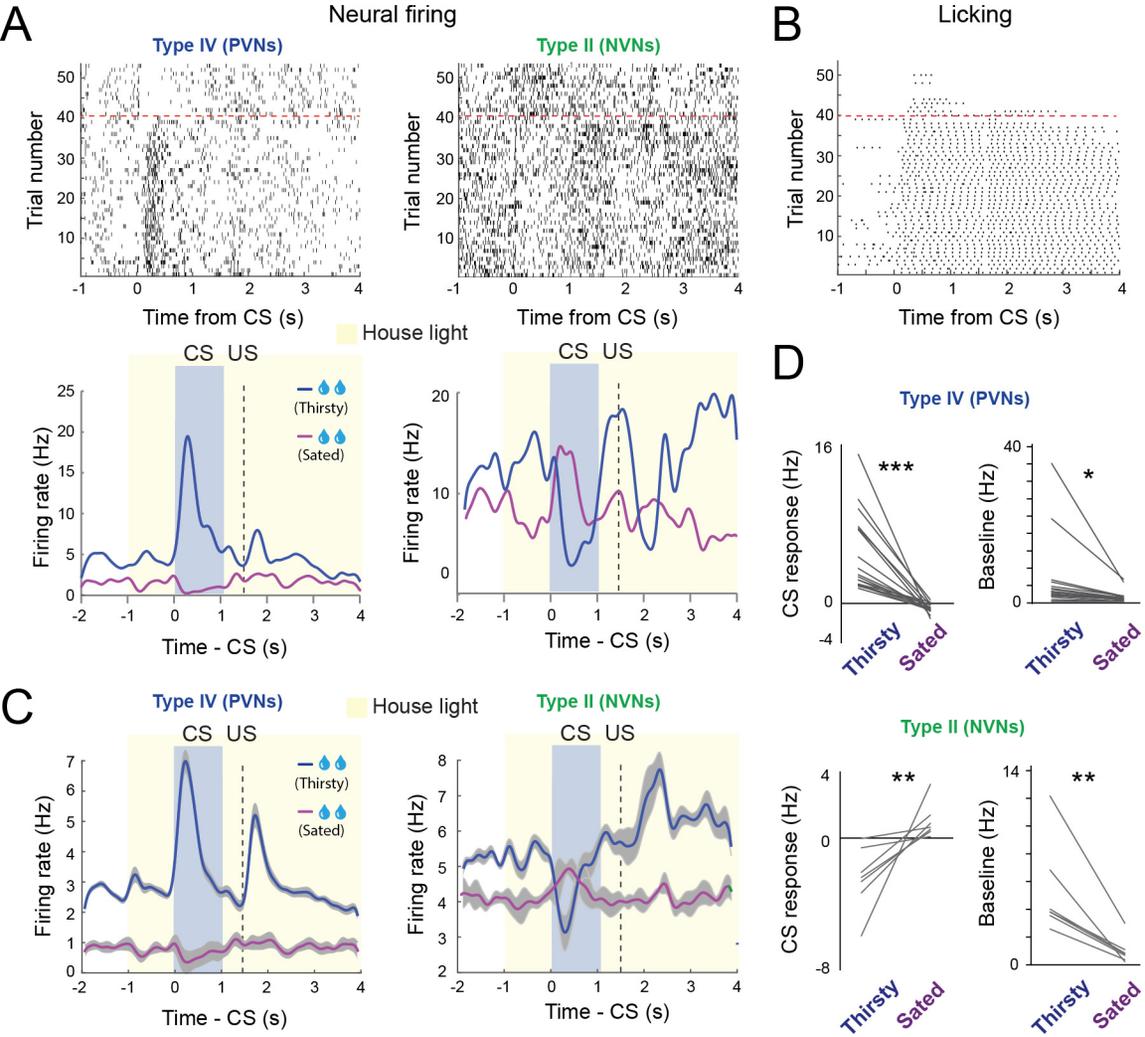


Figure 4

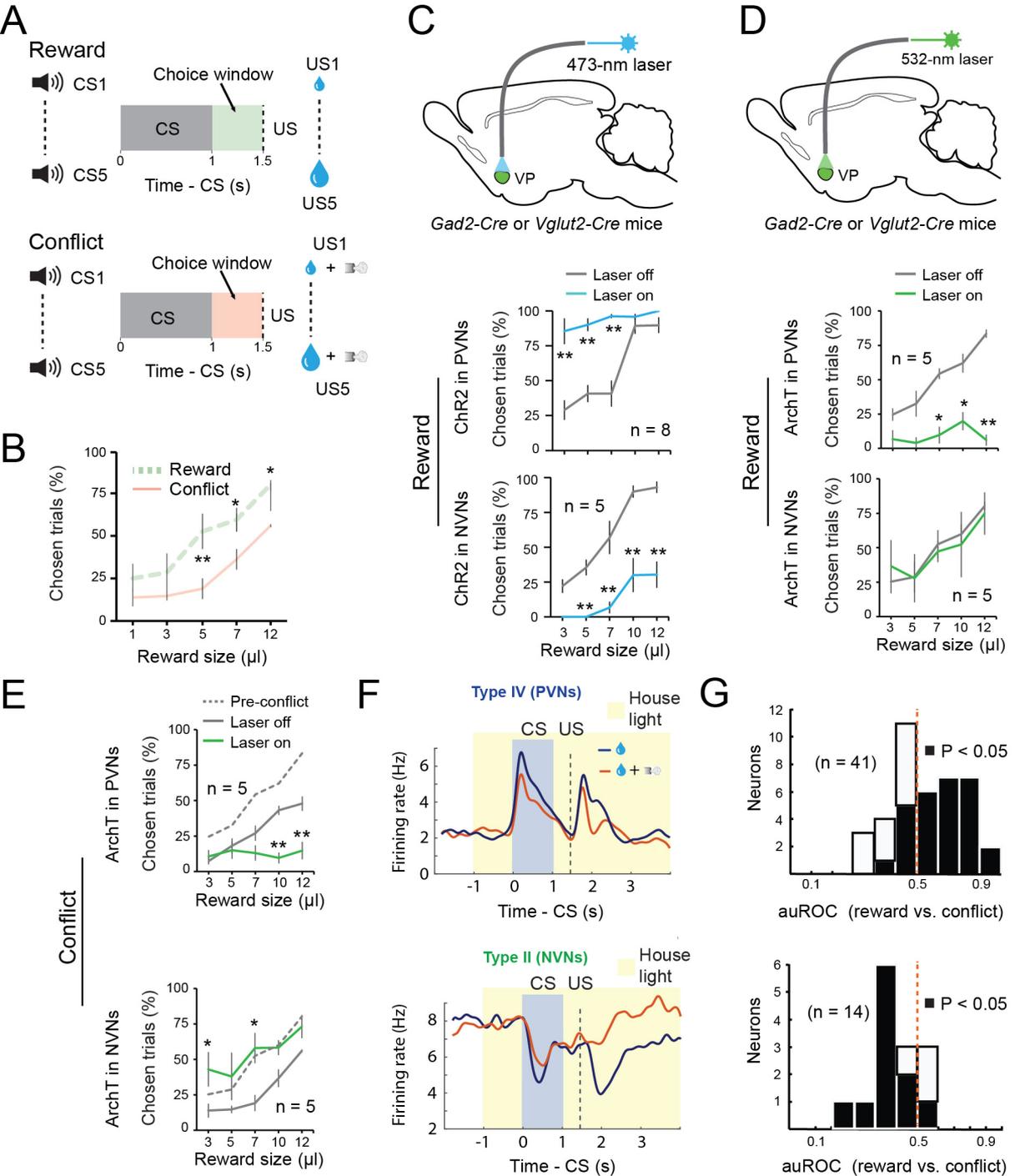


Figure 5

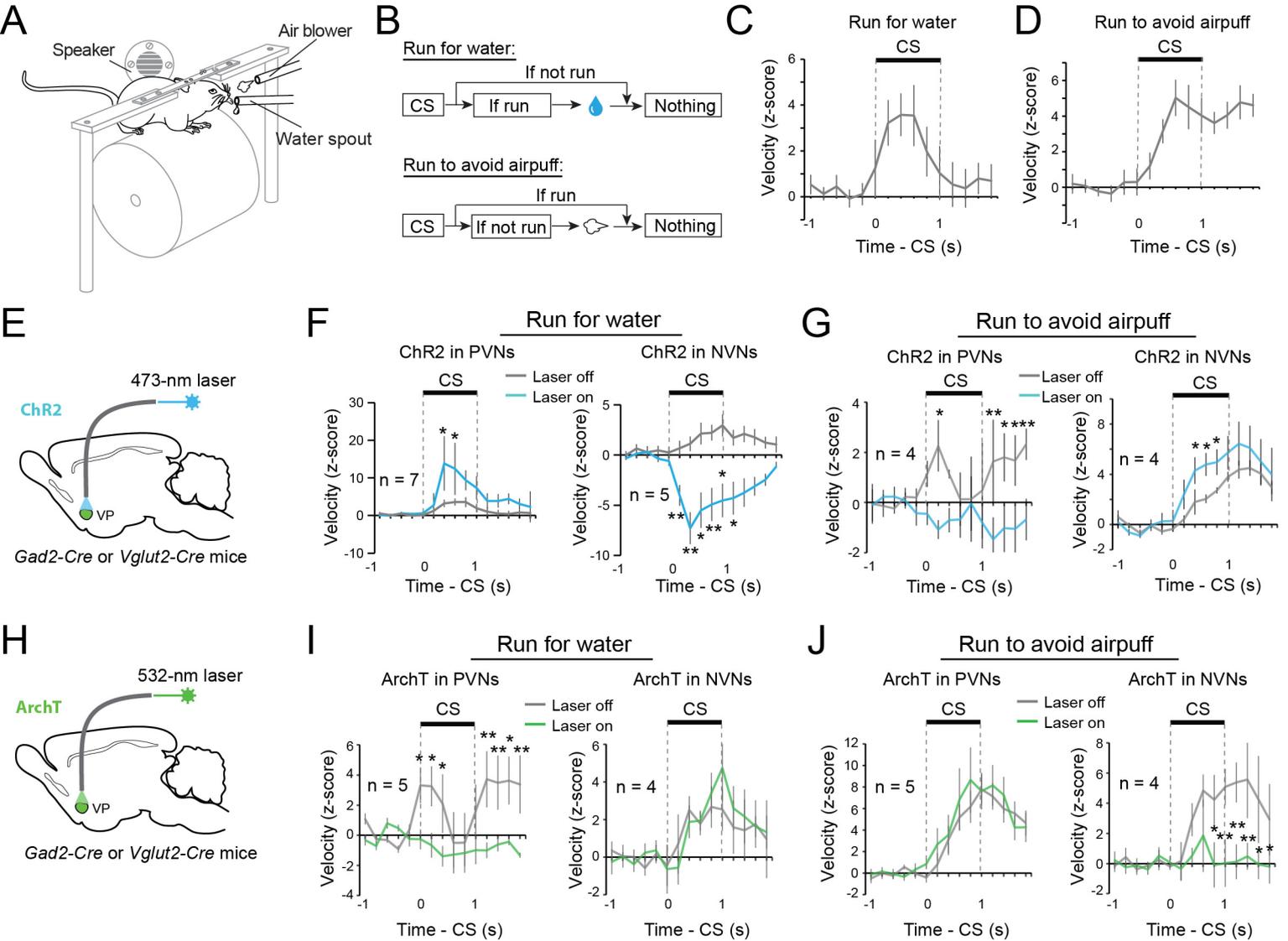


Figure 6

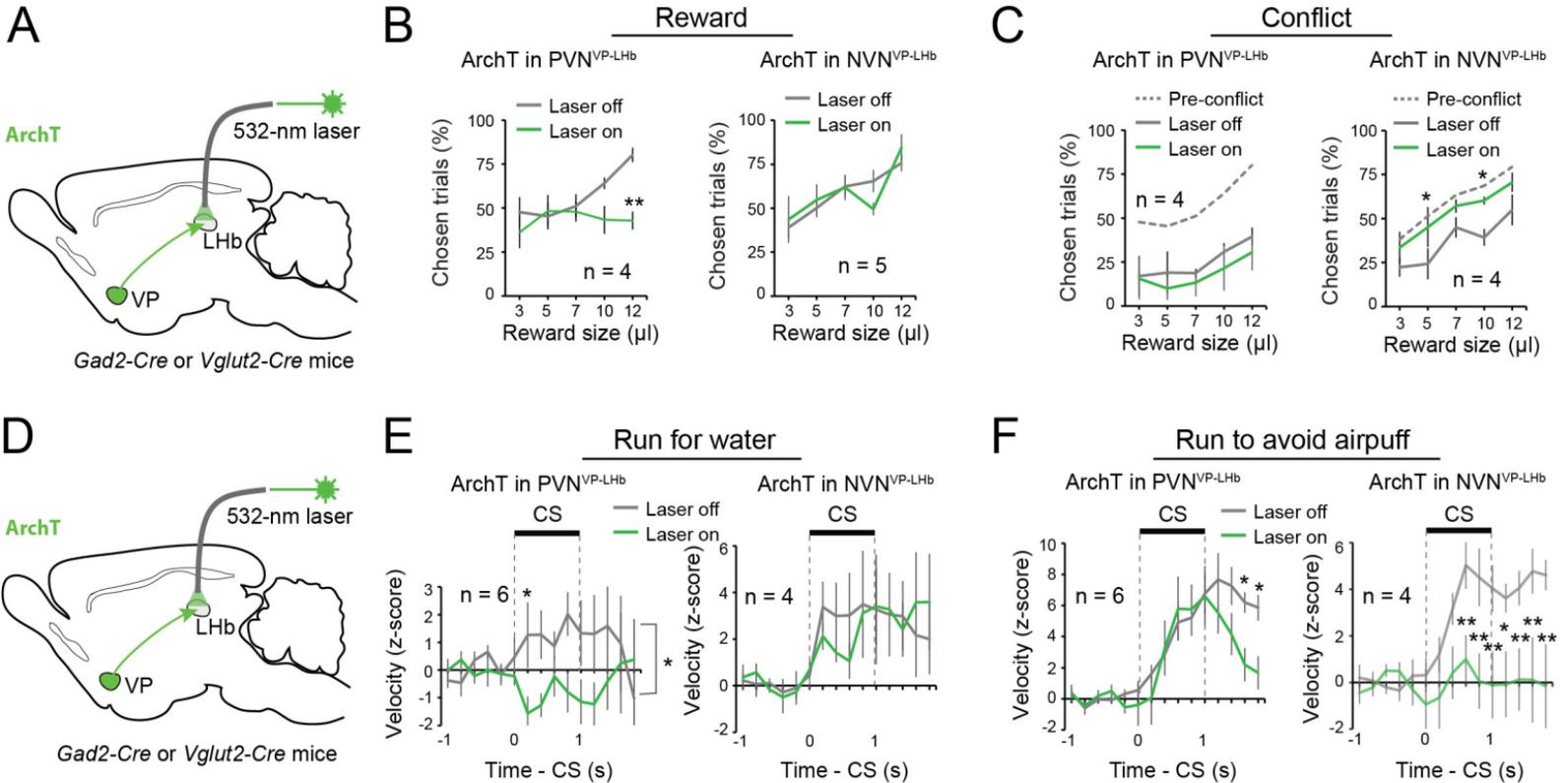
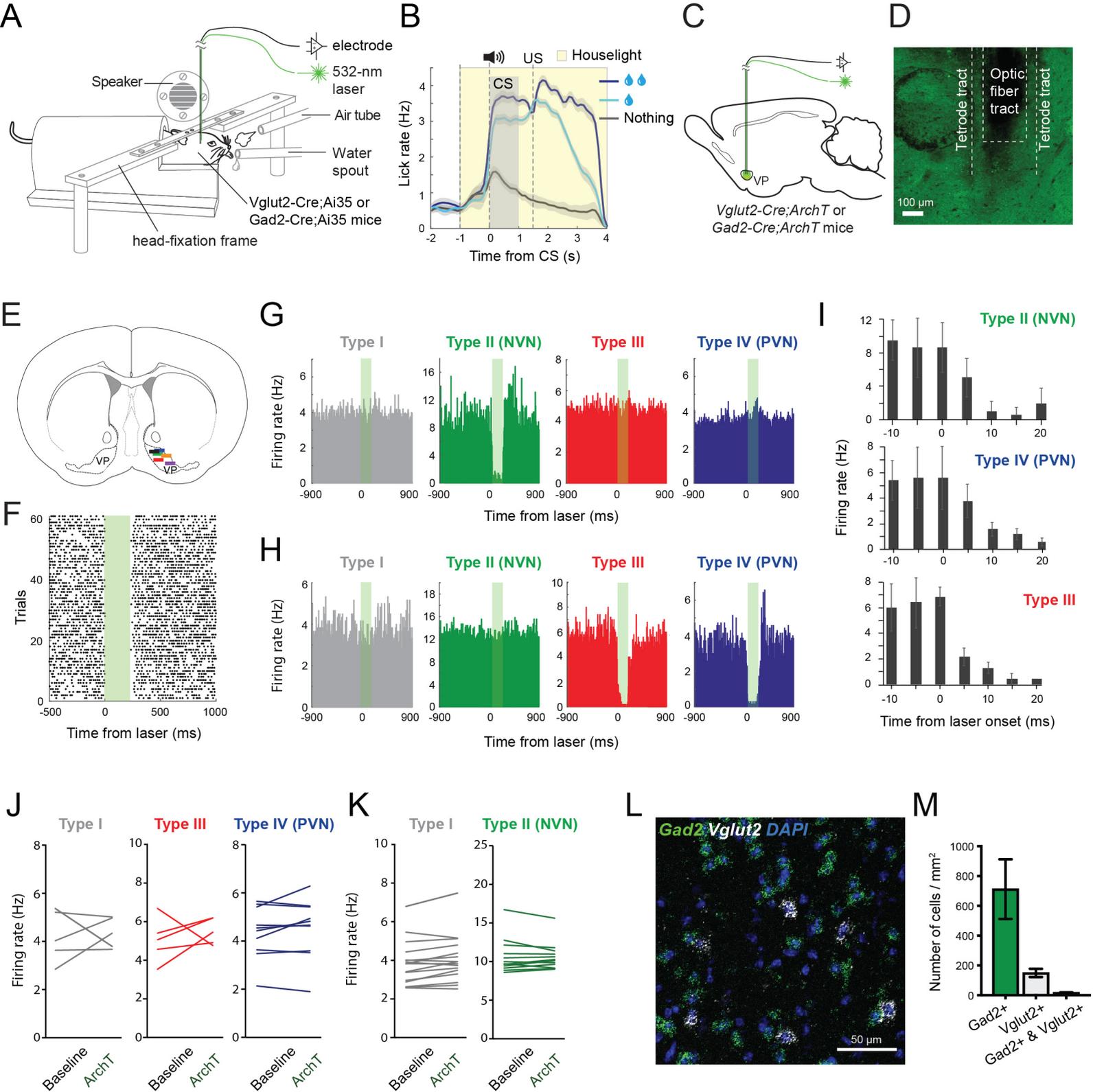
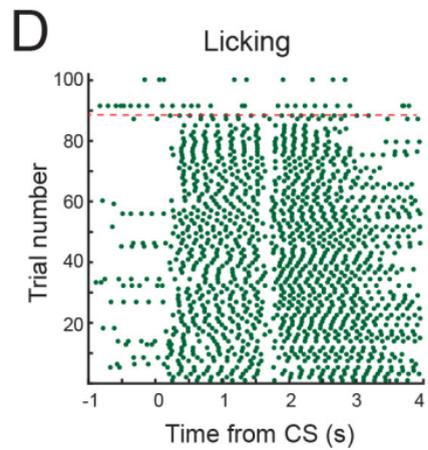
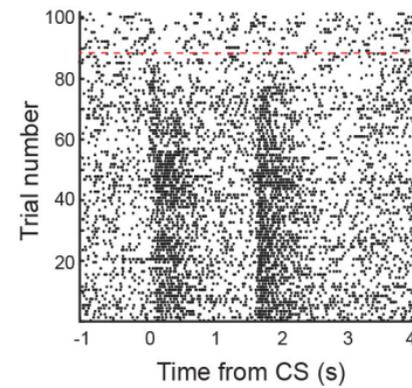
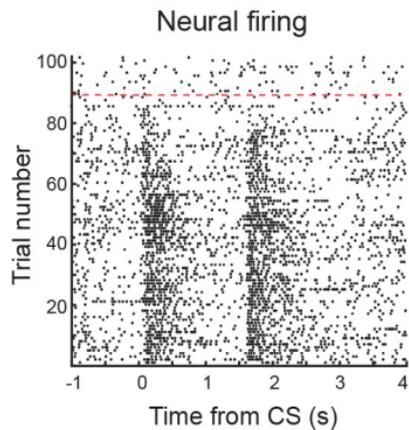
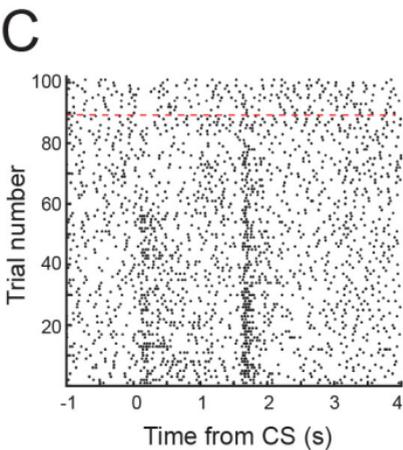
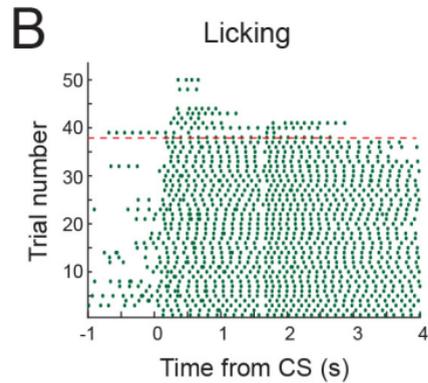
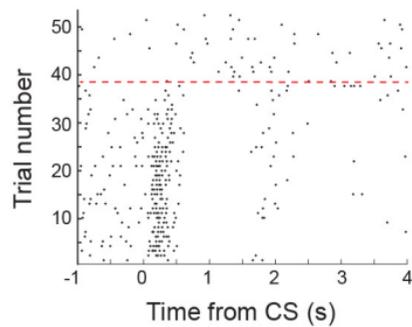
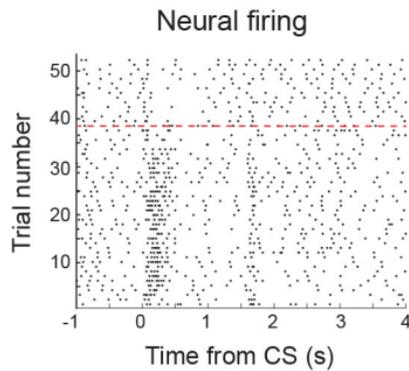
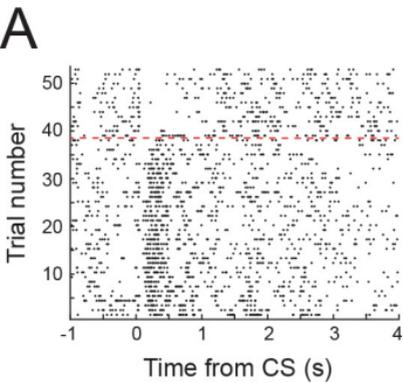


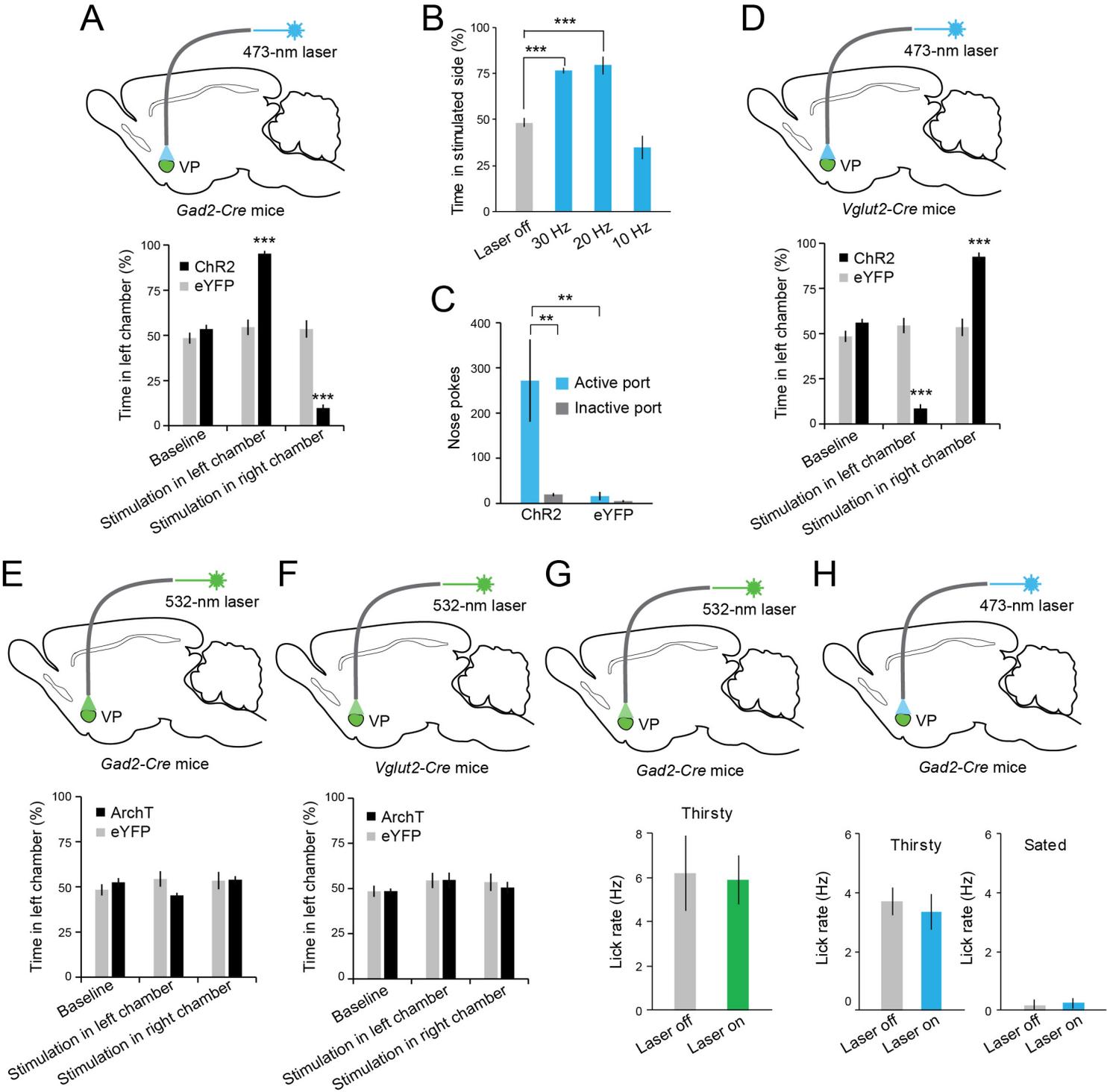
Figure 7

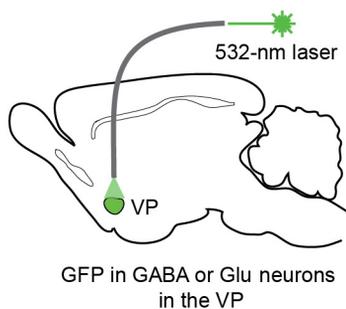
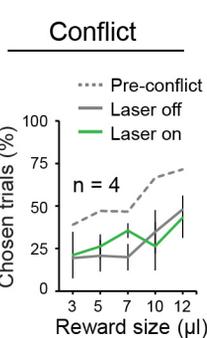
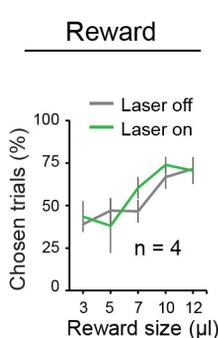
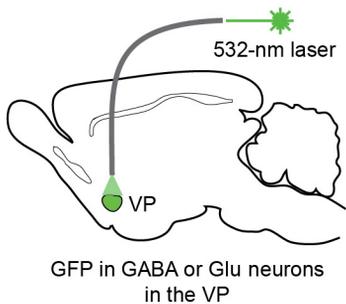
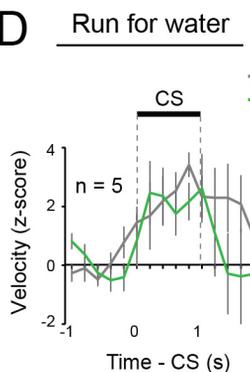
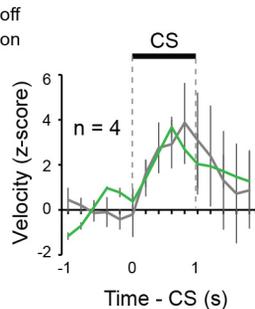
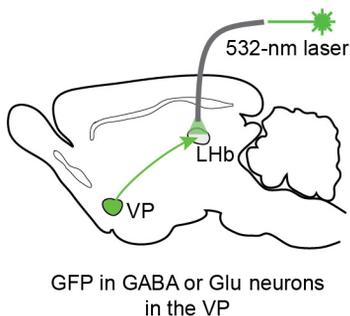
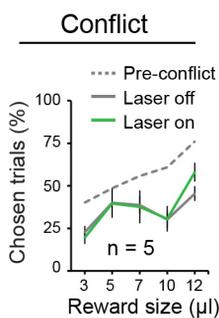
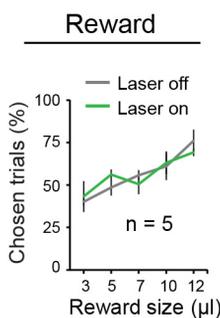
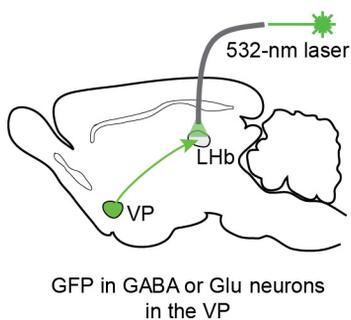
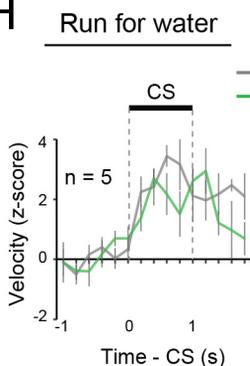
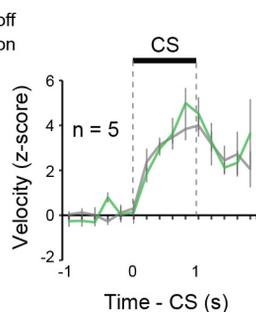


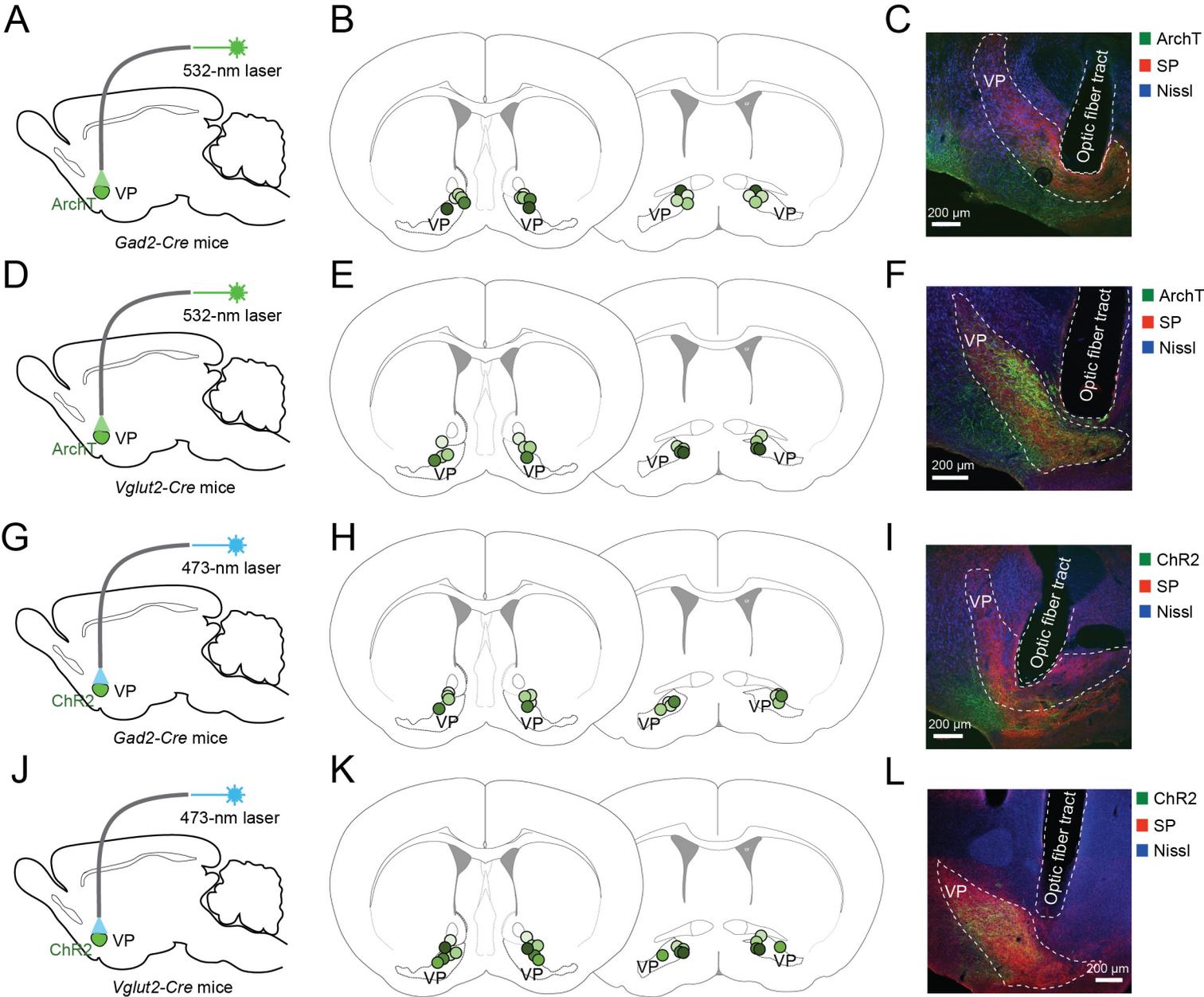
Supplementary figure 1



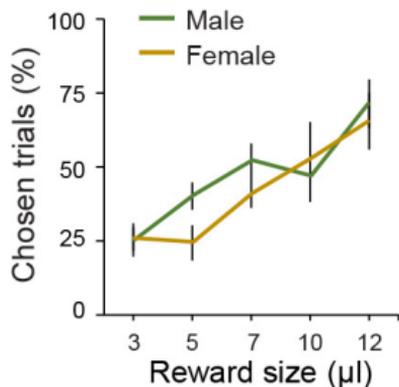
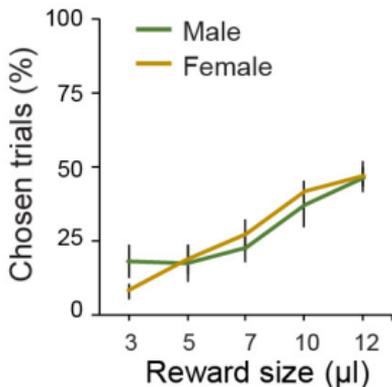
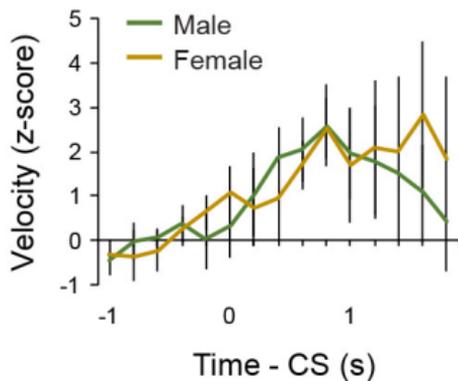
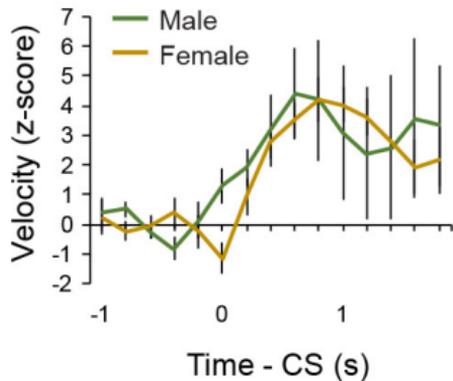
Supplementary figure 2

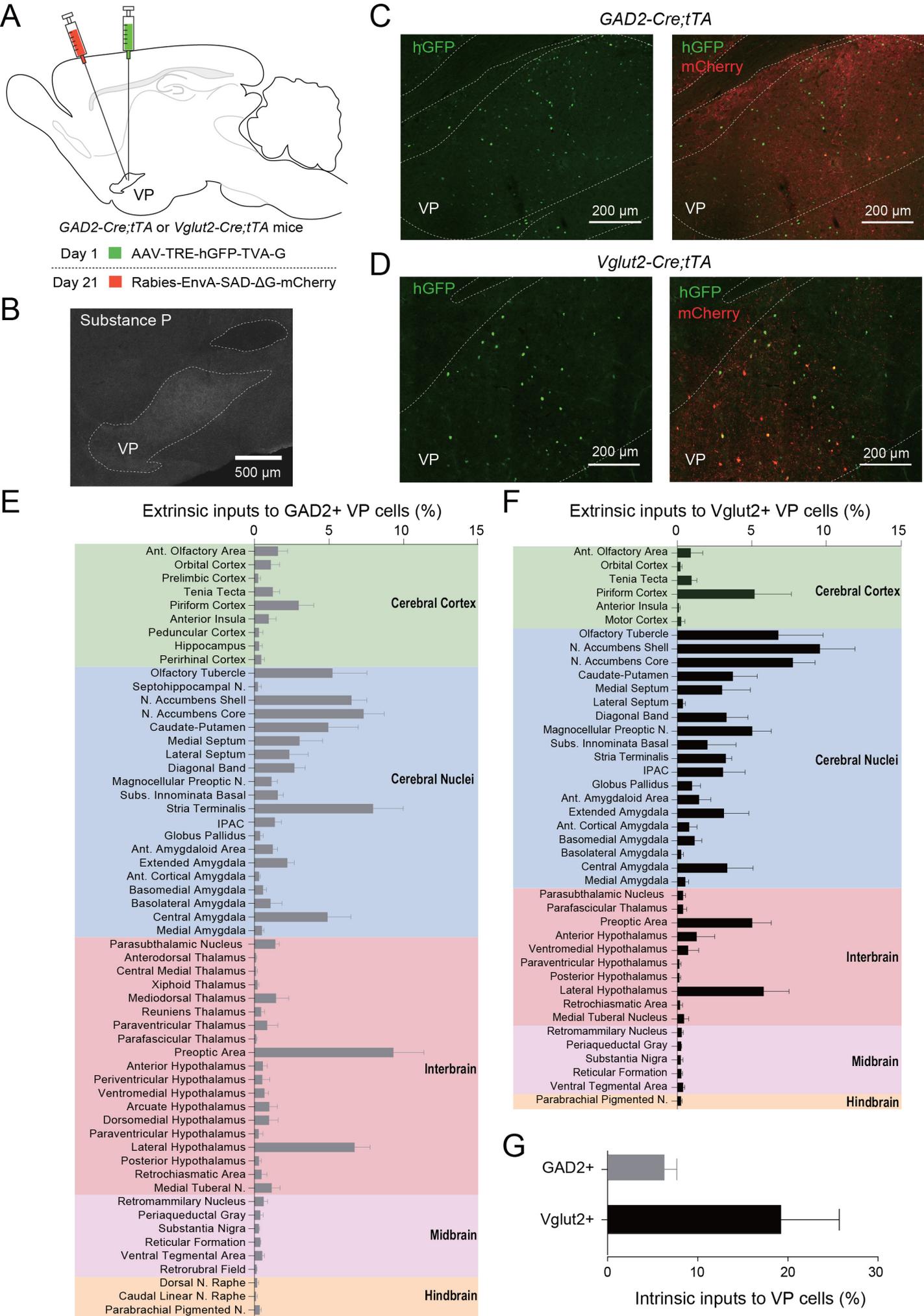


A**B****C****D****Run to avoid airpuff****E****F****G****H****Run to avoid airpuff**

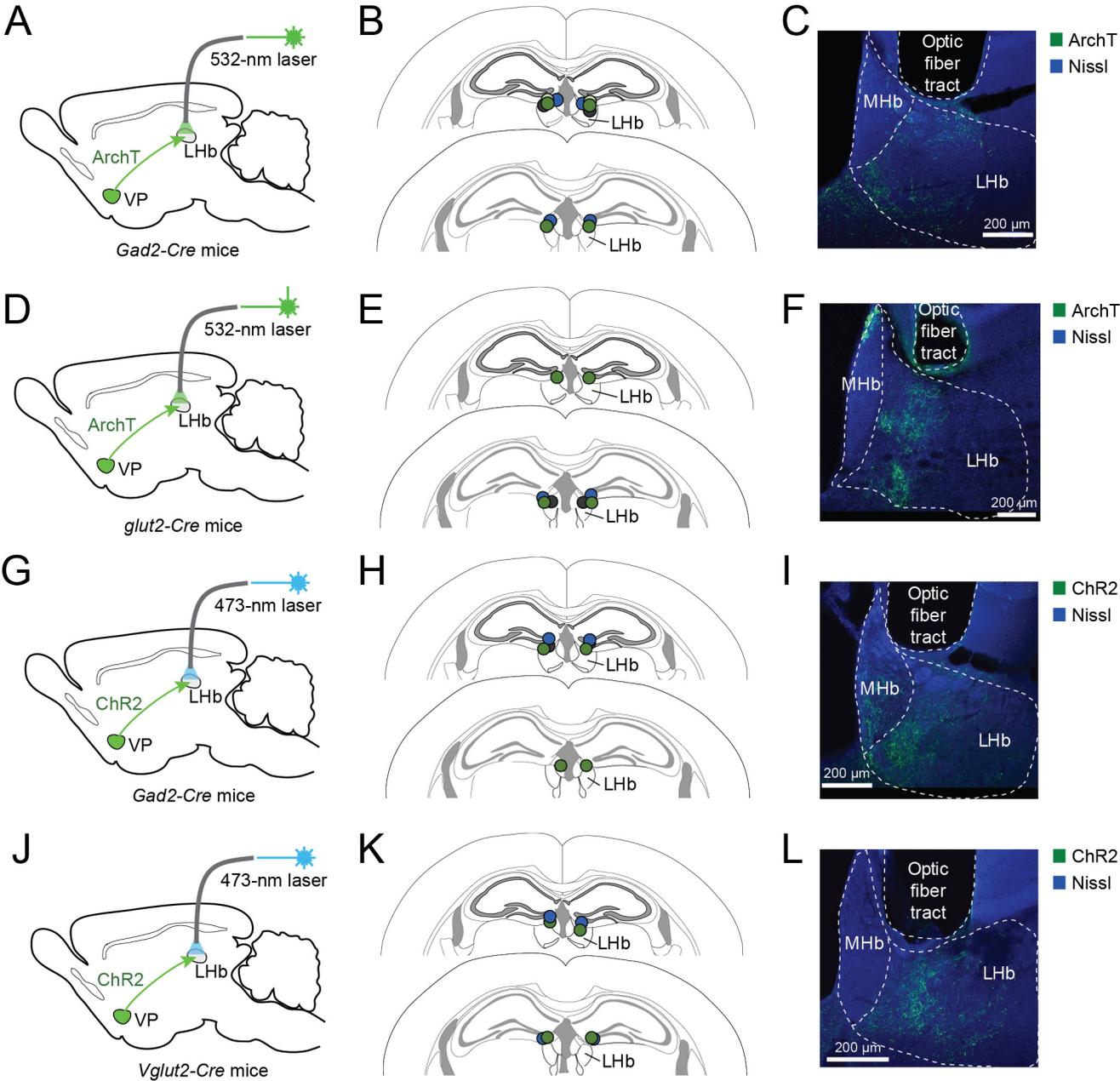


Supplementary figure 5

A**Reward****B****Conflict****C****Run for water****D****Run to avoid airpuff****Supplementary figure 6**

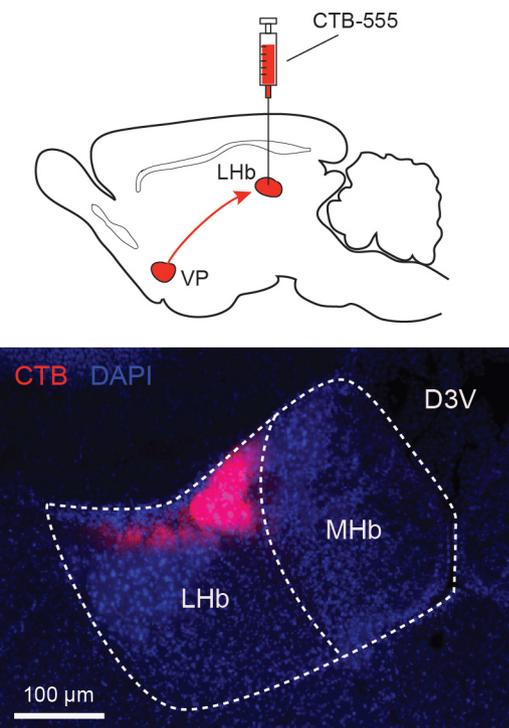


Supplementary figure 7

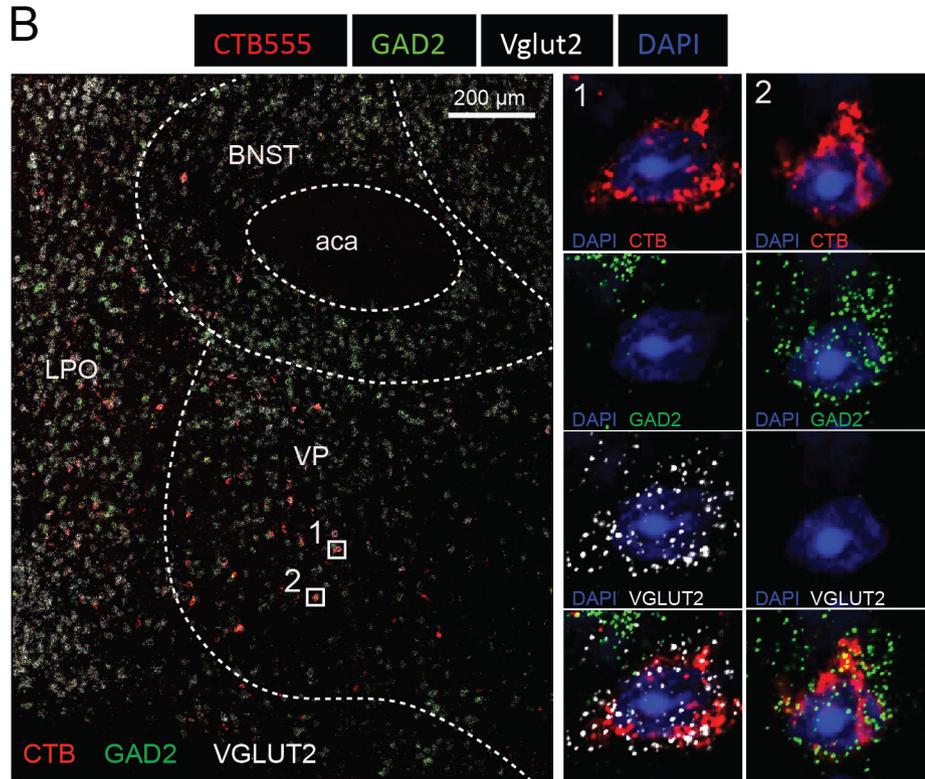


Supplementary figure 8

A



B

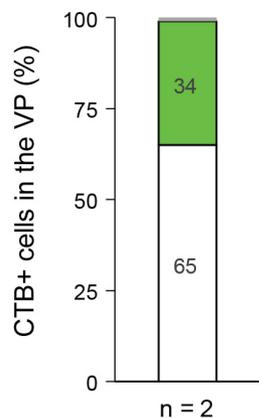


C

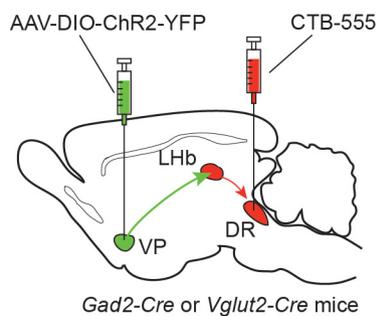
■ GAD2+

□ VGLUT2+

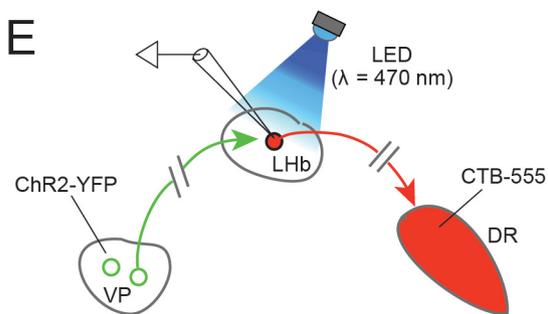
■ Unassigned



D

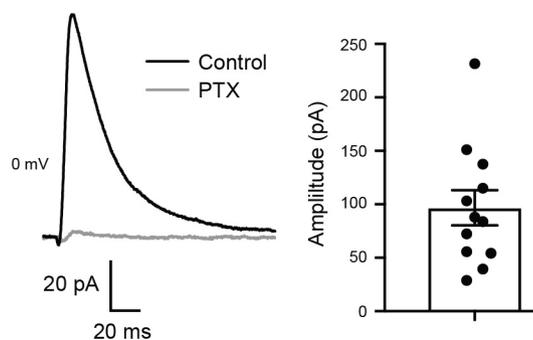


E



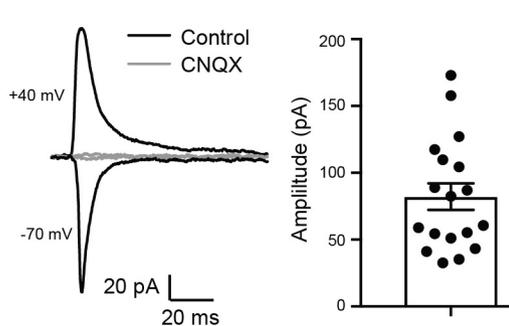
F

Inputs from GABAergic VP neurons

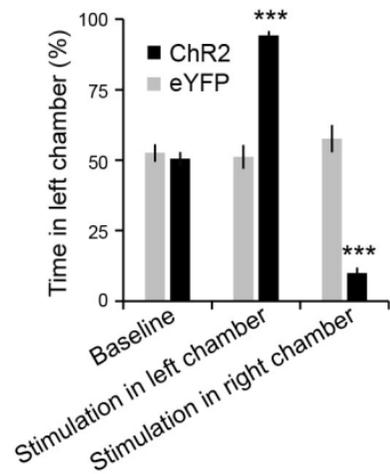
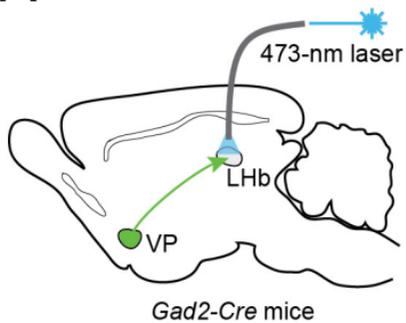


G

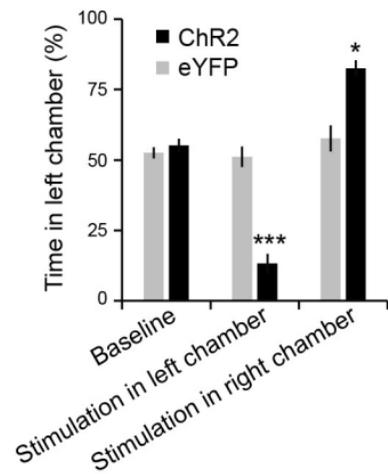
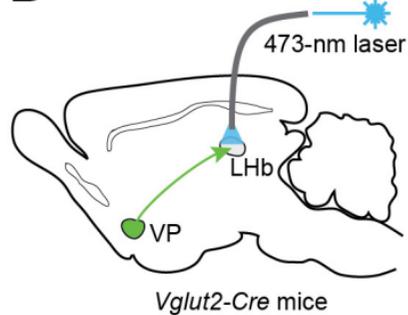
Inputs from glutamatergic VP neurons



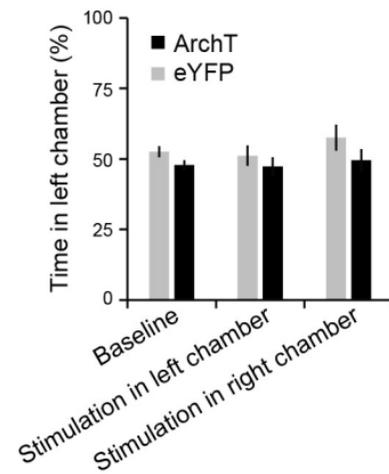
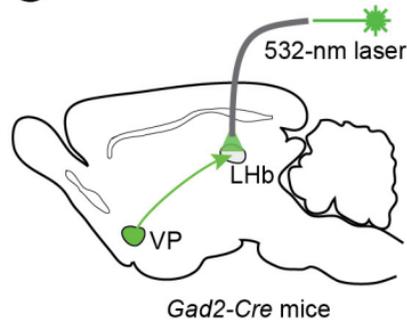
A



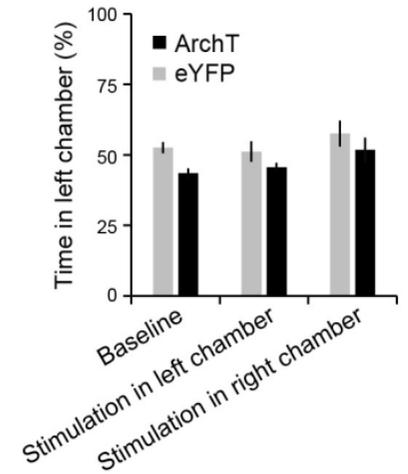
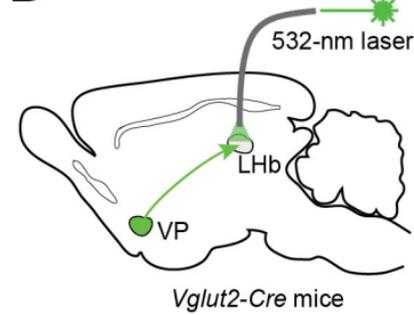
B



C



D



Supplementary figure 10