1 Title

2 Reward timing and its expression by inhibitory interneurons in the mouse primary visual cortex

3

4 Author Names

5 Kevin J Monk[1-2], Simon Allard[1-2], and Marshall G Hussain Shuler[1-2]

6

7 Affiliations and Addresses

8 (1) The Solomon H. Snyder Department of Neuroscience, The Johns Hopkins University

9   School of Medicine, 725 N Wolfe Street, Baltimore, MD 21205

10 (2) Kavli Neuroscience Discovery Institute, The Johns Hopkins University School of Medicine,

11   725 N Wolfe Street, Baltimore, MD 21205

12

13 Running Head

14 Reward timing in mouse visual cortex

15

16 Corresponding Author Address

17 Marshall G Hussain Shuler

18 725 N. Wolfe Street

19 WBSB 914

20 Baltimore, MD 21205

21 shuler@jhmi.edu

22

23 Abstract

24   Primary sensory cortex has historically been studied as a low-level feature detector, but has more

25   recently been implicated in many higher-level cognitive functions. For instance, after an animal

26   learns that a light predicts water at a fixed delay, neurons in primary visual cortex (V1) can

27   produce "reward timing activity" (i.e., spike modulation of various forms that relate the interval

28   between the visual stimulus and expected reward). The manner by which V1 produces these

29   representations is unknown. Here, we combine behavior, *in vivo* electrophysiology, and

30   optogenetics to investigate the characteristics of and circuit mechanisms underlying V1 reward

31   timing in the head-fixed mouse. We find that reward timing activity is present in mouse V1, that

32   inhibitory interneurons participate in reward timing, and that these representations are consistent

33   with a theorized network architecture. Together, these results deepen our understanding of V1

34   reward timing and the manner by which it is produced.

35

36   Introduction

37   Primary sensory cortex is classically regarded as a low-level feature detector providing simple

38   representations for higher-order areas. In the visual system, representations in early areas relate

39   to simple features, and through the cortical hierarchy, these signals are transformed into complex

40   representations of the external world (Hubel and Wiesel, 1959, 1965; Felleman and Van Essen,

41   1991). Yet, it is becoming increasingly clear that representations in primary sensory areas are

42   updated when stimuli gain meaning through associative learning (McGann, 2015). Specifically,

43   this is seen in primary gustatory (Vincis and Fontanini, 2016), somatosensory (Wiest et al., 2010;

44   Gdalyahu et al., 2012; Pais-Vieira et al., 2013), auditory (Polley et al., 2004; Rutkowski and

45   Weinberger, 2005; Guo et al., 2019), and visual (Shuler and Bear, 2006; Gavornik and Bear,

46   2014a, 2014b; Makino and Komiyama, 2015; Goard et al., 2016; Goltstein et al., 2018) cortices,

47   and in the olfactory bulb (Kay and Laurent, 1999; Kass et al., 2013; Ross and Fletcher, 2018).

48   Previous work shows that individual neurons in rodent primary visual cortex (V1) express reward

49   timing activity (Shuler and Bear, 2006; Chubykin et al., 2013; Liu et al., 2015; Zold and Hussain

50   Shuler, 2015). Reward timing activity is a representation of time between a visual stimulus and a

51   reward, expressed as one of three forms: 1) a sustained increase (SI) or 2) sustained decrease

52   (SD) of activity until the time of expected reward, or 3) a peak (PK) of activity around the time of

53   the expected reward. Prior studies advance V1 as a substrate in the learning of this timing activity

54   and implicate acetylcholine as a reinforcing signal: lesions of cholinergic axons within V1 block

55   the ability for V1 to learn reward timing activity (Chubykin et al., 2013); pairing visual stimuli with

56   local activation of cholinergic fibers in V1 mimics behaviorally-conditioned reward timing (Liu et

57   al., 2015); V1 timing activity correlates with timing behavior and perturbation of this activity lawfully

58   shifts timing behavior (Namboodiri et al., 2015; Levy et al., 2017).

59

60   These observations implicate V1 as a substrate for learning reward timing activity—thought to

61   occur through a reinforcement learning process (Hussain Shuler, 2016)—but it remains unclear

62   how V1 circuitry produces reward timing activity with its various response forms. A computational

63   model proposes a solution with a specific connectivity motif within a network of excitatory and

64   inhibitory cells (Huertas et al., 2015). This network architecture has two implications: (1) inhibitory

65   interneurons represent time predominantly as the SI form and (2) neurons inhibited by

66   interneurons represent time predominantly as the SD or PK form. In this network, interneurons

67   are treated as one, monolithic group. However, V1 interneurons fall mainly into one of three

68   subpopulations expressing either parvalbumin (PV), somatostatin (SOM), or vasoactive intestinal

69   polypeptide (VIP) (Xu et al., 2010; Tremblay et al., 2016). Each are unique in their connectivity

70   patterns (Pfeffer et al., 2013) and are functionally distinct during stimulus representation in V1

71   (Atallah et al., 2012; Lee et al., 2012; Wilson et al., 2012). It is unknown if either of the model's

72   implications are borne out *in vivo* and, if so, how the diversity of interneuron subtypes intersects
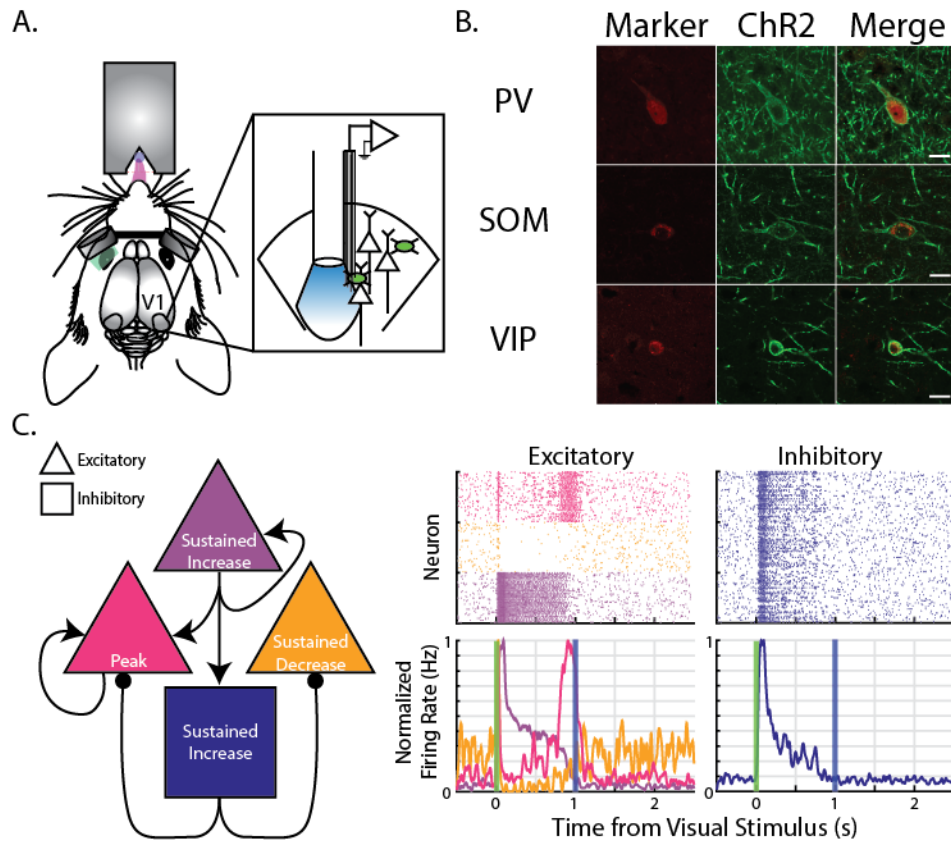
73   with these implications.

3

74 Here we enrich our understanding of V1 reward timing by investigating the manner by which

75 mouse V1 neurons produce reward timing activity, how different interneuron populations express

76 and aid in the production of this activity, and how well biological and computational data accord

77 with one another. In doing so, we find that V1 neurons express reward timing in a manner

78 consistent with a theorized network architecture and that PV+ interneurons fulfill the expectations

79 of the theorized inhibitory population.

80

81 Results

82 The means by which the primary visual cortex produces the various forms of reward timing

83 observed is unknown. We have investigated potential mechanisms through an in-depth

84 characterization of reward timing in the mouse primary visual cortex and how this reward timing

85 activity is expressed by inhibitory interneurons (Figure 1A for a recording schematic). Specifically,

86 we use mice which selectively express channelrhodopsin-2 (ChR2) in interneuron subpopulations

87 to determine how these different cell types produce and aid in the production of reward timing

88 activity (Figure 1B). Finally, with these data we sought to compare the activity of interneuron

89 populations with a proposed network architecture which replicates reward timing activity (Figure

90 1C, (Huertas et al., 2015)).

91

Figure 1: Summary of recording strategy and motivation to investigate inhibitory interneurons. (A) Cartoon schematic showing recording design. Head-fixed mice were required to lick following presentation of visual stimuli to receive a water reward after a fixed delay. Simultaneously, electrophysiological recordings were made in infragranular primary visual cortex. Depending on genotype, mice expressed channelrhodopsin-2 in inhibitory interneurons (schematized in green). (B) Channelrhodopsin-2 (ChR2) expression co-localizes with expected interneuron marker given animal's genotype. Specifically, in PV-ChR2 animals, there was co-expression of PV and ChR2 (top row), in SOM-ChR2 animals there was co-expression of SOM and ChR2 (middle row), and in VIP-ChR2 animals, there was co-expression of VIP and ChR2 (bottom row). Scale bars in right column represent 15μm. (C) Left: Schematic of the theorized network architecture. This architecture is composed of three populations of excitatory cells (triangles) and one population of inhibitory cells (square). Right: Following training of the model, simulated excitatory neurons produce reward timing responses in one of three forms (left), and simulated inhibitory units produce Sustained Increase reward timing responses. Schematic on left adapted from Huertas et al., 2015; plots on right simulated as in Huertas et al., 2015.

92

93

94   *Head-Fixed Mice Associate a Visual Stimulus with a Delayed Reward*

95   Prior to investigating the circuit mechanisms by which primary visual cortex (V1) produces reward

96   timing, we first established its existence within the head-fixed mouse preparation. Mice were

97   trained to associate a water reward with visual stimuli (see Methods). Briefly, head-restrained

98   mice received a 100ms visual stimulus delivered to the left or right eye with equal probability (Cue

99   1 and Cue 2, respectively) via head-mounted goggles and received water from a lick port placed

100 within reach of the tongue. Trials were initiated after a lapse of time comprising a randomly

101 selected interval and a second random interval less than the ITI during which the animal must not

102 lick (a "lick lockout" interval). If an animal licked during this lick lockout, the lockout timer would

103 restart. Such an ITI encourages mice to cease licking and to wait for the onset of the next trial.

104 Upon the initiation of a trial, animals received a monocular visual stimulus delivered to the left or

105 right eye with equal probability, after which the animal was required to make at least one lick

106 within the subsequent delay period so that reward could be delivered at the end of the delay. On

107 half of these trials, if the animals met this behavioral requirement, they received a small water

108 reward (~2µL) at the end of the conditioned interval (so-called "paired" trials). On the other half of

109 these trials, regardless of lick behavior, reward was withheld ("catch" trials). On 20% of trials,

110 neither a visual stimulus nor a reward were delivered although the intertrial interval and lick lockout

111 periods expired successfully; these trials are referred to as "sham" trials and are used to verify

112 that animals are using visual stimuli to guide licking behavior (as opposed to timing lick bouts from

113 events other than a visual cue). The delay time used was the same for both visual stimuli within

114 a recording session and varied across days, as follows: on short delay sessions the delay time

115 was 1 second following visual stimulus offset, and on long delay sessions the delay time was 1.5

116 seconds following visual stimulus offset. A task schematic is shown in Figure 2A and behavior

117 from an example session is shown in Figure 2B. Regardless of trial type, trials in which the animal

118 made a lick during the delay window are defined as Hit trials and trials in which the animal did not

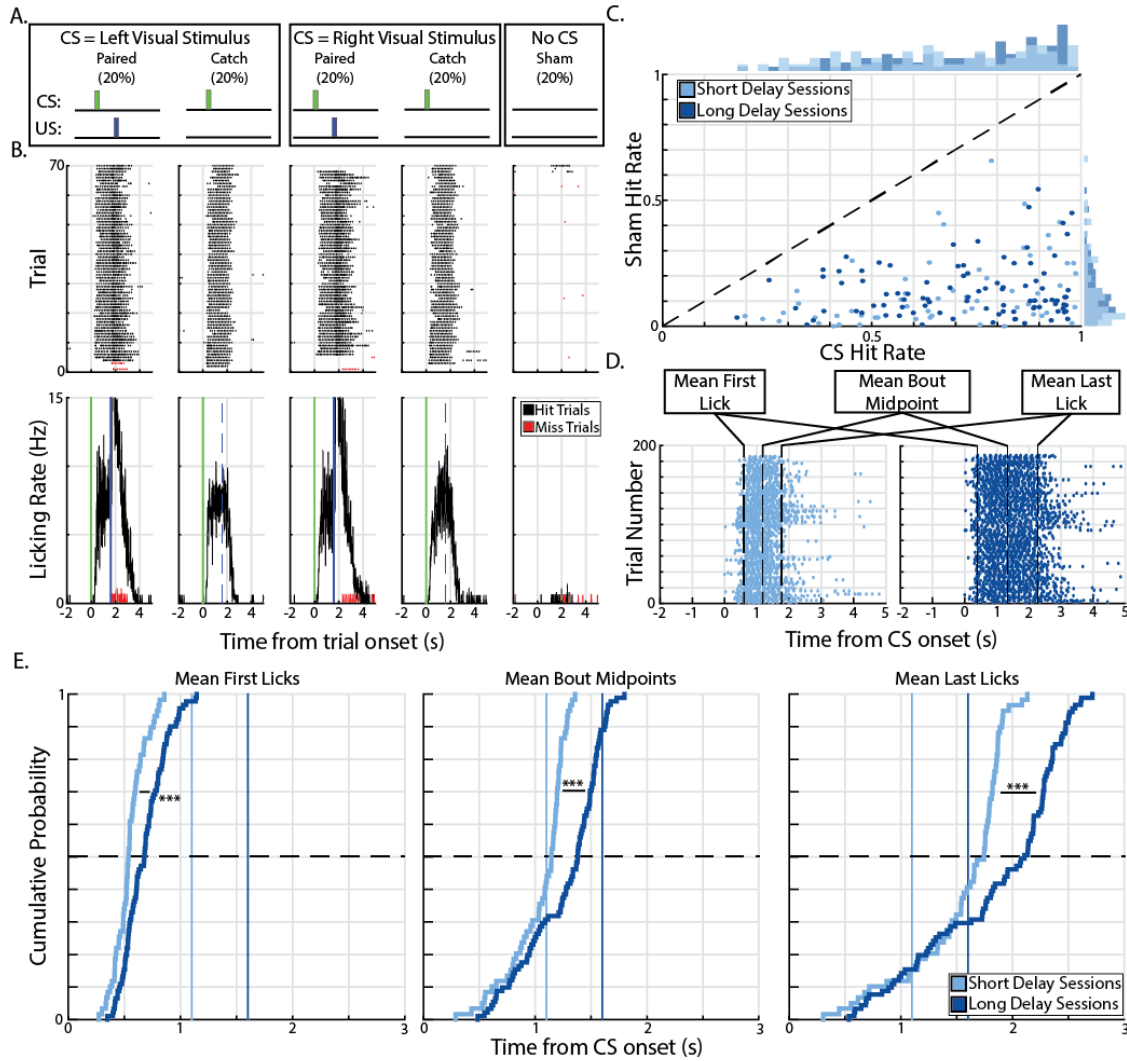119 lick were referred to as Miss trials. All data presented here, unless otherwise noted, are from

6

Figure 2: Head-fixed mice learn the reward timing task. (A) Animals were trained that a monocular visual stimulus (the conditioned stimulus, CS) delivered to the left or right eye predicted a water reward (the unconditioned stimulus, US) at a fixed delay. Upon receipt of the CS, animals were required to lick during the delay period so that reward could be delivered at the end of the short or long delay. Trials in which the reward is delivered are Paired trials and trials in which the reward is withheld are Catch trials. Trials in which no CS nor US were delivered are Sham trials and are used to determine the ability of the animal to use the visual stimulus to guide licking behavior. (B) Example licking behavior for a single animal during a single session of the task for all trial types with respect to the onset of the CS (time 0, green line). Blue lines represent time of reward delivery (solid line representing the receipt of reward on paired trials and dashed line representing expected time of reward on catch trials). Top: raster plot of licking behavior where black dots are licks recorded during "Hit" trials (i.e., trials in which the animal licked during the delay window) and red dots are licks recorded during "Miss" trials (i.e., trials in which the animal did not lick in the delay window). Bottom: PSTH's calculated from licking activity, color scheme as noted above. (C) Scatter plot of session hit rates (probability of licking on a given trial) for all trials in which the CS is delivered (CS Hit Rate) and for all trials in which the CS was withheld (Sham Hit Rate) for short delay sessions (light blue) and long delay sessions (dark blue). Histograms of hit rates for each trial category shown in margins and unity line is the black, dashed line. (D) Behavioral measurements (labelled, black dashed lines) for example licking behavior during Catch+Hit trials from a short delay session (left, light blue raster) and a long delay session (right, dark blue raster). (E) Cumulative distribution plots for the three behavioral measurements measured during Catch+Hit trials during short delay sessions (light blue) and long delay sessions (dark blue). Vertical lines represent reward time and horizontal line indicates the median values. *** - p < 0.001, Wilcoxon rank-sum test.

120

7

121   Catch+Hit trials (i.e., trials in which the animal received a visual stimulus, licked during the delay,

122   and did not receive a reward).

123

124   As expected, animals showed a high probability of licking in the delay period on trials where a

125   visual stimulus was delivered ("CS trials") and a low probability of licking during the sham trials

126   (70.67% and 14.04%, respectively, Figure 2B and 2C). There is a significant effect of trial type

127   (i.e., CS trial vs sham trial, χ^2(1, 286) = 464.11, p = 7.83 x 10^{-62}, Kruskal-Wallis test) on the

128   probability that an animal licks while there is neither a significant effect of session number nor a

129   significant interaction (Session Number: χ^2(8, 286) = 0.51, p = 0.85; Interaction: χ^2 (8, 286) = 1.26,

130   p = 0.26, Kruskal-Wallis test). These results demonstrate that animals lick in response to reward-

131   predicting visual stimulation and that their behavior had reached asymptotic performance at the

132   time of recording.

133

134   We next addressed whether the animals time their behavioral response. To quantify the timing of

135   the licking behavior, we made three measurements: the time of the first lick in a bout, the time of

136   the last lick in a bout, and the mean time between the first and last lick in a bout (Bout Midpoint).

137   The Bout Midpoint is derived from the initiation and cessation of licking, and so is not an

138   independent measure. Rather, its inclusion is simply to determine whether the centering of lick

139   bouts is in good accordance with the expected time of reward. These values were recorded across

140   trials and an average of these values were calculated for a given trial type on a given day (Figure

141   2D for example sessions). When we compare these values, we find that the lick initiation and

142   cessation times (and, consequently, the Bout Midpoint) are significantly smaller for short delay

143   sessions compared to long delay sessions (Figure 2E, Mean First Licks: Z = -6.09, p = 1.11 x 10^{-9};

144   Mean Bout Midpoints: Z = -6.71, p = 2.01 x 10^{-11}; Mean Last Licks: Z = -5.73, p = 9.89 x 10^{-9},

145   Wilcoxon rank-sum test) indicating that animals adapt their licking behavior based on the

146   expected time to reward.

147   *Neurons in Primary Visual Cortex of the Head-Fixed Mouse Express Reward Timing Activity*

148   These behavioral data indicate that animals express an internal sense of the time interval between

149   the visual stimulus and the water reward. To determine what, if any, neural representation of time

150   was present in V1, we recorded single unit activity bilaterally during behavioral sessions. Previous

151   work from our lab has shown that, in similar tasks, neurons in V1 of freely-moving rats and mice

152   represent the time interval to an expected reward in one of three forms: a sustained increase (SI)

153   or sustained decrease (SD) of activity until the time of reward, or as a peak (PK) of activity around

154   the time of reward (Shuler and Bear, 2006; Chubykin et al., 2013; Liu et al., 2015). These

155   response forms were also observed here in the head-fixed mouse (Figure 3A). Using these

156   response forms, we manually classified in a blinded fashion the peristimulus time histograms

157   (PSTHs) of neurons for both Cue 1 and Cue 2 Catch+Hit trials (i.e., trials in which the animal

158   received a visual stimulus, licked during the delay window, and did not receive a water reward at

159   the end of that delay; see Methods). PSTHs created from Sham+Hit trials (that is, trials in which

160   neither CS nor US was delivered, but had licks within the delay window) were also blindly

161   classified as a control. Neurons could be classified as responsive during any of these trial types;

162   as such, we began our analyses by quantifying "neural records" (i.e., a given pattern of activity a

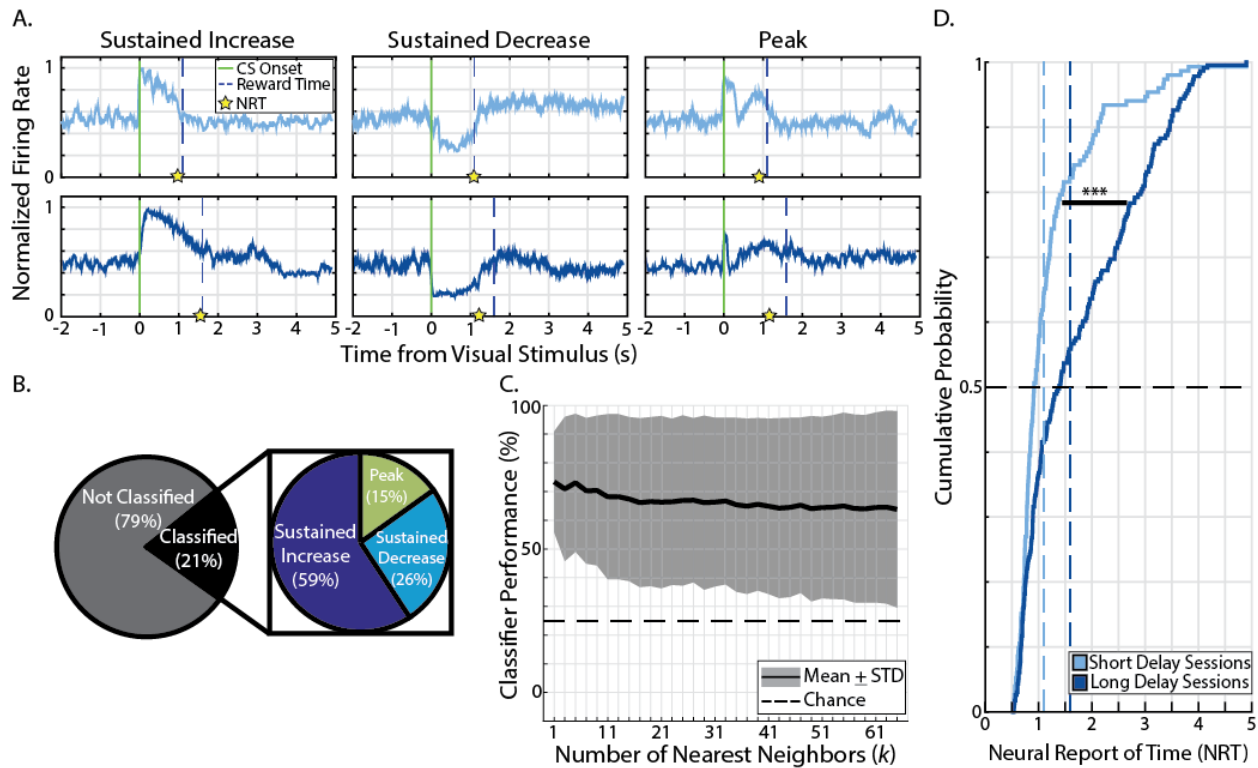163   neuron produced during a trial type).

Figure 3: Neurons in primary visual cortex express reward timing activity in three forms. (A) Six example neurons representing the three forms of reward timing: Sustained Increase (left), Sustained Decrease (middle), and Peak (right) recorded from short delay sessions (light blue, top row) or long delay sessions (dark blue, bottom row). Activity is recorded from Catch+Hit Trials, normalized by AUC, and is plotted with respect to CS onset (green, vertical line). Dashed, vertical lines represent the time of expected reward (i.e., the conditioned interval for the session). Calculated neural report of times (NRT) for these example neurons are shown as yellow stars on x-axis. (B) Pie charts showing proportion of responses which express reward timing (left) and, of those classified responses, the proportion of the three forms of reward timing (right). (C) Performance of kNN classifier across a range of *k* nearest neighbors. Thick black line represents average performance of classifier, shaded region represents the mean ± standard deviation. Dashed horizontal line represents chance performance. (D) Cumulative distribution plots of the calculated neural reports of time calculated from Catch+Hit trials during short delay sessions (light blue) and long delay sessions (dark blue). Vertical lines show time of reward and horizontal line shows median values for distributions. *** - p < 0.001, Wilcoxon rank-sum test.

164

165

166 We recorded from 996 neurons in the primary visual cortex which yielded 1,992 neural records

167 from Catch+Hit trials (each neuron produced two neural records: one in response to Cue 1 and

168 one in response to Cue 2). Of these 1,992 neural records from Catch+Hit trials, 410 (20.58%)

169 were classified as expressing reward timing (i.e., were classified as SI, SD, or PK). These 410

170   records were expressed by 253 neurons (25.40% of the total recorded population). Among the

171   410 records: 243 (59.27%) were classified as SI (47 neurons classified for one CS, 98 neurons

172   classified for both CS's), 105 (25.61%) were classified as SD (33 neurons classified for one CS,

173   36 neurons classified for both CS's), and 62 (15.12%) were classified as PK (30 neurons classified

174   for one CS, 16 neurons classified for both CS's). Only 11 of 996 (1.10%) of the neural records

175   from Sham+Hit trials were classified as one of the forms described above. Figure 3B shows the

176   proportions of neural records classified. We cross-validated these classifications using a k-

177   Nearest Neighbors (kNN) classifier that was trained on data from even trials to predict the reward

178   timing expression of data from odd trials (see Methods). The classifier predictions matched well

179   with the manual classification across a range of parameters ($1 \leq k \leq 65$, Figure 3C) arguing that

180   neural activity within mouse V1 fall into distinct classes of reward timing activity, as previously

181   reported (Shuler and Bear, 2006; Chubykin et al., 2013; Liu et al., 2015).

182

183   Previous reward timing studies have shown that V1 reward timing corresponds to the delay to

184   reward (Shuler and Bear, 2006; Chubykin et al., 2013). By ascribing to each neuron a "neural

185   report of time" (NRT, the moment of time the neuron reports as the delay to expected reward, see

186   Methods), we asked if timing activity to a given reward delay similarly emerges in the head-fixed

187   mouse preparation. Should reward timing responses emerge to visual cues predicting a given

188   delay, the central tendency of those cues' NRT distributions should correspond to that delay.

189   Indeed, we find that the central tendencies for the NRT distributions accord well with the

190   conditioned intervals and are significantly different for short and long delay sessions ($Z = -4.95$, p

191    $= 7.49 \times 10^{-7}$, Wilcoxon rank-sum test – Figure 3D). Furthermore, the NRTs calculated from the

192   cross-validated responses described above also show similar significant changes in distributions

193   (i.e., shorter for the short delay) across the range of values for k (all Z's ≤ -3.43, all p's $\leq 5.97 \times$

194   10^{-4}, Wilcoxon rank-sum test). This shifting of the NRT distributions across conditioned intervals

195   is not explained by differences in licking behavior as licking alone has no effect on recorded neural

196   activity (p = 0.198 Wilcoxon Sign Rank Test) and an animal's within-session licking behavior did

197   not influence the calculated NRT (χ^2(2, 1120) = 1.91, p = 0.385, Kruskal-Wallis Test, Figure 3,

198   Supplemental Figure 1).

199

200   In previous reports of reward timing, neurons showed "cue dominance" (i.e., expressing reward

201   timing activity to one, but not both cues) when the two cues are paired with different delays. Here

202   we find that when the two cues predict a reward at the same delay that neurons can express such

203   cue-specificity in reward timing, but that there is an increase in neurons with reward timing to both

204   cues. Specifically, we find that of the 253 neurons that have reward timing, 157 (62%) express

205   reward timing to both cues and do so with notable similarity across cues (Figure 3, Supplemental

206   Figure 2). The remaining 38% express reward timing to one, but not the other cue, despite the

207   cues foretelling of the same magnitude and delay to reward. Additionally, we find that neurons

208   have stable representations of time across days. Using a previously-described statistic based on

209   waveform shape (the J3 statistic, (Moran and Katz, 2014)) we defined 100 pairs of putative

210   repeatedly-recorded neurons across two consecutive sessions. Of the 77 pairs of reward timing

211   responses across days, we find that 45 (58%) have notably similar reward timing responses

212   across days (Figure 3, Supplemental Figure 3).

213

214   Together, these data demonstrate that mouse V1 neurons are able to express reward timing

215   activity following associative learning. Having established mouse V1 as a locus for such timing

216   activity, we sought to investigate how inhibitory interneurons express and aid in the expression of

217   this reward timing activity. Specifically, we turned to recent theoretical work which proposes a

218   manner by which neurons in primary visual cortex could create such heterogeneous

219   representations of time (Huertas et al., 2015).

220

221   *V1 Neurons Represent Reward Timing in a Manner Consistent with a Theorized Network*

222   *Architecture*

223   Recent computational work posits that a simple network motif can produce reward timing activity

224   with the three known response forms (Huertas et al., 2015). This network motif is derived from a

225   recurrent network of excitatory cells with broad and sparse inhibition; it contains one population

226   of inhibitory cells and three populations of excitatory cells which differ based on levels of recurrent

227   excitation, non-recurrent excitation, and inhibition (schematized in Figure 1C). Two experimentally

228   tractable implications of this network motif are: (1) inhibitory interneurons should represent reward

229   timing predominantly as the sustained increase form and (2) neurons that are inhibited by

230   interneurons should represent time predominantly as sustained decrease or peak forms. Here we

231   test these predictions using mice which selectively express channelrhodopsin (ChR2) exclusively

232   in one of three major interneuron subtypes: those expressing parvalbumin (PV), those expressing

233   somatostatin (SOM), and those expressing vasoactive intestinal polypeptide (VIP, see Methods).

234   By investigating the ability of each interneuron subtype to fulfill the model's implications, we are

235   able to determine how known interneuron diversity intersects with the proposed network

236   architecture.

237

238   With selective ChR2 expression we are able to optogenetically identify interneurons within our

239   recorded population (Figure 4A-4B, see Methods). We identified 35/185 (18.9%) PV+ neurons,

240   15/361 (4.2%) SOM+ interneurons, and 0/203 (0%) VIP+ interneurons (example cells shown in

241   Figure 4B). These proportions match the expected relative distribution given our recording depth

242   (Tremblay et al., 2016). Additionally, a control cohort of animals (not expressing ChR2 in any cell

243   population) resulted in 0/247 (0%) recordings returned as expressing ChR2 from these wildtype
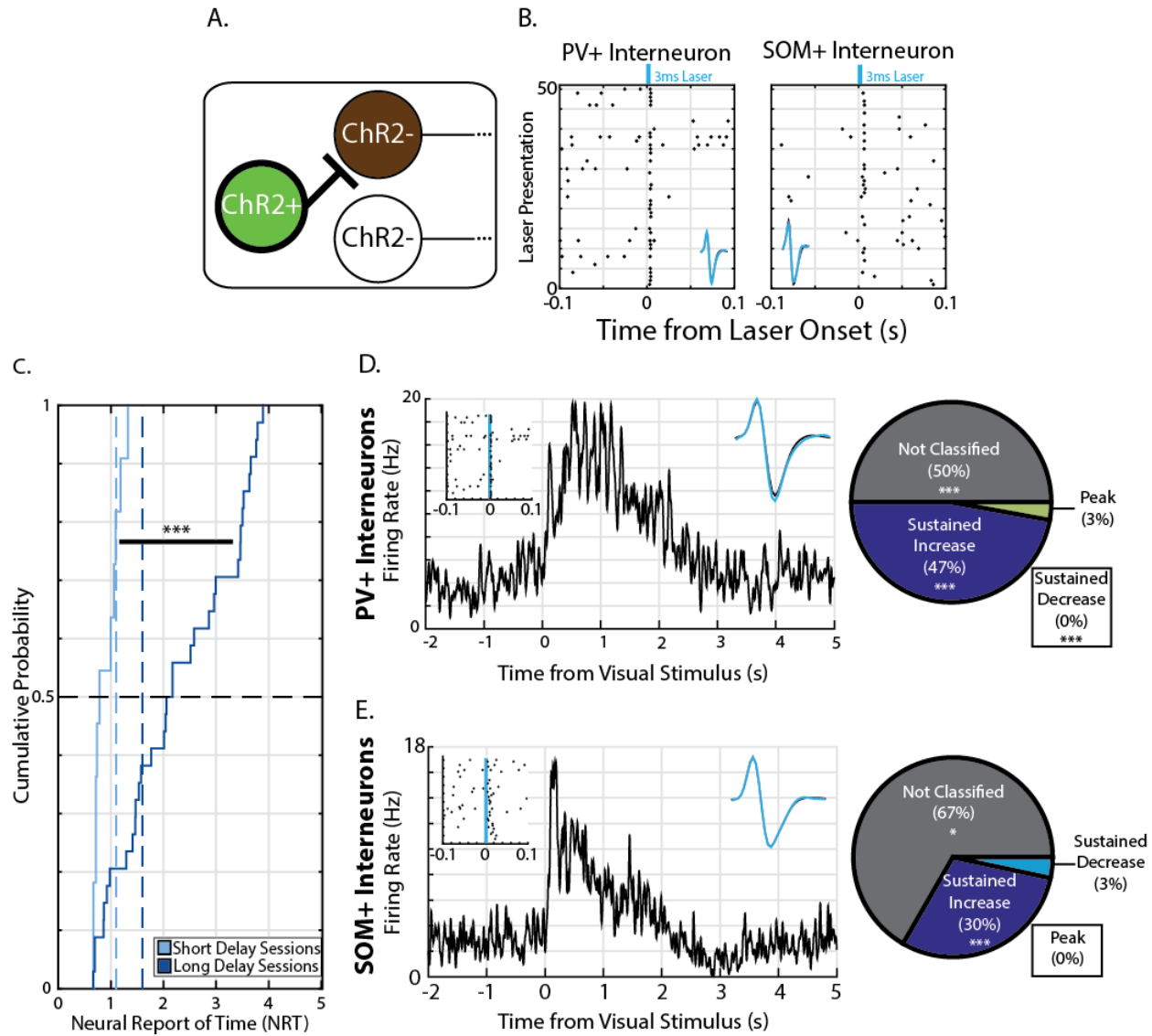
244   animals.

13

Figure 4: Identified interneurons express reward timing as SI form, consistent with theorized network architecture. (A) Cartoon showing the three types of cells we could possibly record from: ChR2-expressing (ChR2+, green), ChR2-negative which are inhibited by ChR2+ cells (ChR2-, brown), and ChR2-negative that are not inhibited by ChR2+ cells (ChR2-, white). Here we investigated reward timing in neurons expressing ChR2 as indicated with the bold lines around the ChR2+ population. (B) Two example raster plots of spikes recorded during "optotagging". Spikes are black dots plotted with respect to the time of a 3ms laser stimulus; this activity is recorded from a putative PV+ interneuron (left) and a putative SOM+ interneuron (right) and insets show average waveforms during laser stimulation (cyan) and average waveforms during spontaneous activity (black). (C) Cumulative distribution plots of NRTs calculated from identified interneurons during short delay sessions (light blue) or long delay sessions (dark blue). Vertical, dashed lines represent time of expected reward and horizontal, dashed line indicates median of the distributions. (D) Left: Example PV+ interneuron expresses reward timing as the Sustained Increase form. Inset in top left shows raster plot of spikes during optotagging and inset in top right shows average laser-evoked waveform and average spontaneous waveform (conventions as in B). Right: Pie chart showing proportions of responses from all PV+ interneurons which express reward timing and in what forms. (E) Same as E, but for SOM+ interneurons. * - p < 0.05, *** - p < 0.001, Wilcoxon Rank-Sum test (C) or Bootstrap (D – E).

245

14

246   As we optogenetically identified PV+ and SOM+ interneurons, we were then able to ascertain

247   their reward timing capabilities. We first determined their ability to produce representations of time

248   and found that their NRTs shift across conditioned intervals (Z = -3.605, p = 3.117 x 10^{-4}, Wilcoxon

249   rank-sum test; Figure 4C). Again, we verified that these representations of time are unlikely to be

250   explained by licking behavior as licking, by itself, had no significant effect on ongoing spiking

251   activity (p = 0.355, Wilcoxon Sign Rank Test; Figure 4, Supplemental Figure 1). Having

252   demonstrated that identified interneurons are expressing reward timing, we then asked what the

253   distribution of reward timing forms are for the subpopulation of interneurons. We found that 1)

254   PV+ interneurons are significantly more likely to represent the time interval than non-identified

255   counterparts, and, that 2) they are significantly more likely to represent time as a sustained

256   increase of activity (p = 8.10 x 10^{-12}, p = 6.33 x 10^{-25}, respectively, bootstrap – Figure 4D). We

257   then asked the same of SOM+ interneurons and found that, again, they are more likely to express

258   reward timing and are significantly more likely to represent time as a sustained increase of activity

259   (p = 0.018, p = 3.61 x 10^{-4}, respectively, bootstrap – Figure 4E). Additionally, although we did not

260   optogenetically identify excitatory cells, we have identified a subpopulation of putative pyramidal

261   cells using waveform shape. We did so, specifically, by using the spike width of a waveform to

262   define a population of recorded cells as wide-spiking (Barthó et al., 2004). To determine the

263   reward timing expression of putative pyramidal cells, we looked at neurons within the top quartile

264   of the spike width distribution. As expected, we find that these neurons express reward timing in

265   all forms (Figure 4, Supplemental Figure 2). These data are in accordance with the proposed

266   network architecture which suggests that inhibitory interneurons should represent the time interval

267   as the sustained increase form.

268

269   An additional component of the theorized network architecture is that neurons whose spiking is

270   inhibited by inhibitory neurons should express reward timing as either the sustained decrease or

271   peak form. To investigate this prediction, we defined cells as "suppressed" by presenting

15

272   extended laser stimuli (100ms) and recording responses (see Methods). Consistent with this

273   prediction, we found that neurons which are inhibited by PV+ activation are significantly more

274   likely to represent the time interval as the sustained decrease form (p = 1.85 x 10^{-4}, bootstrap;

275   Figure 5A). However, neurons that are inhibited by SOM+ activation were less likely to express

276   reward timing (i.e., were more likely to be not

277   classified, p = 1.61 x 10^{-5}, bootstrap) and, contrary to

278   the model's prediction, were significantly less likely to

279   be classified as sustained decrease or peak (p = 7.49

280   x 10^{-5} and p = 0.02, respectively, bootstrap; Figure

281   5B). Additionally, we find that neurons that are

282   inhibited by VIP+ activation are significantly more

283   likely to express reward timing (p = 0.018, bootstrap)

284   and have a significant enrichment of the sustained

285   increase form (p = 0.014, bootstrap; Figure 5C).

286   Together, these results show that PV+ interneurons

287   fulfill the expectations of the theorized interneuron

288   population and provide evidence in favor of the

289   proposed network architecture wherein the reward

290   timing forms arise due to the connectivity among

291   excitatory and inhibitory neurons within V1. In

292   addition, these results also point to functional

293   distinctions of interneuron subtypes in the production
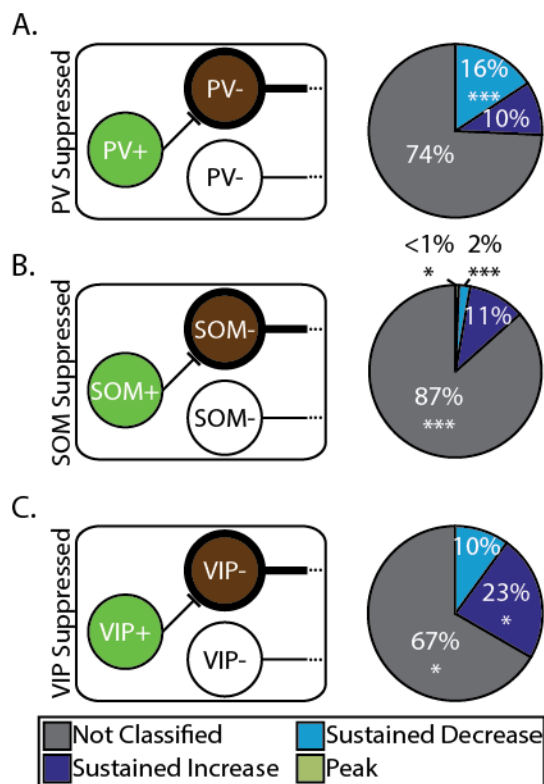
294   of reward timing activity.

295

296

297

Figure 5: Neurons inhibited by different interneuron subtypes express reward timing forms in varying proportions. (A) Left: Cartoon schematizing specific PV- neurons of interest (i.e., those inhibited by PV+ interneurons); note that though cartoon shows monosynaptic inhibition, we are unable to address the mono- or polysynaptic nature of observed inhibition. Right: Pie chart showing proportion of responses which express reward timing and in what forms. (B) Same as A but for SOM- neurons which are inhibited by SOM+ activation. (C) Same as A but for VIP- neurons which are inhibited by VIP+ activation. * - p < 0.05, ** - p < 0.01, *** - p < 0.001, bootstrap.

16

298   <u>Discussion</u>

299   Reward timing in the primary visual cortex is a network phenomenon that requires coordinated

300   activity of various cell types. Here we have shown that reward timing exists in three forms in V1

301   of head-fixed mice; that reward timing activity is expressed within identified interneuron

302   subpopulations; and, that neurons that are inhibited by different interneuron subpopulations differ

303   in their expression of reward timing. These findings are consistent with a theorized network

304   architecture.

305

306   *Reward Timing in the Primary Visual Cortex of Head-Fixed Mice*

307   Head-fixed mice were trained to associate a visual stimulus with a delayed reward (Figure 2) and

308   V1 neurons reflected this learned association in one of three forms (Figure 3). These results

309   replicate previous reward timing reports in the primary visual cortex of freely-moving rats (Shuler

310   and Bear, 2006; Chubykin et al., 2013) and mice (Liu et al., 2015), extend these reports to the

311   head-fixed preparation, and add to reports of non-sensory representations within V1 (Ji and

312   Wilson, 2007; Poort et al., 2015; Fiser et al., 2016; Pakan et al., 2018). As other sensory areas

313   express altered representations following associative learning (McGann, 2015), our

314   understanding of V1 reward timing allows for greater insight into how cortical circuits, generally,

315   can create predictions of future events.

316

317   Though we contend that the production of reward timing in V1 is the result of interactions among

318   cells within it, might it be that V1 is reflecting some non-specific global input signal (e.g., arousal

319   or attention)? Our data argue that this alternate explanation is unlikely to be the case. First, a

320   substantial fraction of neurons with reward timing show "cue dominance" (i.e., express reward

321   timing to one, but not both cues, Figure 3 Supplemental Figure 2). Such specificity in reward

322   timing is difficult to explain if V1 neurons were reflecting some non-specific, global signal. Second,

323   we find that the expression of reward timing is unaffected by how engaged an animal is in our

324 task, as indicated by its licking behavior (a measure known to co-vary with other measures of

325 arousal (Lee and Margolis, 2016)). The dissociation between licking and reward timing, then, is

326 not consistent with a global signal being the cause of V1 reward timing activity.

327

328 *V1 Neurons Express Reward Timing in a Manner Consistent with the Core Network Architecture*

329 Recent computational work has theorized a manner by which a network of cells can produce

330 reward timing activity in the various forms observed (Huertas et al., 2015). The results of this

331 formal model suggest that the heterogeneity of reward timing forms can be captured by a "core

332 network architecture" where the relative amount of inhibition, recurrent excitation, and non-

333 recurrent excitation differs according to a simple motif (Figure 1C). Here we sought to address

334 two key implications of this model to determine potential biological validity of the proposed

335 network architecture: 1) inhibitory interneurons should reflect reward timing predominantly as the

336 sustained increase form and 2) neurons inhibited by interneurons should express reward timing

337 predominantly as the sustained decrease or peak forms. Using selective expression of

338 channelrhodopsin-2 (ChR2) in interneuron subpopulations, we are able to define these two

339 populations (interneuron and suppressed) outside of behavioral conditioning and probe the

340 reward timing expression in such populations. A simplifying assumption of the model is that all

341 interneurons behave in a similar manner; however, it is known that there are various different

342 interneuron subpopulations within V1 and that they have been shown to have different functional

343 roles when the network represents sensory information (Atallah et al., 2012; Lee et al., 2012;

344 Wilson et al., 2012; Kvitsiani et al., 2013). By selectively expressing ChR2 in specific interneuron

345 subpopulations, we are able to address the model's implications and determine how these activity

346 patterns intersect with the various interneuron subtypes.

347

348 Here we identify putative PV+ and SOM+ interneurons (Figure 7B-D) and find that PV+

349 interneurons adhere to the model implications. They produce reward timing predominantly as the

350 sustained increase form (Figure 4D), and neurons that they inhibit produce reward timing with an

351 enrichment of the sustained decrease form (Figure 5A). These results can be contrasted with

352 SOM+ interneurons which, while expressing reward timing predominantly with the sustained

353 increase form (Figure 4E), largely do not inhibit neurons which express reward timing (Figure 5B).

354 Finally, the manner by which VIP+ interneurons express reward timing is unknown, but we have

355 shown that those neurons inhibited by VIP+ activation express reward timing predominantly with

356 the sustained increase form (Figure 5C). These results can be understood when known

357 connectivity is incorporated into the network architecture, as discussed below.

358

359 Inhibitory interneurons are known to have distinct connectivity among other interneurons and

360 pyramidal cells (Pfeffer et al., 2013; Tremblay et al., 2016). For example, it is known that VIP+

361 interneurons predominantly innervate other inhibitory interneurons (Pfeffer et al., 2013; Tremblay

362 et al., 2016). As such, it follows that the neurons we defined as inhibited by VIP+ interneurons

363 express reward timing in a manner similar to PV+ and SOM+ interneurons (i.e., with an

364 enrichment of the sustained increase form). Additionally, the finding that PV+ and SOM+

365 interneurons produce reward timing in a similar manner may be surprising as they are thought to

366 perform different functions in stimulus representation (Atallah et al., 2012; Lee et al., 2012; Wilson

367 et al., 2012). However, according to the theorized network architecture, interneurons express

368 reward timing as the sustained increase form because they receive input from the excitatory

369 population of sustained increase neurons (Figure 1C). Cortical interneurons are known to receive

370 convergent input from local excitatory cells (Bock et al., 2011; Fino and Yuste, 2011; Hofer et al.,

371 2011; Packer and Yuste, 2011). Specifically, a population of deep layer pyramidal cells has been

372 shown to target both PV+ and SOM+ interneurons (West et al., 2006). Perhaps the similarity in

373 reward timing expression in PV+ and SOM+ populations arises from similar pyramidal cell input

374 to these interneurons. Additionally, if this input is shared with VIP+ interneurons, it would posit

375 that VIP+ interneurons could also express reward timing as the sustained increase form. Thus,

19

376   functional differences would be borne out in downstream neurons (e.g., those neurons whose

377   activity is suppressed by interneurons, Figure 5). Finally, when comparing between PV+ and

378   SOM+ interneurons, PV+ interneurons act in a manner more consistent with the inhibitory

379   population proposed in the core network architecture. This also can be understood when one

380   considers that PV+ interneurons are the most abundant type of interneuron and provide the

381   majority of inhibition to pyramidal cells (Markram et al., 2004; Tremblay et al., 2016). Taken

382   together, we are able to overlay known connectivity patterns with the data shown here to

383   hypothesize amendments to the core network architecture (Figure 6).

384
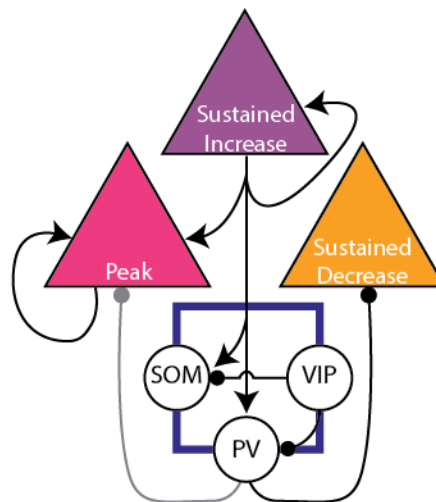
Figure 6: Predictions for core network architecture. The network architecture adapted from Figure 1 showing a potential manner by which simulated inhibitory neurons can be divided based on reward timing responses and reported connectivity. The gray line connecting PV+ interneurons is purported to exist given computational work; while not evidenced in the data presented here, it cannot be ruled out. In the current proposed architecture, PV+ interneurons fulfill expectations of the theorized inhibitory population as they produce reward timing with the SI form and inhibit neurons expressing reward timing with the SD form.

385

386

387

388   *Concluding Remarks*

389   Reward timing in the primary visual cortex, like many complex responses in the brain, requires

390   the coordinated activity of many cells. Here we have enriched the phenomenological

391   characterization of reward timing, extended that work to incorporate the heterogeneity of cell types

392   within V1, and provided a greater comprehension of how V1 produces reward timing activity

393   through its circuitry. We now better understand how the various cell types come together to

394   produce such a representation of time between a cue and a reward.

395

396   <u>Materials and Methods</u>

397   *Experimental Design and Statistical Analyses*

398   All procedures performed on animals were in accordance with the US NIH Guide for the Care and

399   Use of Laboratory Animals and were approved by the Animal Care and Use Committee at the

400   Johns Hopkins University School of Medicine. Male mice (N = 14, between 2 and 6 months old)

401   were used in this study. For this study, we used four genetic backgrounds which differ in their

402   expression of the light-activated cation channel, channelrhodopsin-2 (ChR2). One cohort

403   expressed no ChR2 (wildtype or WT, n = 3 animals), one cohort expressed ChR2 in parvalbumin-

404   positive interneurons (PV-ChR2, n = 4 animals), one cohort expressed ChR2 in somatostatin-

405   positive interneurons (SOM-ChR2, n = 4 animals), and one cohort expressed ChR2 in vasoactive-

406   intestinal-polypeptide-positive interneurons (VIP-ChR2, n = 3 animals). The WT cohort was

407   composed of C57BL/6 mice (Strain Code: 027, Charles River Laboratories). ChR2-expressing

408   mice were derived by selectively breeding the following genetic lines: a parvalbumin-Cre

409   recombinase line (PV-Cre; 008069, Jackson Laboratory, (Hippenmeyer et al., 2005)), a

410   somatostatin-Cre recombinase line (SOM-Cre; 013044, Jackson Laboratory, (Taniguchi et al.,

411   2011)), a vasoactive-intestinal-polypeptide-Cre recombinase line (VIP-Cre; 010908, Jackson

412   Laboratory, (Taniguchi et al., 2011)), and a loxP-STOP-loxP-channelrhodopsin-2-eYFP Cre-

413   dependent line (ChR2-eYFP, Ai32; 012569, Jackson Laboratory, (Madisen et al., 2010)). The

414 mice of these crosses were on mixed backgrounds composed primarily of C57BL/6 and CD-1. All

415 mice from all cohorts underwent identical training and training occurred in the light cycle during a

416 12h light/dark schedule (lights provided between 0700 and 1900).

417

418 Non-parametric tests (e.g., Kruskal-Wallis test and Wilcoxon Rank-Sum test) are used throughout

419 the text with an alpha of 0.05 unless otherwise noted. Other statistical analyses include the use

420 of the Stimulus Associated Latency Test (SALT) as described previously (Kvitsiani et al., 2013)

421 and bootstrap analyses; both methods are described in more detail below.

422

423 *Surgical Procedures*

424 Surgical procedures were performed under aseptic conditions and were in accordance with the

425 Animal Care and Use Committee at the Johns Hopkins University School of Medicine. Animals

426 underwent two surgeries spaced at least two weeks apart from one another. Prior to either

427 surgery, mice were anesthetized using a cocktail of ketamine (Ketaset, 80mg/kg) and xylazine

428 (Anased, 10 mg/kg) and eyes were covered with ophthalmic ointment (Puralube). The first surgery

429 was performed to affix a head-restraint bar to the animal's skull for training purposes and to mark

430 sites for future craniotomies. In the first surgery, the hair covering the skull was removed (Nair),

431 the skin cleaned with alternating 70% ethanol and iodine, then the skin was cut away. Following

432 this, the periosteum was removed and the skull cleaned with alternating 70% ethanol and

433 hydrogen peroxide, then the skull was dried with canned air. A total of four sites were marked for

434 future craniotomies: two for ground screws (arbitrarily marked over the anterior parietal bone) and

435 two for primary visual cortex (measured as 3mm lateral to lambda, bilaterally). Sites for future

436 craniotomies were covered in a silicone elastiomer (Smooth-On Body Double) and a head-post

437 was affixed to the anterior portion of the mouse's skull with super glue (Loctite 454). The remaining

438 bone was covered in super glue. A second surgery was performed to implant recording

439 electrodes. Briefly, small craniotomies were performed using a dental drill for ground screws and

22

440   screws were implanted into sites. Next, craniotomies were performed over V1, the dura cleaned

441   with sterile paper points, and electrodes were brought to the surface of the brain, then implanted

442   500μm below the cortical surface in accordance with stereotaxic measurements of V1 (Franklin

443   and Paxinos, 2008). Wires were covered in sterile ophthalmic ointment (Puralube) and encased

444   in dental cement (Orthojet). Ground screws and ground wires were connected and a headcap

445   was built of dental cement.

446

447   *Behavioral Task Design*

448   Prior to electrode implantation (between the first and second surgeries), animals were habituated

449   to head-fixation over the course of 2-3 days, and then were trained that a visual stimulus predicted

450   a water reward at a fixed delay for 2-3 weeks. Visual stimuli were full-field retinal flashes delivered

451   monocularly to the left (Cue 1) or right eye (Cue 2) via head-mounted goggles. These goggles

452   are custom made and consist of a miniature LED glued to the back of a translucent, plastic

453   hemidome. Licks were recorded on a lickometer via an infrared beam break (IslandMotion);

454   experiments were controlled through an Arduino Mega microcontroller board (Arduino) and events

455   were recorded with Neuralynx. In every session, trials were separated by an inter-trial interval

456   (ITI, between 3 and 8 seconds, uniformly distributed). In order to initiate the next trial (and exit the

457   ITI), animals had to cease licking for a random interval during the later portion of the ITI (deemed

458   a "lick lockout"). This lick lockout period was the same across conditioned delays and was used

459   to discourage non-stimulus-evoked licking, as licks within this period caused the timer to restart

460   and, thus, a longer ITI. Upon trial initiation, a monocular visual stimulus was either delivered (CS

461   trials) or withheld (Sham trials), followed by a delay window. CS's were visual stimuli which lasted

462   100ms and were delivered, with equal probability, to the left or right eye. The delay to reward was

463   the same for both CS's within a session and was held constant for several consecutive sessions

464   as either the short (1s) or the long (1.5s) delay. Sessions conditioned with the short delay

465   constitute the "short delay sessions"; those with the long delay, the "long delay sessions". CS

23

466   trials were further divided into "paired" and "catch" trials; paired trials being trials in which a small

467   water reward (~2μL) became available following the delay period, provided that the animal made

468   at least one lick on the lick port within the delay. Catch trials, however, were trials in which the

469   reward was withheld regardless of behavior. Licks were never rewarded during Sham trials. At

470   the conclusion of the delay window on both CS and Sham trials, the animal re-entered the ITI.

471   Trials in which the animal licked during the delay window are defined as "Hit" trials and trials in

472   which the animal did not lick during the delay window are defined as "Miss" trials. Unless otherwise

473   noted, data presented here are from Catch+Hit trials (i.e., trials in which the animal received a

474   visual stimulus, licked during the delay window, and did not receive a water reward at the end of

475   that delay).

476

477   The relative proportion of paired/catch trials and sham trials was systematically varied across

478   behavioral shaping as well as the requirement to lick within the delay window. In the final form of

479   the task (and in all sessions reported here), 80% of trials were CS trials (with equal probability of

480   being paired or catch), with the remaining trials being Sham trials.

481

482   *Behavioral Measurements*

483   The timing of individual licks was recorded using a lickometer (IslandMotion) and were recorded

484   simultaneously with neural data. During the task, animals tended to make one lick bout following

485   delivery of the CS; the timing of this bout is quantified as the time of the first lick within the bout,

486   the time of the last lick within the bout, and the mean time between these two licks ("Bout

487   Midpoint").

488

489   *Electrophysiology*

490   Neural activity was recorded bilaterally from primary visual cortex using custom-built recording

491   electrodes. Per recording electrode, 16 channels of neural data were recorded at a sampling rate

492   of 32,556 Hz through commercial hardware (Neuralynx). Neurons were offline identified through

493   manual, 3D cluster-cutting methods through commercial software (Offline Sorter, Plexon).

494   Electrodes were composed of a connector with 16 recording channels and two ground wires

495   (Omnetics). Bundles were cut at a ~45° bias to allow for sampling across a depth of approximately

496   250μm. In order to optogenetically identify interneuron subtypes, an optic fiber—composed of a

497   200μm core diameter glass multimode fiber (ThorLabs) and a 1.25mm ceramic stick ferrule

498   (Precision Fiber Products)—was glued next to the wire bundle such that the tip of the optic fiber

499   was abutted next to the majority of the electrode tips of the bundle (<200μm tip-to-tip distance

500   with some wires above optic fiber tip, some next to, and some below optic fiber tip). A schematic

501   showing the recording strategy can be found in Figure 1A.

502

503   *Neural Data Analysis*

504   The following neural data analyses were performed using custom scripts and functions in

505   MATLAB (Mathworks).

506

507   *Reward Timing Classification.* The form with which a neuron expressed reward timing was

508   determined using manual classification in a blinded fashion. Specifically, a neuron was randomly

509   selected from a random session. Then, a peri-stimulus time histogram (PSTH) calculated from

510   trials that were either Cue 1 Catch+Hit trials, Cue 2 Catch+Hit trials, or Sham+Hit trials was

511   randomly presented to an experimenter (KJM). This PSTH was then classified as "Not Classified"

512   (NC), "Sustained Increase" (SI), "Sustained Decrease" (SD), or "Peak" (PK). The remaining

513   PSTH's were presented, followed by the remaining neurons. These classifications were

514   performed without knowledge of animal identity, recording session, or delay time.

515

516   *k-Nearest Neighbors Classification.* We sought to cross-validate the human classification

517   of reward timing neurons. To do so, we implemented a *k*-Nearest Neighbors (kNN) classifier.

518   Briefly, kNN takes classified data as a "training example" to then determine the identity of

519   unclassified "query points" based on the proximity of query points to classified training example

520   points. Identity of the query point is defined as a plurality vote of its k nearest neighbors in the

521   training example. In our case, we first split data from reward timing neurons into two halves: neural

522   activity from even trials and neural activity from odd trials. Then, we normalized neural activity

523   using the area under the ROC curve (AUC, see below) and used principal components analysis

524   (PCA) for dimensionality reduction. Specifically, we reduced the normalized firing activity from

525   even trials to the first eight principal components which explained >85% of the variance within the

526   neural activity; the projection in eight dimensions and human-classified identity of the responses

527   recorded in even trials served as the training example for the kNN classifier. Then, data from the

528   odd trials were projected into the 8-dimension subspace (acting as the query points) and were

529   classified across a range of k. Specifically, we varied the number of neighbors between 1 and 65;

530   to avoid ties, we only used an odd number of neighbors in our classification.

531

532   *Neural Report of Time Classification.* To attribute a time in which neurons with reward

533   timing activity reported the expected delay to reward, we calculated the Neural Report of Time

534   (NRT). The NRT is the moment taken as the time which neurons return to a baseline level of

535   activity, for SI and SD response forms, or the time of maximum firing rate from baseline (after the

536   visual-evoked response), for the PK response form.

537

538   To calculate such a time, neural activity was normalized to the baseline firing rate by calculating

539   the area under the ROC curve (AUC) using a sliding 100ms window (Cohen et al., 2015; Sadacca

540   et al., 2018). An AUC value of 0.5 means that the ideal observer would be at chance level to tell

541   apart two distributions and values above or below 0.5 reflect greater dissimilarity among two

542   distributions. For our purposes, we found the AUC value between the distribution of spike counts

543   from a 100ms window of baseline pre-stimulus activity, and a given 100ms of spiking activity

26

544   across all trials of the same type (e.g., Paired, Cue 1 trials; Catch, Cue 1 trials; etc.). In this way,

545   we do not rely on the averaging of spike counts in the same way that a PSTH does and thus the

546   resultant value is more robust against a small subset of trials with many spikes or other forms of

547   inter-trial spiking variability. Furthermore, this method normalizes the firing rate to a value

548   bounded by 0 and 1 for every set of trials. As the AUC-normalized firing rate is the magnitude of

549   difference and not the sign of the difference between an AUC value and 0.5 (which determines

550   how dissimilar two distributions are), we found the absolute value of the difference between the

551   AUC vector and a value of 0.5. In doing so, neurons with sustained activation or suppression (SI

552   or SD neurons, respectively) could be treated with the same algorithm to calculate an NRT. We

553   operationally defined a difference threshold of 0.15 (true AUC value of 0.35 or 0.65), and, using

554   this threshold, we then defined the NRT as the first moment in time when the AUC difference

555   vector fell below the threshold for at least 100ms. For classified PK neurons, the NRT was defined

556   as the time of the maximum of this AUC difference vector. To avoid conflating reward timing

557   responses with general visual responses, we set a minimum value for valid NRTs as 0.5s after

558   stimulus offset. Though only Catch+Hit and Sham+Hit trials were classified, we were able to use

559   the Catch classifications to calculate a response's NRT in two other trial types: Paired+Hit and

560   CS+Miss. The algorithm for calculating these NRTs was identical across trial types.

561

562   *Calculation of ΔSpikes.* This value is used to determine the average change in spike rate

563   based on an animal's first lick in Sham+Hit trials. It is defined as follows:

$$\Delta Spikes = \frac{Spikes_{Pre} - Spikes_{Post}}{N}$$

565   Where $Spikes_{Pre}$ is the number of spikes in the 100ms preceding the first lick within a Sham+Hit

566   trial, $Spikes_{Post}$ is the number of spikes in the 100ms following the first lick within a Sham+Hit trial,

567   and N is the number of trials of Sham+Hit trials within the session.

568

569   *Calculation of the J3 Statistic.* This statistic was developed to determine whether neurons

570   are the same from one recording session to the next (Moran and Katz, 2014). First the waveforms

571   of all spikes recorded from two recordings are projected onto reduced dimensions using PCA.

572   Then, values are calculated as follows:

573
$$J_1 = \sum_k \sum_{k_i} \left\| s_{k_i} - m_k \right\|^2$$

574
$$J_2 = \sum_k N_k \left\| m_k - m \right\|^2$$

575
$$J_3 = J_2 / J_1$$

576   Where s_{k_i} is the projection in two dimensions of spike i in session k, m_k is the mean vector of all

577   spikes (N_k) from the kth session, m is the overall point mean of the projection, and $\|\cdot\|$ represents

578   the Euclidean Distance. In essence, the J3 value is a ratio between the Euclidean distance

579   between each spike's waveform and the center of the cluster of all other spikes' waveforms from

580   that neuron to the distance between the two clusters (i.e., a ratio of the inter- and intra-cluster

581   distance). J3 is maximal when two recordings are tightly packed and far away from one another

582   in PC space; this reflects that two recordings are unique from one another. However, we utilized

583   this statistic to determine whether a neuron recorded on one day was the same as a recording

584   made on the same channel the subsequent day. To do so, we defined a J3 threshold by finding

585   all "within" J3 values (that is, the J3 value between the first third of the recording's spikes and the

586   last third of the recording's spikes). The threshold was defined as the 95th percentile of this

587   distribution. That is to say, any neurons which were recorded from the same animal and on the

588   same channel which had a J3 value that was less than this threshold was deemed the same.

589

590   *Similarity Measurements of Reward Timing Responses.* We sought to assess the similarity

591   of reward timing responses of a given neuron across the two CS's to assess the consistency of

592   reward timing responses when different cues predicted the same reward occurring at the same

28

593   delay. Furthermore, where possible, we sought to assess the stability of a neuron's reward timing

594   response to the same stimulus across sessions. Reward timing responses of a neuron could differ

595   (or persist) between cues or across sessions in their presence, form, timing, and shape. For

596   instance, within a session, reward timing responses may be present within a given neuron to both

597   cues, exhibit the same response form (e.g., SI) with an overall similar response shape, and report

598   nominally similar NRT's. Additionally, neurons can express similar responses to the same

599   stimulus across days. To determine how similar these responses are, we first calculated the

600   concordance of reward timing forms (for example, how often a SI cell expresses reward timing as

601   SI for the opposite CS or on a following day). Among the responses which are concordant, we

602   then determined the similarity in the neuron's report of time by calculating the absolute difference

603   in NRT's. Finally, within these responses, we quantified the similarity in shape by calculating the

604   Euclidean distance between the evoked responses. These values were compared with a shuffled

605   control distribution. Shuffling distributions were calculated by shuffling across neurons that

606   expressed reward timing in the same form for the same conditioned interval.

607

608   *Neuron Identification*

609   Mice in this study (with the exception of the WT cohort) expressed channelrhodopsin-2 in one of

610   three interneuron populations: PV+, SOM+, and VIP+ interneurons (Figure 1B). This selective

611   expression allows us to determine how the diversity of V1 inhibitory interneurons intersects with

612   the theorized network architecture (as schematized in Figure 1C).

613

614   *Optogenetic Interneuron Identification.* Outside of conditioning, brief (1 or 3ms) laser

615   stimuli were randomly delivered to V1 with an inter-pulse interval randomly drawn from a

616   distribution (between 5 and 10 seconds, uniformly distributed) while recording from neurons. To

617   identify putative neurons expressing ChR2, we used the latency to the first spike and the

618   probability that a laser evoked a spike. To determine significant latencies to the first spike, we

619   used the calculated p-value from the Stimulus Associated Latency Test (SALT). This test has

620   been previously described (Kvitsiani et al., 2013); briefly, this test compares the latencies to a first

621   spike after a laser stimulus to the latencies to a first spike after arbitrary moments in time without

622   a laser presentation. Specifically, a raster of spiking activity is divided into N 10ms bins and the

623   time to a first spike within each bin is recorded. Of the N bins created, one bin is the "test bin" and

624   begins with the laser stimulus onset and one other bin is the "baseline bin" (a bin from the pre-

625   laser time period). For all N bins, a histogram of first-spike latency is created and a modified

626   Jensen-Shannon divergence is calculated between pairs of these distributions. The divergence

627   between the "baseline bin" and all other non-test bins creates a null distribution against which the

628   divergence between the "baseline" and "test" bin is compared. The resultant p-value represents

629   the probability that the divergence between the baseline and test bins falls within the null

630   distribution; we have set a conservative alpha of 0.01 as was used in the first description of the

631   method (Kvitsiani et al., 2013). In this way, neurons which have fast and consistent spikes (i.e.,

632   fire quickly and with low jitter) after a laser stimulus will be deemed significant. A caveat to this

633   statistical measure occurs when a neuron has a relatively low baseline firing rate. In such a

634   neuron, due to very low firing rates, random, spontaneous activity occurring within the test window

635   may result in a highly-significant p-value. For this reason, we also required a neuron to have an

636   action potential in the window immediately following the laser at least 20% of all laser stimulus

637   presentations.

638

639   *Identification of Pyramidal Cells via Spike Width.* In addition to interneuron identification

640   we sought to define a population of putative pyramidal cells. We did so by calculating a neuron's

641   spike width where the spike width is defined as time difference when the average waveform first

642   crosses 20% of its peak amplitude and last crosses 20% of its valley amplitude. We then set a

643   threshold at the 75th percentile of non-identified interneurons to define a population of putative

644   pyramidal cells.

645   *Optogenetic Identification of Suppressed Neurons.* Additionally, we were interested in

646   classifying neurons whose responses were inhibited by activating ChR2-expressing interneurons

647   ("suppressed neurons"). Specifically, we sought to classify those neurons that putatively do not

648   express ChR2 (i.e., did not pass one or both of the thresholds set to define interneurons, see

649   above). To determine this, we also presented 100ms laser pulses after the brief laser

650   presentations (with the same inter-pulse interval parameters). We then compared the distribution

651   of spike counts in the 100ms immediately prior to and during laser stimulation with the Wilcoxon

652   signed-rank test (WSRT). If a significant difference was found, we then compared the total number

653   of spikes between these two windows across all presentations. Significantly inhibited neurons are

654   those neurons which passed the WSRT and had fewer spikes during laser presentation than

655   before laser presentation. Although we cannot resolve the exact nature of this inhibition (either

656   mono- or polysynaptic), we are able to assess whether populations affected by inhibitory subtype

657   activation follow predictions of the computational model and whether they reveal functional

658   specialization of various interneurons. Additionally, we have limited our analysis to only those

659   neurons which are inhibited by interneuron activation as neurons which are activated during this

660   stimulation could be activated for one of at least two reasons: (1) they become disinhibited upon

661   activation of interneurons or (2) they express ChR2 but do not pass our statistical thresholds to

662   be defined as expressing ChR2.

663

664   *Bootstrap Procedures.* To determine significant changes in the proportion of neurons

665   expressing reward timing in the various forms, we used bootstrap analyses. Specifically, for a

666   given population of "test" neurons (e.g., interneurons or suppressed neurons), we randomly

667   selected a sample of neurons of the same size (with replacement) from all other neurons recorded

668   from animals of the same genotype. We then determined the expression of reward timing in this

669   subsampled distribution and created a bootstrap distribution by repeating the process 1,000

670   times. *P*-values are the probability that values found in the "test" sample would fall in the bootstrap

671   distribution.

672

673   *Histology*

674   Animals were deeply anesthetized using sodium pentobarbital (200mg/kg, Vedco). After which,

675   animals were transcardially perfused with ice cold phosphate-buffered saline (PBS) followed by

676   ice cold 4% paraformaldehyde (PFA). Brains were immersion fixed overnight in 4% PFA and were

677   transferred to 30% sucrose until sectioning. Brains were sectioned on a cryostat into 60µm slices.

678   Electrode location was verified using Nissl staining, as follows. Sections containing V1 were

679   selected and mounted on gelatin subbed slides and air dried. These slides were then immersed

680   in a solution containing 0.1% Cresyl violet and 1% glacial acetic acid dissolved in water for 5

681   minutes, followed by a 2-minute wash in distilled water, then by 2 minutes in 50% ethanol, then 2

682   minutes in 70% ethanol. Stained and washed sections were air dried, immersed in xylenes then

683   coverslipped with Permount Mounting Medium (Electron Microscopy Sciences).

684

685   Expression of ChR2 in interneuron subpopulations was verified with immunohistochemistry, as

686   follows. Brain sections containing V1 were selected for immunohistochemistry. On day 1 the

687   sections were washed three times for ten minutes each (3x10 minutes) with PBS then were

688   blocked in 10% normal goat serum (NGS) in PBS + Triton 0.1% to permeabilize and reduce

689   background binding to antibodies for 1h. Sections were then incubated with two primary

690   antibodies overnight at 4°C. Sections for all animals were incubated with a primary GFP antibody

691   to recognize the eYFP tag of the ChR2 (Chicken polyclonal, 1:2000, Aves Labs (Catalog Number:

692   GFP-1020)) and one primary antibody to recognize one of three interneuron markers: PV (rabbit

693   polyclonal, 1:2000, Swant (Catalog Number: PV27)), SOM (rat monoclonal, 1:800, EMD Millipore

694   (Catalog Number: MAB354)), or VIP (rabbit polyclonal, 1:2000, Immunostar (Catalog Number:

695   20077)). Sections were then washed 3x10 minutes with PBS, then incubated overnight at 4°C

696 with secondary antibodies: Alexa 488 Goat Anti-Chicken (1:500, Jackson ImmunoResearch

697 (Catalog Number: 103-545-155)) and Alexa 568 Goat anti Rabbit (PV and VIP) or Rat (SOM)

698 (1:500, Jackson ImmunoResearch (Catalog Numbers: 111-065-144 and 112-585-143)). Sections

699 were washed with PBS, mounted on glass slides, and coverslipped with Fluoromount-G mounting

700 medium (Electron Microscopy Sciences). To control for unspecific staining, sections were stained

701 in an identical manner except primary antibodies were omitted. Co-expression was as expected

702 given animal's genotype (Figure 1B).

703 Acknowledgements

706

707 Grants

711

712 Disclosures

713 The authors declare no competing financial interests.

714

715   <u>References</u>

716   Atallah BV, Bruns W, Carandini M, Scanziani M (2012) Parvalbumin-Expressing Interneurons
717   Linearly Transform Cortical Responses to Visual Stimuli. Neuron 73:159–170.

718   Barthó P, Hirase H, Monconduit L, Zugaro M, Harris KD, Buzsáki G (2004) Characterization of
719   Neocortical Principal Cells and Interneurons by Network Interactions and Extracellular
720   Features. Journal of Neurophysiology 92:600–608.

721   Bock DD, Lee W-CA, Kerlin AM, Andermann ML, Hood G, Wetzel AW, Yurgenson S, Soucy ER,
722   Kim HS, Reid RC (2011) Network anatomy and in vivo physiology of visual cortical
723   neurons. Nature 471:177–182.

724   Chubykin AA, Roach EB, Bear MF, Hussain Shuler MG (2013) A Cholinergic Mechanism for
725   Reward Timing within Primary Visual Cortex. Neuron 77:723–735.

726   Cohen JY, Amoroso MW, Uchida N (2015) Serotonergic neurons signal reward and punishment
727   on multiple timescales. eLife 4 Available at:
728   https://www.ncbi.nlm.nih.gov/pubmed/25714923.

729   Felleman DJ, Van Essen DC (1991) Distributed hierarchical processing in the primate cerebral
730   cortex. Cereb Cortex 1:1–47.

731   Fino E, Yuste R (2011) Dense Inhibitory Connectivity in Neocortex. Neuron 69:1188–1203.

732   Fiser A, Mahringer D, Oyibo HK, Petersen AV, Leinweber M, Keller GB (2016) Experience-
733   dependent spatial expectations in mouse visual cortex. Nature Neuroscience 19:1658–
734   1664.

735   Franklin KBJ, Paxinos G (2008) The Mouse Brain in Stereotaxic Coordinates, Third Edition.
736   New York, NY: Academic Press.

737   Gavornik JP, Bear MF (2014a) Higher brain functions served by the lowly rodent primary visual
738   cortex. Learn Mem 21:527–533.

739   Gavornik JP, Bear MF (2014b) Learned spatiotemporal sequence recognition and prediction in
740   primary visual cortex. Nature Neuroscience 17:732–737.

741   Gdalyahu A, Tring E, Polack P-O, Gruver R, Golshani P, Fanselow MS, Silva AJ, Trachtenberg
742   JT (2012) Associative Fear Learning Enhances Sparse Network Coding in Primary
743   Sensory Cortex. Neuron 75:121–132.

744   Goard MJ, Pho GN, Woodson J, Sur M (2016) Distinct roles of visual, parietal, and frontal motor
745   cortices in memory-guided sensorimotor decisions Mrsic-Flogel TD, ed. eLife 5:e13764.

746   Goltstein PM, Meijer GT, Pennartz CM (2018) Conditioning sharpens the spatial representation
747   of rewarded stimuli in mouse primary visual cortex Lee D, King AJ, Lee D, eds. eLife
748   7:e37683.

749   Guo W, Robert B, Polley DB (2019) The Cholinergic Basal Forebrain Links Auditory Stimuli with
750   Delayed Reinforcement to Support Learning. Neuron 0 Available at:

751   https://www.cell.com/neuron/abstract/S0896-6273(19)30594-X [Accessed August 6,
752   2019].

753   Hippenmeyer S, Vrieseling E, Sigrist M, Portmann T, Laengle C, Ladle DR, Arber S (2005) A
754   Developmental Switch in the Response of DRG Neurons to ETS Transcription Factor
755   Signaling. PLoS biology 3:e159.

756   Hofer SB, Ko H, Pichler B, Vogelstein J, Ros H, Zeng H, Lein E, Lesica NA, Mrsic-Flogel TD
757   (2011) Differential connectivity and response dynamics of excitatory and inhibitory
758   neurons in visual cortex. Nature Neuroscience 14:1045–1052.

759   Hubel DH, Wiesel TN (1959) Receptive fields of single neurones in the cat's striate cortex. The
760   Journal of Physiology 148:574–591.

761   Hubel DH, Wiesel TN (1965) Receptive fields and functional architecture in two nonstriate visual
762   areas (18 and 19) of the cat. Journal of Neurophysiology 28:229–289.

763   Huertas MA, Hussain Shuler MG, Shouval HZ (2015) A Simple Network Architecture Accounts
764   for Diverse Reward Time Responses in Primary Visual Cortex. The Journal of
765   Neuroscience: The Official Journal of the Society for Neuroscience 35:12659–12672.

766   Hussain Shuler MG (2016) Timing in the visual cortex and its investigation. Current Opinion in
767   Behavioral Sciences 8:73–77.

768   Ji D, Wilson MA (2007) Coordinated memory replay in the visual cortex and hippocampus
769   during sleep. Nature Neuroscience 10:100–107.

770   Kass MD, Rosenthal MC, Pottackal J, McGann JP (2013) Fear Learning Enhances Neural
771   Responses to Threat-Predictive Sensory Stimuli. Science 342:1389–1392.

772   Kay LM, Laurent G (1999) Odor- and context-dependent modulation of mitral cell activity in
773   behaving rats. Nature Neuroscience 2:1003–1009.

774   Kvitsiani D, Ranade S, Hangya B, Taniguchi H, Huang JZ, Kepecs A (2013) Distinct behavioural
775   and network correlates of two interneuron types in prefrontal cortex. Nature 498:363–
776   366.

777   Lee CR, Margolis DJ (2016) Pupil Dynamics Reflect Behavioral Choice and Learning in a
778   Go/NoGo Tactile Decision-Making Task in Mice. Front Behav Neurosci 10 Available at:
779   https://www.frontiersin.org/articles/10.3389/fnbeh.2016.00200/full [Accessed July 10,
780   2019].

781   Lee S-H, Kwan AC, Zhang S, Phoumthipphavong V, Flannery JG, Masmanidis SC, Taniguchi
782   H, Huang ZJ, Zhang F, Boyden ES, Deisseroth K, Dan Y (2012) Activation of specific
783   interneurons improves V1 feature selectivity and visual perception. Nature 488:379–383.

784   Levy JM, Zold CL, Namboodiri VMK, Hussain Shuler MG (2017) The Timing of Reward-Seeking
785   Action Tracks Visually Cued Theta Oscillations in Primary Visual Cortex. J Neurosci
786   37:10408–10420.

787 Liu C-H, Coleman JE, Davoudi H, Zhang K, Hussain Shuler MG (2015) Selective Activation of a
788   Putative Reinforcement Signal Conditions Cued Interval Timing in Primary Visual Cortex.
789   Current Biology 25:1551–1561.

790 Madisen L, Zwingman TA, Sunkin SM, Oh SW, Zariwala HA, Gu H, Ng LL, Palmiter RD,
791   Hawrylycz MJ, Jones AR, Lein ES, Zeng H (2010) A robust and high-throughput Cre
792   reporting and characterization system for the whole mouse brain. Nature Neuroscience
793   13:133–140.

794 Makino H, Komiyama T (2015) Learning enhances the relative impact of top-down processing in
795   the visual cortex. Nature Neuroscience 18:1116–1122.

796 Markram H, Toledo-Rodriguez M, Wang Y, Gupta A, Silberberg G, Wu C (2004) Interneurons of
797   the neocortical inhibitory system. Nature Reviews Neuroscience 5:793–807.

798 McGann JP (2015) Associative learning and sensory neuroplasticity: how does it happen and
799   what is it good for? Learn Mem 22:567–576.

800 Moran A, Katz DB (2014) Sensory Cortical Population Dynamics Uniquely Track Behavior
801   across Learning and Extinction. The Journal of neuroscience : the official journal of the
802   Society for Neuroscience 34:1248–1257.

803 Namboodiri VMK, Huertas MA, Monk KJ, Shouval HZ, Hussain Shuler MG (2015) Visually Cued
804   Action Timing in the Primary Visual Cortex. Neuron 86:319–330.

805 Packer AM, Yuste R (2011) Dense, Unspecific Connectivity of Neocortical Parvalbumin-Positive
806   Interneurons: A Canonical Microcircuit for Inhibition? J Neurosci 31:13260–13271.

807 Pais-Vieira M, Lebedev MA, Wiest MC, Nicolelis MAL (2013) Simultaneous Top-down
808   Modulation of the Primary Somatosensory Cortex and Thalamic Nuclei during Active
809   Tactile Discrimination. J Neurosci 33:4076–4093.

810 Pakan JMP, Currie SP, Fischer L, Rochefort NL (2018) The Impact of Visual Cues, Reward, and
811   Motor Feedback on the Representation of Behaviorally Relevant Spatial Locations in
812   Primary Visual Cortex. Cell Reports 24:2521–2528.

813 Pfeffer CK, Xue M, He M, Huang ZJ, Scanziani M (2013) Inhibition of inhibition in visual cortex:
814   the logic of connections between molecularly distinct interneurons. Nature Neuroscience
815   16:1068–1076.

816 Polley DB, Heiser MA, Blake DT, Schreiner CE, Merzenich MM (2004) Associative learning
817   shapes the neural code for stimulus magnitude in primary auditory cortex. PNAS
818   101:16351–16356.

819 Poort J, Khan AG, Pachitariu M, Nemri A, Orsolic I, Krupic J, Bauza M, Sahani M, Keller GB,
820   Mrsic-Flogel TD, Hofer SB (2015) Learning Enhances Sensory and Multiple Non-
821   sensory Representations in Primary Visual Cortex. Neuron 86:1478–1490.

822 Ross JM, Fletcher ML (2018) Learning-Dependent and -Independent Enhancement of
823   Mitral/Tufted Cell Glomerular Odor Responses Following Olfactory Fear Conditioning in
824   Awake Mice. J Neurosci 38:4623–4640.

825   Rutkowski RG, Weinberger NM (2005) Encoding of Learned Importance of Sound by Magnitude
826   of Representational Area in Primary Auditory Cortex. Proceedings of the National
827   Academy of Sciences of the United States of America 102:13664–13669.

828   Sadacca BF, Wied HM, Lopatina N, Saini GK, Nemirovsky D, Schoenbaum G (2018)
829   Orbitofrontal neurons signal sensory associations underlying model-based inference in a
830   sensory preconditioning task. eLife 7 Available at:
831   https://www.ncbi.nlm.nih.gov/pubmed/29513220.

832   Shuler MG, Bear MF (2006) Reward Timing in the Primary Visual Cortex. Science 311:1606–
833   1609.

834   Taniguchi H, He M, Wu P, Kim S, Paik R, Sugino K, Kvitsani D, Fu Y, Lu J, Lin Y, Miyoshi G,
835   Shima Y, Fishell G, Nelson SB, Huang ZJ (2011) A Resource of Cre Driver Lines for
836   Genetic Targeting of GABAergic Neurons in Cerebral Cortex. Neuron 71:995–1013.

837   Tremblay R, Lee S, Rudy B (2016) GABAergic Interneurons in the Neocortex: From Cellular
838   Properties to Circuits. Neuron 91:260–292.

839   Vincis R, Fontanini A (2016) Associative learning changes cross-modal representations in the
840   gustatory cortex Mrsic-Flogel TD, ed. eLife 5:e16420.

841   West DC, Mercer A, Kirchhecker S, Morris OT, Thomson AM (2006) Layer 6 Cortico-thalamic
842   Pyramidal Cells Preferentially Innervate Interneurons and Generate Facilitating EPSPs.
843   Cereb Cortex 16:200–211.

844   Wiest MC, Thomson E, Pantoja J, Nicolelis MAL (2010) Changes in S1 Neural Responses
845   During Tactile Discrimination Learning. Journal of Neurophysiology 104:300–312.

846   Wilson NR, Runyan CA, Wang FL, Sur M (2012) Division and subtraction by distinct cortical
847   inhibitory networks *in vivo*. Nature 488:343–348.

848   Xu X, Roby KD, Callaway EM (2010) Immunochemical characterization of inhibitory mouse
849   cortical neurons: Three chemically distinct classes of inhibitory cells. Journal of
850   Comparative Neurology 518:389–404.

851   Zold CL, Hussain Shuler MG (2015) Theta Oscillations in Visual Cortex Emerge with Experience
852   to Convey Expected Reward Time and Experienced Reward Rate. J Neurosci 35:9603–
853   9614.

854

855

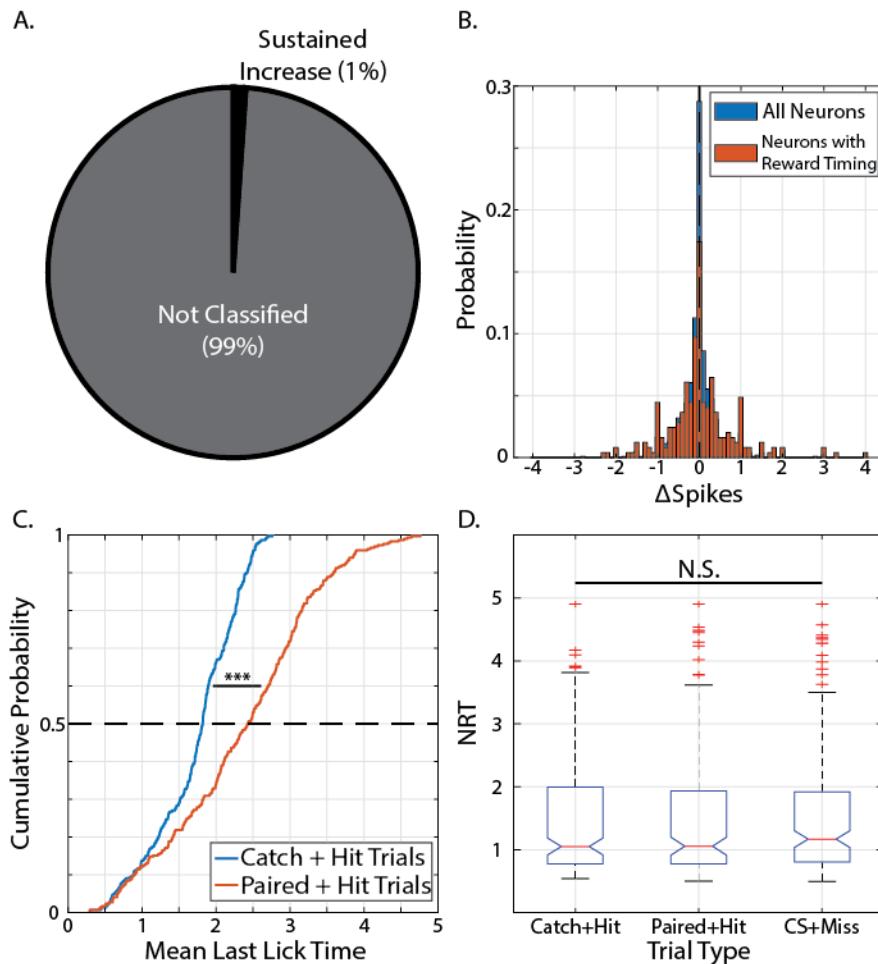856 <u>Figure Supplements</u>

Figure 3, Supplemental Figure 1: Reward timing activity is not explained by licking behavior. (A) Pie chart showing classification of responses recorded during Sham+Hit trials (trials in which no CS was given, but an animal licked). (B) Distribution of ΔSpikes (a variable which reflects the average change in the number of spikes before and after a first lick, see Methods) for all recorded neurons (blue) and neurons with reward timing responses (orange). Vertical, dashed line represents medians of distributions; neither distribution is significantly different from zero (p's > 0.05, Wilcoxon Signed-Rank test). (C) Cumulative distribution plots of Mean Last Licks within Catch+Hit trials (blue) and Paired+Hit Trials (orange) show that animals have longer lick bouts for Paired+Hit Trials compared to Catch+Hit trials (Z = -13.34, p = 1.30 x 10^{-40}, Wilcoxon Sign-Rank test). (D) Boxplot of distributions of NRTs calculated from the different trial types from neurons with reward timing responses. Box limits in box plots represent 25th and 75th percentiles, lines correspond to roughly +/-2.7σ. N.S.: Not Significant; ***: p < 0.001, Wilcoxon Sign-Rank test.
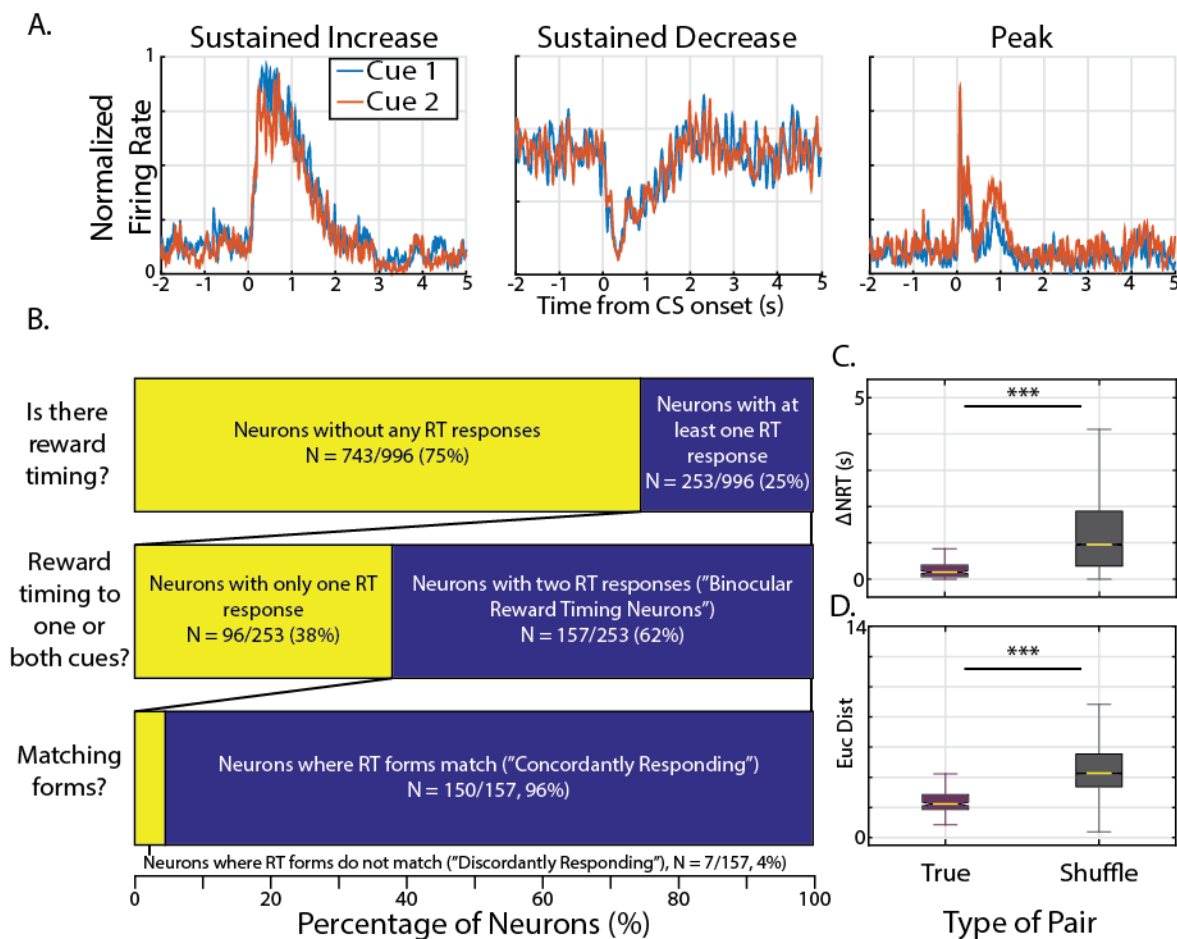
857

858

39

Figure 3, Supplemental Figure 2: Different cues elicit responses similar in form, timing, and shape within binocular reward timing neurons. (A) Normalized activity from three example neurons which have similar responses to Cue 1 as they do for Cue 2. (B) Bar chart representing the proportion of neurons with reward timing responses (top); of those cells, the proportion of cells with two classified responses ("Binocular Reward Timing Neurons", middle); and, of the Binocular Reward Timing Neurons, the proportion of neurons which express reward timing as the same form to either CS ("Concordantly Responding"). (C - D) Box plots showing differences in the absolute difference of calculated NRT's ("ΔNRT", C) and Euclidean distances (D) for "true pairs" (Cue 1 vs Cue 2 of the same neuron) and "shuffle pairs" (Cue 1 vs Cue 2 of neurons with same reward timing form and same conditioned interval). Box limits in box plots represent 25th and 75th percentiles, lines correspond to roughly +/-2.7σ. For demonstration purposes, outliers have been removed from plot. *** in panel C: Z = -12.67, p = 8.60 x 10^{-37}, Wilcoxon rank-sum test. *** in panel D: Z = -16.97, p = 1.39 x 10^{-64}.
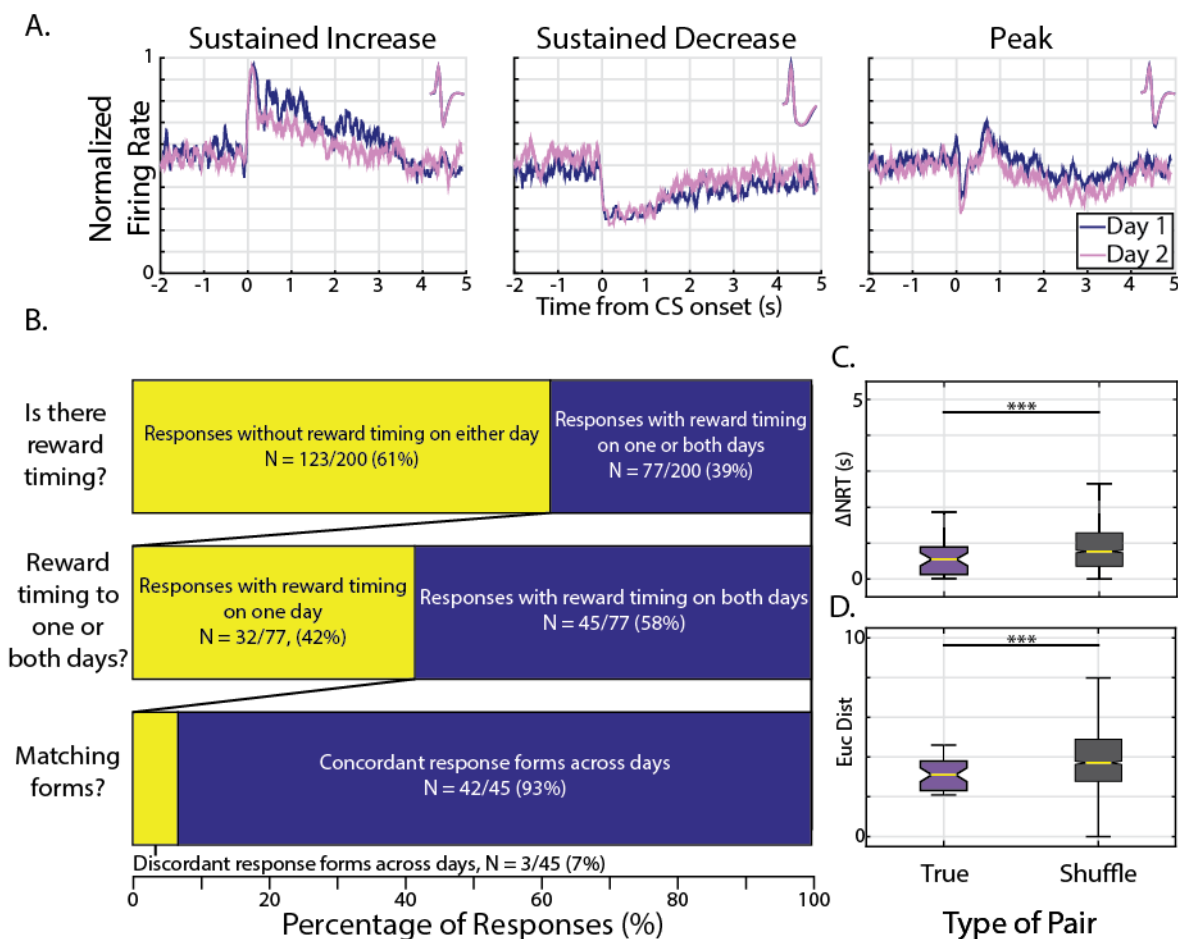
859

860

Figure 3, Supplemental Figure 3: Reward timing is stable in form, timing, and shape across recording sessions. (A) Normalized activity from three example neurons deemed to be the same neuron from different recording sessions which express reward timing similarly on day 1 (purple) and day 2 (pink); average waveforms from entire session shown in insets. (B) Bar chart representing, of all CS responses from repeat neurons, the proportion of responses with reward timing on at least one day (top); of those responses with reward timing on at least one day, the proportion of responses which had reward timing on both days; and of those responses, the proportion of responses which were classified as having the same form ("Concordant Responses"). (C - D) Box plots showing differences in the absolute difference of calculated NRT's ("ΔNRT", C) and Euclidean distances (D) for "true pairs" (Day 1 vs Day 2 for the same response) and "shuffle pairs" (Day 1 vs Day 2 of responses with same reward timing form and same conditioned interval). Box limits in box plots represent 25th and 75th percentiles, lines correspond to roughly +/-2.7σ. For demonstration purposes, outliers have been removed from plot. *** in panel C: Z = -2.73, p = 6.40 x 10^{-3}, Wilcoxon rank-sum test. *** in panel D: Z = -3.61, p = 3.07 x 10^{-4}.
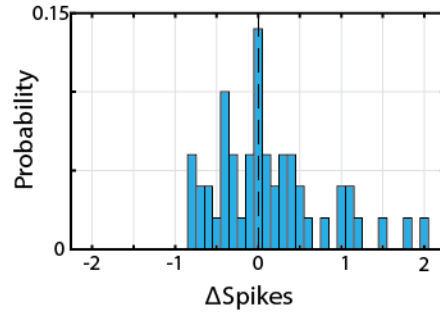
861

862

41

Figure 4, Supplemental Figure 1: Identified interneuron activity is not influenced by licking behavior. Distribution of ΔSpikes (a variable which reflects the average change in the number of spikes before and after a first lick, see Methods) for all identified interneurons. Vertical, dashed line represents median of distributions which is not significantly different from zero (Z = 0.926, p = 0.355, Wilcoxon Signed-Rank test).
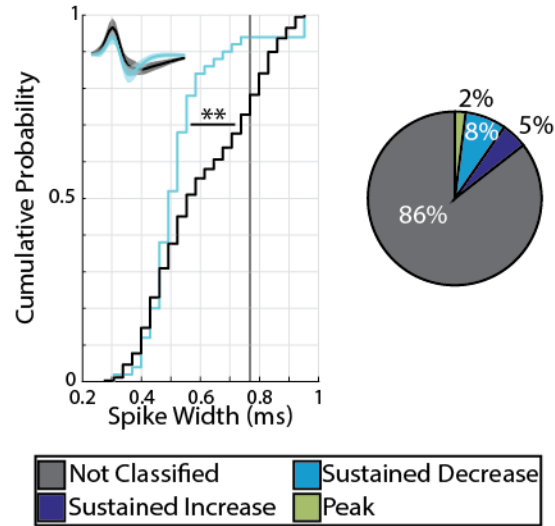
863

864

Figure 4, Supplemental Figure 2: Putative pyramidal neurons express reward timing in all forms. Left: Cumulative distribution plots of calculated Spike Width for identified interneurons (cyan) and all other neurons (black). Identified interneurons have significantly narrower spike widths compared to unidentified counterparts ($Z = -2.61$, $p = 9.2 \times 10^{-3}$, Wilcoxon Rank-Sum test). Vertical, black line shows threshold value to define putative pyramidal cells (neurons with spike widths in the top quartile of distribution). Inset at top left shows mean $\pm$ standard deviation of average waveform for identified interneurons (cyan) or putative pyramidal cells (black). Right: Pie chart showing distribution of reward timing forms within putative excitatory cells. **: $p < 0.01$, Wilcoxon Rank-Sum test

865

866