

# Timing & Time Perception

A recurrent neural network model accounts for both timing and working memory components of an interval discrimination task.

--Manuscript Draft--

<b>Manuscript Number:</b>	
<b>Full Title:</b>	A recurrent neural network model accounts for both timing and working memory components of an interval discrimination task.
<b>Short Title:</b>	Timing and working memory.
<b>Article Type:</b>	Special Issue: For Warren
<b>Corresponding Author:</b>	Dean Buonomano University of California, Los Angeles Los Angeles, CA UNITED STATES
<b>Corresponding Author's Institution:</b>	University of California, Los Angeles
<b>First Author:</b>	Rehan B Chinoy
<b>Order of Authors:</b>	Rehan B Chinoy  Ashita Tanwar  Dean Buonomano
<b>Abstract:</b>	Interval discrimination is of fundamental importance to many forms of sensory processing, including speech and music. Standard interval discrimination tasks require comparing two intervals separated in time, and thus include a working memory (WM) component as well as the need to tell time. Standard models of interval discrimination invoke separate circuits for the timing and WM components. Here we examine if, in principle, the same recurrent neural network can implement both. Using human psychophysics, we first explored the role of the WM component by varying the interstimulus delay. Consistent with previous studies, discrimination was significantly worse for a 250 ms delay, compared to 750 and 1500 ms delays, suggesting that for longer delays WM stably stores the first interval. We next successfully trained a recurrent neural network (RNN) on the task, demonstrating that the same network can implement both the timing and WM components. Many units in the RNN were tuned to specific intervals during the sensory epoch, and others encoded the first interval during the delay period. Overall, the encoding strategy was consistent with the notion of mixed selectivity. Units generally encoded more interval information during the sensory epoch than in the delay period, reflecting categorical encoding of short versus long in WM, rather than encoding of the specific interval. Our results demonstrate that in contrast to standard models of interval discrimination that invoke a separate memory module, the same network can, in principle, solve the timing, WM, and comparison components of an interval discrimination task.
<b>Keywords:</b>	Timing Interval discrimination working memory computational model RNN
<b>Suggested Reviewers:</b>	Devin Terhune d.terhune@gold.ac.uk Expertise: interval discrimination  Matthew Matell matthew.matell@villanova.edu expertise: timing and computational models.  Andre Cravo andrecravo@gmail.com Expertise: timing and working memory
<b>Opposed Reviewers:</b>	

**Funding Information:**

National Science Foundation  
(RI:2008741)

Dr Dean Buonomano

1  
2  
3  
4  
5  
6 **A recurrent neural network model accounts for both timing and working memory**  
7 **components of an interval discrimination task.**

8  
9  
10  
11  
12  
13  
14  
15 Rehan B. Chinoy  
16  
17 Ashita Tanwar  
18  
19 Dean V. Buonomano\*

20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36 Departments of Neurobiology and Psychology, Brain Research Institute, and Integrative  
37 Center for Learning and Memory, University of California, Los Angeles, CA,  
38 USA.  
39

40  
41 \* To whom correspondence should be addressed.  
42  
43

44  
45  
46  
47  
48 **Acknowledgments:** We thank Michael Seay for advice and technical assistance. This  
49 research was supported by NSF grant RI:2008741.  
50

51  
52  
53 **Keywords:** Interval discrimination, computational models, RNNs, working memory,  
54 interstimulus interval.  
55

56  
57 **Short Title:** Timing and working memory.  
58  
59  
60  
61  
62  
63  
64  
65

1  
2  
3  
4  
**Abstract**  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28

Interval discrimination is of fundamental importance to many forms of sensory processing, including speech and music. Standard interval discrimination tasks require comparing two intervals separated in time, and thus include both working memory (WM) and timing components. Models of interval discrimination invoke separate circuits for the timing and WM components. Here we examine if, in principle, the same recurrent neural network can implement both. Using human psychophysics, we first explored the role of the WM component by varying the interstimulus delay. Consistent with previous studies, discrimination was significantly worse for a 250 ms delay, compared to 750 and 1500 ms delays, suggesting that the first interval is stably stored in WM for longer delays. We next successfully trained a recurrent neural network (RNN) on the task, demonstrating that the same network can implement both the timing and WM components. Many units in the RNN were tuned to specific intervals during the sensory epoch, and others encoded the first interval during the delay period. Overall, the encoding strategy was consistent with the notion of mixed selectivity. Units generally encoded more interval information during the sensory epoch than in the delay period, reflecting categorical encoding of short versus long in WM, rather than encoding of the specific interval. Our results demonstrate that, in contrast to standard models of interval discrimination that invoke a separate memory module, the same network can, in principle, solve the timing, WM, and comparison components of an interval discrimination task.

29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65

1  
2  
3  
4  
5  
**Summary**  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65

The ability to discriminate the duration of sensory stimuli on the time scale of hundreds of milliseconds is critical to most forms of sensory processing. For example, processing the duration of phonemes, voice-onset times, and pauses between words is required for speech recognition. Similarly, music perception and Morse code are defined in large part by their temporal structure on the subsecond scale. In many cases, it is also necessary to compare the duration of different intervals, as in, for example, a two-interval discrimination task. In such cases, one interval must be transiently stored for comparison with the next interval in working memory. Standard models of interval discrimination invoke separate circuits for the timing and WM components. Here we used psychophysical and computational approaches to study the relationship and mechanisms underlying timing and working memory. We first showed that short delays impair interval discrimination, but medium or long delays do not impair discrimination. We next successfully trained a recurrent neural network (RNN) on the same task humans performed and demonstrated that the same network can implement the timing, working memory, and comparison components of the task.

1  
2  
3  
4 **INTRODUCTION**  
5  
6

7     *Time and space are the fundamental dimensions of our existence. Although space is*  
8     *gradually losing its value in a world of computer networks, cellular phones and virtual*  
9     *libraries, time is becoming the essence of our times, as is reflected by ever increasing*  
10    *speed, rate of return and productivity — concepts that are intrinsically related to time.*

11    Buhusi and Meck, 2005  
12  
13

14    Over the past few decades, it has become increasingly clear that we cannot  
15    understand the brain without understanding how the brain tells and represents time.  
16    Consequently, neuroscientific and psychological studies of time have expanded  
17    dramatically in recent decades. The widespread recognition of the importance of the  
18    problem of time and the growth of the timing field is Warren Meck's legacy. In addition to  
19    his long list of scientific contributions to the field of timing, he fostered interest in the  
20    problem of time and nurtured the generation of scientists responsible for driving the  
21    expansion of the timing field.  
22  
23

24    We now understand that the problem of time is not a single problem, but many  
25    interrelated problems (Meck and Ivry, 2016). Here, as scientists that have been influenced  
26    by Warren Meck's legacy, we focus on one of these sub-problems: interval discrimination  
27    on the subsecond scale. Interval discrimination on a subsecond scale is of fundamental  
28    importance to sensory processing across the animal kingdom, ranging from  
29    communication in crickets, electric fish and frogs, echolocation in bats, and speech,  
30    music, and Morse code in humans (Covey and Casseday, 1999; Doupe and Kuhl, 1999;  
31    Bueti and Buonomano, 2014; Rose, 2014; Matthews and Meck, 2016; Motanis et al.,  
32    2018). In a standard two-interval forced choice (2IFC) interval discrimination task,  
33    subjects first listen to an interval bounded by two tones and, after an interstimulus delay,  
34    a second comparison interval is presented, following which subjects must make a  
35    judgement as to whether the first or second interval was longer (Bueti and Buonomano,  
36    2014). While there are several variants of the task, this 2IFC format generally requires a  
37    working memory component to store the first interval for comparison with the second.  
38    This working memory component and the interaction between timing and memory has  
39    generally not been a focus of interval discrimination studies. Here we take steps in that  
40    direction by using human psychophysics to explore the effects of the delay on interval  
41    discrimination performance and develop a recurrent neural network (RNN) based  
42    computational model that performs both the timing and working memory components of  
43    the 2IFC interval discrimination task.  
44  
45

46    A number of different models have been put forth regarding how the brain may tell  
47    time in the subsecond range, including pacemaker-accumulators, ramping activity, the  
48    oscillator based beat-frequency model, neural population clocks, and state-dependent  
49    network models (Creelman, 1962; Treisman, 1963; Gibbon et al., 1984; Meck, 1996;  
50    Mauk and Buonomano, 2004; Buhusi and Meck, 2005; Merchant et al., 2013; Balci and  
51    Simen, 2016; Buhusi et al., 2016; Paton and Buonomano, 2018; Gupta et al., 2022).  
52    Among these, the state-dependent network (SDN) model, which is constrained to the  
53    subsecond range, has made specific predictions regarding the importance of the  
54    interstimulus interval, or what can be considered the WM epoch. As originally proposed,  
55  
56

time-varying changes in network state produced by neuronal and synaptic properties, such as short-term synaptic plasticity, changed the hidden state of a network (that is, in the absence of ongoing neural activity) in a manner that allowed networks to respond to stimuli in a time-dependent manner (Buonomano and Merzenich, 1995; Buonomano, 2000; Karmarkar and Buonomano, 2007). An early prediction of the SDN model was that short interstimulus delays or the presence of cross-trial temporal uncertainty would degrade performance because the network would not have time to reset to the same baseline state between the first and second intervals. A number of studies have tested these predictions, for example, by using short interstimulus delays between the first and second stimuli and have generally concluded that, at least for very short intervals, e.g., below 300 ms, predictions of the SDN model generally hold true (Karmarkar and Buonomano, 2007; Buonomano et al., 2009; Spencer et al., 2009; Fornaciai et al., 2018; Sadibolova et al., 2021).

In varying the delay between the intervals being discriminated these studies are potentially also varying the working memory requirements of the task—i.e., the amount of time available to store interval information and the amount of time this information must be stored for. It is unlikely that the decreased performance observed for short interstimulus delays are produced because there is too little time to store the intervals in working memory because, as predicted by the SDN model, if the first and second intervals are presented at different auditory frequencies or visual locations, short interstimulus intervals do not impair performance (Karmarkar and Buonomano, 2007; Buonomano et al., 2009; Fornaciai et al., 2018).

The SDN model, however, does not address the problem of working memory—the transient storage of the first interval—or the mechanisms that underlie how the first and second interval are compared. Early models of timing however, explicitly invoked distinct circuits for the storage and comparison of the intervals. For example, scalar timing models hypothesized a distinct memory module for storing the reference interval and a comparator circuit which ultimately drove the behavioral response (Gibbon, 1991; Meck, 1996; Gibbon et al., 1997; Matell and Meck, 2000). While these modules were often embedded within a neuroanatomical framework, they did not provide neurobiological interpretations for how these functions were performed.

Here we first examined not only the effects of short interstimulus delays, but also the effects of long interstimulus delays on interval discrimination. We also examined the effects of the order of the short, medium, and long delay blocks on performance. We show that while performance is strongly affected by short delays, performance for the medium and long delays was similar. This result suggests that, consistent with standard models of working memory, interval information might be stored in steady state activity during the delay period (Wang, 2001; D'Esposito and Postle, 2015). Because some brain areas, most notably the prefrontal cortex, have been implicated in some forms of timing (Oshio et al., 2008; Kim et al., 2013; Xu et al., 2014; Bakurin et al., 2017; Emmons et al., 2017) and working memory (Funahashi et al., 1989; Quintana and Fuster, 1992; D'Esposito and Postle, 2015; Stokes, 2015), we trained a supervised RNN on the same task used in the human studies. Analyses of the RNN showed that units can encode sensory, timing, and

1  
2  
3  
4 94 working memory components and that, consistent with the notion of mixed selectivity,  
5 95 (Buonomano and Maass, 2009; Rigotti et al., 2013) some units could encode more than  
6 96 one feature of the task.  
7  
8 97  
9 98

10 99 **MATERIALS AND METHODS**  
11 100

12 101 *Subjects.* Subjects consisted of eighteen paid individuals (thirteen females) between the  
13 102 ages of 18 and 39 years. The participants provided written informed consent prior to the  
14 103 experiment and were compensated monetarily for their participation (\$8), supplemented  
15 104 with a bonus (\$7) if the percentage of correct responses exceeded 65%. The study was  
16 105 approved by the Institutional Review Board at the University of California, Los Angeles.  
17 106 No participants reported having any previous experience in psychoacoustic or interval  
18 107 discrimination tasks from our research lab.  
19  
20  
21  
22

23 108 All participants were recruited through Prolific ([www.prolific.co](http://www.prolific.co)), an online data  
24 109 collection service, and reported residency in the United States and fluency in English. The  
25 110 experiment was hosted by Gorilla ([www.gorilla.sc](http://www.gorilla.sc)), an online experiment builder (Anwyl-  
26 111 Irvine et al., 2021). For consistency and precision purposes, participants were not allowed  
27 112 to complete the experiment using mobile phones or tablets and were limited to the  
28 113 Chrome and Firefox browsers.  
29  
30  
31

32 114 *Experiment.* Participants performed an auditory discrimination task consisting of nine  
33 115 blocks, split into sets of three for each experimental condition. Each set was characterized  
34 116 by three blocks of 80 trials each with the same interstimulus delay type (experimental  
35 117 condition): 250, 750, or 1500 ms. Participants were randomly assigned the order in which  
36 118 the experimental conditions appeared, counterbalanced across the six possible  
37 119 presentation orders (250-750-1500, 250-1500-750, 750-250-1500, 750-1500-250, 1500-  
38 120 250-750, and 1500-750-250). Additionally, participants had the option to take a brief break  
39 121 between blocks.  
40  
41  
42

43 122 Each trial began with an intertrial interval (ITI) sampled from a uniform distribution  
44 123 between 1 and 1.5 seconds. Following the ITI, subjects were presented with two auditory  
45 124 intervals. Each interval was bounded by two 15 ms tone pips of 1 kHz with 5 ms on and  
46 125 off ramps—all stimuli were generated in MATLAB and uploaded to Gorilla. The interval  
47 126 was measured from the offset of the first tone to the onset of the second. These stimuli  
48 127 were separated by an interstimulus delay period corresponding to the experimental  
49 128 condition of either 250, 750, or 1500 ms. The delay period for each trial was sampled  
50 129 from a uniform distribution between delay  $\pm 0.25 \times$  delay. The temporal jitter of the delay  
51 130 introduced temporal uncertainty and eliminates the introduction of a fixed time interval  
52 131 that could potentially be used as a reference to perform the task.  
53  
54  
55

56  
57 132 In order to use the same task for the human psychophysics experiments and the RNN  
58 133 simulations, we used the method of constant stimuli rather than an adaptive method  
59 134 (Lapid et al., 2008). Each pair of intervals included a standard interval of 200 ms and a  
60  
61  
62  
63  
64  
65

1  
2  
3  
4 135 comparison interval of 120, 160, 180, 190, 210, 220, 240, or 280 ms. Each block  
5 136 contained 10 trials of each comparison interval, presented in random order. Subjects were  
6 137 asked to click a button to indicate whether the second stimulus was shorter or longer than  
7 138 the first. Visual feedback was presented immediately after an incorrect response.  
8  
9 139

10 140 **Analysis.** Acceptance criteria for inclusion required subjects to have at least 65%  
11 141 accuracy for each experimental condition. Consistent with the increased variability of  
12 142 online psychophysical studies, there was a significant drop-out rate. Of the thirty-two  
13 143 participants that completed the online study, eighteen met criteria. Six of the fourteen  
14 144 participants that did not meet criteria were excluded solely based on the performance on  
15 145 the 250 ms condition. Thus, our reported discrimination performance estimates on the  
16 146 250 ms condition are likely an overestimation.  
17  
18 147

19 148 The interval discrimination threshold and the point-of-subjective equality (PSE) were  
20 149 estimated by fits to the logistic function:  
21  
22 150

23 151 
$$f(x) = \frac{1}{1 + e^{\frac{(x_0-x)}{k}}}$$
  
24  
25 152 where  $x_0$  and  $k$  correspond to the PSE and discrimination threshold (difference limen).  
26  
27 153

28 154 Statistical analyses relied on one-way repeated measures ANOVAs and post-hoc  
29 155 tests for all analyses were conducted as multiple comparisons using a Tukey-Kramer  
30 156 critical value. We used the ANOVAN and MULTCOMPARE MATLAB functions for these  
31 157 analyses.  
32  
33 158

34 159 *RNN simulations*  
35  
36

37 160 **Network structure.** Each RNN contained  $N_{rec} = 256$  units, with 204 excitatory and 52  
38 161 inhibitory units. Before time-discretization, the network activity, represented by the vector  
39 162  $r$ , followed a continuous dynamical equation (Yang et al., 2019):  
40  
41 163

$$\frac{\tau dr}{dt} = -r + f(W^{rec}r + W^{in}u + b + \sqrt{2\tau\sigma_{rec}^2}\xi)$$

42  
43 164

44 165 In this equation,  $\tau = 50$  ms represents the neuronal time constant,  $u$  is the input to the  
45 166 network,  $b$  is the bias or background input,  $f(\cdot)$  is the activation function,  $\xi$  are  $N_{rec}$   
46 167 independent Gaussian white noise processes with zero mean and unit variance and  $\sigma_{rec} =$   
47 168 0.05 is the strength of the noise.  $W^{in}$  and  $W^{rec}$  represent the input and recurrent weight  
48 169 matrices, respectively. The activation function  $f$  was a standard ReLU function:  
49  
50 170

$$f(x) = \max(x, 0)$$

51  
52 171

53 172 An output unit  $z$  provided a linear readout of the network:  $z = W^{ou} r$ .  
54 173  
55 174

After using a first-order Euler approximation with a time-discretization step  $\Delta t$ , we have:

$$r_t = (1 - \alpha)r_{t-1} + \alpha f(W^{rec}r_{t-1} + W^{in}u_t + b + \sqrt{2\alpha^{-1}\sigma_{rec}^2} N(0,1))$$

where  $\alpha = \Delta t/\tau$ , and  $N(0,1)$  represents the standard normal distribution. We used a discretization step of  $\Delta t = 10$  milliseconds, thus  $\alpha = 0.2$ .

The input,  $u$ , is composed of one input channel that feeds into the recurrent layer.  $U = 0$  in the absence of a stimulus and  $U = 2$  in the presence of a stimulus. Each interval was demarcated by two “tones” of 20 ms each. Noise was added to the input channel, so  $u = U + u_{noise}$ , where  $u_{noise} = \sigma_u N(0,1)$  and  $\sigma_u = 0.005$ .

*Training Procedure.* The loss  $L_{mse}$  to be minimized was computed by time-averaging the squared errors between the network output  $z(t)$  and the target output  $\hat{z}(t)$ :

$$L_{mse} = m_t(z_t - \hat{z}_t)^2$$

The squared errors at different time points were differentially weighted by a non-negative mask matrix  $m_t$ . Before the response epoch,  $m_t = 1$ . During the first 50 ms of the response epoch, there was a “grace period” of  $m_t = 0$ , while for the rest of the response epoch,  $m_t = 2$ . We also included an L2 regularization on rates:

$$L_{L2} = \beta \frac{1}{N_{rec}} \sum_{i,t} (r_{i,t})^2$$

where we chose  $\beta = 10^{-6}$ . Thus, the total loss to be minimized was:

$$L = L_{mse} + L_{L2}$$

$W^{rec}$  was initialized as a random orthogonal matrix with a gain of 0.1. Next, Dale’s was applied by:

$$W_{init}^{rec} = DA|W_{ortho}|$$

where  $D$  is a diagonal matrix composed of +1 or -1, representing excitatory and inhibitory units, respectively. The inhibitory weights were multiplied by a factor of four (the excitatory/inhibitory ratio) to start in a balanced excitatory/inhibitory regime.  $A$  is the autapse mask, composed of ones everywhere except a diagonal of zeros imposing no self-connections.  $|W_{ortho}|$  is the absolute value of the initial orthogonal recurrent weight matrix.

Both recurrent and output weights were trained. Recurrent biases and output biases were initialized to be zero. The input weights, recurrent biases, and output biases were

not trained. Training was performed with Adam, a learning rate of 0.001, and the decay rate for the first and second moment estimates were 0.9 and 0.999, respectively. We trained each network for 4000 epochs of 400 trials segmented into 20 batches of 50 trials. The network and the training were implemented in TensorFlow.

*2IFC interval discrimination task and performance.* The 2IFC interval discrimination task used to train the RNN was based on the human psychophysical task (see above) and included a standard interval of 200 ms and comparison intervals of 120, 160, 180, 190, 210, 220, 240, and 280 ms. The delay period was from the onset of the second tone of the first interval to the onset of the first tone of the second interval. Finally, the response epoch was measured from the onset of the second tone to the end of the trial. For non-response trials (second interval short), if the network responded above a predefined threshold at any point during the trial, the trial was considered an incorrect response. For response trials (second interval long), the network was required to respond above threshold during the response epoch; if it failed to respond during this epoch, or responded at any other time, the trial was considered incorrect.

*Interval tuning.* To quantify interval tuning, we used a sensory epoch defined by a window 50 ms before to 50 ms after the offset of the first interval across trials for each first interval condition, and a WM epoch defined by the last 200 ms of the delay period. To focus on the most active units, we used the top quartile of units based on maximum activity during the sensory epoch and separately for the WM epoch. For each unit, the average activity within an epoch was normalized by its maximal activity.

*Mutual Information.* To calculate the mutual information (MI) for each unit, we first discretized the mean analog rate values into ten bins, ranging from 0 to the maximal value for that unit. We then calculated the MI for unit  $i$  as (Buonomano, 2005):

$$MI_i = H(R_i) + H(Int) - H(R_i, Int)$$

where  $H(R_i)$  is the conventional entropy measure of the binned activity.  $H(Int)$  reflects the entropy of the interval distribution, which were equiprobable (1/9).  $H(R_i, Int)$  is the joint entropy.

## RESULTS

### *Effect of the interstimulus delay on performance.*

To study the effect of short and long interstimulus delays on interval discrimination, we used an auditory 2IFC task with a standard interval of 200 ms (see Methods). Three delay conditions were examined: 250, 750, and 1500 ms (**Fig. 1A**). Importantly, to ensure temporal uncertainty (i.e., that the state of the network at the onset of the second interval was variable), the standard stimulus could be presented first or second and the delays themselves were jittered  $\pm 25\%$ . The three delays were presented in consecutive blocks and presentation order was counterbalanced across subjects.

1  
2  
3  
4 263  
5 264 There was an overall effect of delay on the interval discrimination threshold but not on  
6 265 the point-of-subjective equality (**Fig. 1B-D**). Specifically, a one-way ANOVA revealed a  
7 266 significant effect of delay ( $F_{2,34}=12.42$ ,  $p<10^{-4}$ ), and a multiple comparison analysis  
8 267 confirmed that the threshold in the 250 ms condition was significantly higher than in the  
9 268 750 and 1500 ms conditions ( $p<0.001$ ). These results are consistent with previous studies  
10 269 that reported that a 250 ms interstimulus delay impairs interval discrimination of 100 and  
11 270 200 ms standards (Karmarkar and Buonomano, 2007; Sadibolova et al., 2021), and  
12 271 establish that there is no negative impact of long delays potentially causing degradation  
13 272 of information in WM.  
14 273  
15

16 274 We also examined whether there was any short-term learning or adaptation effects by  
17 275 asking whether performance for 750 ms delay was influenced by the delay that preceded  
18 276 it (**Fig. 1E**). A one-way analysis of variance on the threshold of the 750 ms delay across  
19 277 conditions in which the 750 ms block came after the 250 or 1500 ms blocks, or was the  
20 278 first block, revealed a small but significant effect ( $F_{2,8} = 4.5$ ,  $p<0.05$ ). This effect was driven  
21 279 by the condition in which the 750 ms delay was preceded by 250 ms block. Because the  
22 280 250 ms delay is significantly harder, this result is consistent with the notion that perceptual  
23 281 learning benefits from exposure to easier stimuli first (Ahissar and Hochstein, 2004).  
24 282  
25 283

### 26 284 *A RNN model of interval discrimination.*

27 285 Artificial neural networks and supervised RNNs in particular have proven to be a  
28 286 valuable tool towards understanding how neural circuits perform an array of different  
29 287 computations (Laje and Buonomano, 2013; Mante et al., 2013; Chaisangmongkon et al.,  
30 288 2017; Yang and Wang, 2020; Kim and Sejnowski, 2021). But as far as we are aware,  
31 289 RNN models of 2IFC interval discrimination that include both a timing and working  
32 290 memory component have not been developed. Thus, we explored if it is possible for a  
33 291 single recurrent circuit to perform this task and, if so, how time is both discriminated and  
34 292 stored in WM (**Fig. 2A**). For biological plausibility, the RNN consisted of subpopulations  
35 293 of excitatory and inhibitory units (see Methods).  
36 294  
37 295

38 296 The RNN task closely matched the psychophysical interval discrimination task.  
39 297 Initially, the RNN was trained with a standard interval of 200 ms and a delay of 750 ms.  
40 298 The comparison intervals were 120, 160, 180, 190, 210, 220, 240, or 280 ms, and the  
41 299 presentation order of the standard and comparison intervals were randomized. Training  
42 300 and testing were performed on all comparison intervals.  
43 301  
44 302

45 303 After training, the performance of five RNNs was near perfect at the trained 750 ms  
46 304 delay (**Fig. 2B**). That is, almost all trials with a short comparison interval were classified  
47 305 as short (no output response) and trials with long comparison intervals were classified as  
48 306 long (output response). Next, these RNNs were tested on the same task but with  
49 307 interstimulus delays of 250 or 1500 ms. Performance was not significantly decreased for  
50 308 the 1500 ms delay compared to the 750 ms delay but was dramatically worse for the 250  
51 309 ms delay ( $p<0.01$ ), percent correct performance was approximately 66%. The decreased  
52 310 performance mirrors the prediction of the SDN model because the RNN has not stabilized  
53 311  
54 312  
55 313  
56 314  
57 315  
58 316  
59 317  
60 318  
61 319  
62 320  
63 321  
64 322  
65 323

1  
2  
3  
4 309 or reset into a consistent network state at the onset of the second interval. Importantly,  
5 310 however, in the absence of short-term synaptic plasticity (which was not incorporated into  
6 311 the model), this RNN does not explicitly test prediction of the SDN model and confounds  
7 312 potential state-dependent effects and the time it takes the RNN to “store” a memory of  
8 313 the first interval. The generalization from 750 to 1500 ms indicates that information about  
9 314 the duration of the first interval is stored over the delay period as a fixed-point attractor in  
10 315 which the activity of the units is not changing significantly.  
11 316  
12 317

13 318 *Mechanisms underlying timing and WM in the RNN.*

14 319 The overall dynamics of all the active units of one RNN in response to a Short-Long  
15 320 (120-200 ms) and a Long-Short (280-200) trial are shown in **Fig. 2C**. Units are sorted  
16 321 according to the latency of their peak response. We can see that most cells have  
17 322 properties of sensory units, in that they respond to both the first and second “tone” of the  
18 323 first and second intervals. But it is also possible to see that the response to each tone is  
19 324 modulated by temporal context and that some units are active during an interval (i.e.,  
20 325 during the inter-tone-interval).  
21 326  
22 327

23 328 Within interval activity is ultimately responsible for timing, i.e., for implementing a  
24 329 timer. Specifically, timing of an interval requires a trace or memory of the first tone, which  
25 330 affects the response of the second tone. There are two general mechanisms by which  
26 331 neural circuits can maintain information about the past: in the hidden or active state of a  
27 332 network (Buonomano and Maass, 2009). The hidden state includes short-term synaptic  
28 333 plasticity, and the active state consists of ongoing suprathreshold neural activity. In the  
29 334 absence of short-term synaptic plasticity, the temporal information is maintained in the  
30 335 active state. Note that, in principle, the hidden state also includes the membrane time  
31 336 constant—here, however, with  $\tau = 50$  ms, passive decay should not significantly  
32 337 contribute to the timing of intervals of 200 ms.  
33 338  
34 339

35 340 We can also see that population activity decreases dramatically during the  
36 341 interstimulus delay. But critically, a small subset of units maintained an approximately  
37 342 constant level of activity during the latter part of the delay and this activity was dependent  
38 343 on the duration of the first interval. These units are encoding the WM of the first interval.  
39 344  
40 345

41 346 To better understand the timing and WM mechanisms of the RNN, it is useful to look  
42 347 at the activity of sensory, interval-tuned, and WM units across all trial types. **Fig. 3A**  
43 348 shows an example of a “pure” sensory unit in which the response to each tone of each  
44 349 interval is approximately the same—i.e., independent of the interval. **Fig. 3B** provides an  
45 350 example of an interval-tuned inhibitory unit that, at the offset of the second tone, responds  
46 351 more to the longer intervals. This unit has mixed selectivity properties because it also  
47 352 encodes (at a lower activity rate) long intervals during the WM epoch. **Fig. 3C** shows a  
48 353 WM neuron that fires tonically during the latter half of the delay period. Importantly, it fires  
49 354 at a higher rate when the first interval was short and approximately the same amount  
50 355 regardless of whether the first interval was 120, 140, 160, or 180 ms. This neuron is thus  
51 356 storing WM critical to the task, rather than the precise duration of the first interval. In  
52 357 essence, it is categorically encoding whether the first interval was short or long. **Fig. 3D**  
53 358  
54 359  
55 360  
56 361  
57 362  
58 363  
59 364  
60 365

provides an example of a “pure” interval-tuned unit that displays its tuning primarily during the sensory epoch, responding more at the offset of longer intervals.

#### *Mixed selectivity of interval tuning and interval storage in WM.*

The notion of mixed selectivity reflects experimental and computational observations that, within recurrent circuits, neurons often contribute to more than one computational component of a task (Buonomano and Maass, 2009; Rigotti et al., 2013; Fusi et al., 2016). We next determined if units in the RNN exhibit mixed selectivity, i.e., if they multiplexed timing and WM. We first examined interval selectivity during the sensory epoch (a window straddling the offset of the second tone of the first interval) and the end of the delay period (the WM epoch). **Fig. 4A** shows the tuning curves of the active units of one RNN sorted according to which interval elicited the largest response during the sensory epoch. We can see that most units exhibited preferred interval tuning to either the shortest or longest interval and that most units exhibited graded tuning to interval. **Fig. 4B** shows the interval tuning during the last 100 ms of the delay period. The plot reveals that in contrast to the sensory epoch, encoding during the WM epoch was more categorical in nature (**Supplemental Figure 1**). Thus, the strategy the RNN took to solve the task was not to store the absolute first interval in WM, but whether it was short or longer than 200 ms. In other words, the key computation of whether the first stimulus was short or long was performed before the arrival of the second interval. This observation is consistent with psychophysical studies that suggest that subjects in some cases can make short or long decisions before the arrival of the second interval, and that subjects can perform interval discrimination tasks when only one interval is presented (Buonomano et al., 2009).

To quantify the interval tuning, we calculated the mutual information (MI) each unit contained about the nine possible intervals (max MI = 3.17). This analysis focused on the units that were most active during either the sensory or WM epochs (see Methods). Overall, units contained significantly more interval information during the sensory epoch compared to the WM epoch ( $p < 0.01$ ) (**Fig. 5A-B**). To examine whether the units exhibited mixed selectivity, i.e., if they contained interval information during the sensory and WM epochs, we examined a MI scatterplot for all five RNNs (**Fig. 5A**), which revealed a highly heterogeneous distribution of units. Units primarily encoded a lot of information about the first interval, but little or no interval information during the WM epoch (upper left quadrant). In contrast, other units contained interval information during the WM, but not sensory, epoch (lower right quadrant). Other units contained more interval information during both the sensory and WM epochs (upper right quadrant). Note that the overrepresentation of units with  $MI \geq 1$  bit during the WM epoch reflect the categorical encoding of the binary information: short x long.

## **DISCUSSION**

Interval discrimination tasks often require two distinct computations: the ability to measure elapsed time and the ability to store the first interval for comparison with the second. To date, models of interval discrimination have assumed that interval

1  
2  
3  
4 401 discrimination and storage of the interval are distinct computations performed by different  
5 402 brain areas. Here we first explored whether the amount of time information is required to  
6 403 be stored in WM influences performance, and whether, in principle, the same network  
7 404 can accomplish both the timing and WM components of a 2IFC discrimination task.  
8 405  
9 406

10 406 *Interstimulus delay and working memory*  
11 407

12 408 Several studies have explored the role of the interstimulus delay in interval  
13 409 discrimination (Karmarkar and Buonomano, 2007; Buonomano et al., 2009; Fornaciai et  
14 410 al., 2018; Sadibolova et al., 2021). These studies were mostly motivated by the SDN  
15 411 model, which predicts that when there is not enough time for the network state to reset  
16 412 and there is temporal variability across trials, the first and second intervals will interfere  
17 413 with each other because the network will be in different states within each trial. For the  
18 414 most part, studies have confirmed SDN predictions for standard intervals and delays  
19 415 below 300 ms, but there is ongoing debate as to the time scale at which these predictions  
20 416 hold true (Spencer et al., 2009; Fornaciai et al., 2018; Sadibolova et al., 2021).  
21 417  
22 418

23 419 The interstimulus delay also has important implications in the context of the WM  
24 420 component of interval discrimination. The timing and working memory components of  
25 421 standard interval discrimination tasks have in general been considered to be distinct  
26 422 computations performed by distinct circuits (Meck, 1996). Recent studies, however, have  
27 423 suggested potential links between timing and working memory. For example,  
28 424 psychophysical studies in humans have shown that the temporal structure of tasks is implicitly  
29 425 learned during WM tasks and that WM is impaired when information has to be retrieved at  
30 426 unexpected times (Cravo et al., 2017; van Ede et al., 2017; Nobre and van Ede, 2018; Jin et al.,  
31 427 2020).

32 428 Here we analyzed the performance of subjects on a 2IFC discrimination task with a  
33 429 standard interval of 200 ms and explored the effect of short (250 ms), medium (750 ms),  
34 430 and long (1500 ms) interstimulus delays. Our findings again confirm that short delays  
35 431 impair performance. We did not observe any effects of long delays, as there was no  
36 432 significant difference in interval discrimination thresholds between 750 and 1500 ms  
37 433 delays. This result is consistent with the notion that once a memory of the first interval is  
38 434 stored in WM, it is stored as a fixed-point attractor and is thus time-independent.  
39 435  
40 436

41 437 Following the work of Nobre and colleagues (Cravo et al., 2017; van Ede et al., 2017;  
42 438 Nobre and van Ede, 2018; Jin et al., 2020), future studies should address the question of  
43 439 implicit timing of the interstimulus delay—specifically, whether subjects create an implicit  
44 440 expectation of the duration of the delay during interval discrimination tasks. This question  
45 441 can be addressed by using a 2IFC task with a standard delay interspersed with  
46 442 unexpected shorter or longer interstimulus delays on a small number of trials.

47 443 *An RNN model of interval discrimination*  
48 444

49 445 Here we presented what is, to the best of our knowledge, the first RNN implementation  
50 446 of a 2IFC interval discrimination task. Our goal was to determine if the timing and WM  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65

1  
2  
3  
4 447 components of the task could be performed by the same network and, if so, how the units  
5 448 of the RNN accomplished these distinct computations. Our results establish that an RNN  
6 449 can effectively learn the same 2IFC task used in our psychophysical studies. RNNs  
7 450 trained on the 750 ms delay performed the task almost perfectly in terms of percent  
8 451 correct. We did not report the RNN results in terms of their psychophysical thresholds  
9 452 because they are determined primarily by the intervals used during the training—and the  
10 453 threshold can probably be arbitrarily low when an RNN is trained with intervals very close  
11 454 to the standard. When the RNNs trained on the 750 ms delay were tested on the 250 and  
12 455 1500 ms interstimulus delays, we observed a significantly impaired performance at the  
13 456 250 ms delay. This result loosely mirrors the impairment at 250 ms in the psychophysical  
14 457 studies. In the RNN, the effect is due to the interference between the first and second  
15 458 interval because the network has not fully converged to steady state activity during the  
16 459 delay period. However, we stress that this result does not comprise a test of the SDN  
17 460 model because short-term synaptic plasticity was not incorporated into the model and  
18 461 retraining on the 250 ms resulted in near perfect performance.  
19 462  
20 463

21 464 Both in the context of timing tasks and the wide variety of other psychophysical tasks  
22 465 that rely on working memory, it has generally been assumed that working memory is a  
23 466 distinct computation performed by specific neural circuits. However, converging  
24 467 experimental (Miller et al., 1996; Rainer et al., 1999; Mante et al., 2013; Carnevale et al.,  
25 468 2015) and computational data (Mante et al., 2013; Rigotti et al., 2013; Fusi et al., 2016;  
26 469 Goudar and Buonomano, 2018; Orhan and Ma, 2019; Yang et al., 2019) suggest that the  
27 470 critical computations for a given task (e.g., interval discrimination or motion integration  
28 471 during a random-dot motion task) and WM may be performed and encoded not only with  
29 472 the same circuits, but within the same neurons.  
30 473  
31 474

32 475 Early models of interval discrimination have invoked three separate computations,  
33 476 each performed by a separate module: a pacemaker responsible for timing per se, a  
34 477 memory module responsible for transiently storing a reference interval, and a comparator  
35 478 responsible determining whether the current or stored values is larger and thus if the first  
36 479 or second interval is longer (Gibbon, 1991; Meck, 1996; Gibbon et al., 1997; Matell and  
37 480 Meck, 2000). Here we have established that, in principle, a single circuit can perform a  
38 481 2IFC task and thus all three computations can be performed in a multiplexed fashion by  
39 482 the same group of units.  
40 483  
41 484

42 485 Key to the ability to perform these three components is the notion of state-dependent  
43 486 computations (Buonomano and Maass, 2009; Goudar and Buonomano, 2018). After each  
44 487 interval, the network converges to a steady-state fixed point attractor during the delay  
45 488 period, which encodes information about the first interval. This fixed point serves as the  
46 489 starting state at the arrival of the second interval. And depending on the interaction  
47 490 between this starting state and the duration of the second interval, the network generates  
48 491 a short or long response—effectively implementing the comparator function.  
49 492  
50 493

51 494 Our goal is not to propose that the RNNs simulated here represent a biologically  
52 495 realistic implementation of the 2IFC task. Indeed, consistent with the SDN model of  
53 496 timing, a converging body of data suggest that sensory areas contribute to the timing of  
54 497  
55 498

1  
2  
3  
4 493 intervals on the subsecond scale, and higher-order areas underlie the WM component  
5 494 (Shuler and Bear, 2006; Karmarkar and Buonomano, 2007; Chubykin et al., 2013;  
6 495 Namboodiri et al., 2014; Motanis et al., 2018; Paton and Buonomano, 2018; Monk et al.,  
7 496 2020). Nevertheless, higher-order prefrontal areas have been implicated in both timing  
8 497 (Oshio et al., 2008; Kim et al., 2013; Xu et al., 2014; Bakhurin et al., 2017; Emmons et  
9 498 al., 2017) and working memory (Funahashi et al., 1989; Quintana and Fuster, 1992;  
10 499 D'Esposito and Postle, 2015; Stokes, 2015). Thus, future work should determine if interval  
11 500 discrimination in the subsecond range is performed in a modular fashion, as well as  
12 501 whether, in some cases, the same neural circuits are indeed capable of performing both  
13 502 timing and WM components in a multiplexed fashion. Finally, our results show how the  
14 503 same network can implement three computations—timing, memory, and comparison—  
15 504 that were previously proposed to be performed by three distinct modules.  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65

1  
2  
3  
4 505 **FIGURE LEGENDS**  
5 506  
6  
7 507  
8 508  
9 509  
10 510  
11 511  
12 512  
13 513  
14 514  
15 515  
16 516  
17 517  
18 518  
19 519  
20 520  
21  
22  
23  
24 521  
25 522  
26  
27 523  
28 524  
29 525  
30 526  
31 527  
32 528  
33 529  
34 530  
35 531  
36 532  
37 533  
38 534  
39  
40  
41 535  
42 536  
43 537  
44 538  
45 539  
46 540  
47 541  
48 542  
49 543  
50  
51 544  
52 545  
53 546  
54 547  
55 548  
56  
57 549  
58  
59  
60  
61  
62  
63  
64  
65

**Figure 1.** Psychophysics. (A) Schematic illustration of the task. Each trial consisted of two auditory intervals bounded by 15 ms tones (standard interval = 200 ms). The inter-stimulus delay was either 250 ms, 750 ms, or 1500 ms. Participants responded by indicating whether the second interval was shorter or longer than the first. (B) Threshold data for all participants by delay condition. The interval discrimination threshold was significantly higher for the 250 ms interstimulus delay condition compared to the 750 and 1500 ms conditions ( $p < 0.01$ ). There was no significant difference between the 750 and 1500 ms condition. (C) Point of subjective equality (PSE) data for all participants. There was no significant difference between the average PSE between the three delay types. (D) Fitted psychometric functions for all 18 subjects on the 250 ms, 750 ms, and 1500 ms delay conditions. (E) Threshold data for the 750 ms delay condition separated by whether it was preceded by a 250 ms block, a 1500 ms block, or no blocks. A significant increase in threshold for the 750 ms delay condition was observed when it was preceded by the 250 ms delay condition ( $p < 0.05$ ).

**Figure 2.** (A) Schematic of RNN model. The input unit projects to the RNN composed of 256 units (204 excitatory, 52 inhibitory). The recurrent units project to the output unit. The network is trained to produce a motor response for trials in which the second interval is longer than the first interval following the offset of the second interval. (B) Performance. Percent correct trials for five networks on three different delay conditions. Networks were trained on the 750 ms condition. There was a significant decrease in performance in the 250 compared to the 750 ms condition. ( $p < 0.01$ , Wilcoxon rank-sum test). (C) Sample population activity of a trained RNN tested on two different trial conditions. Red bars indicate input pulses. Left panel shows network activity for the short interval (120 ms) followed by the standard interval (200 ms) case, and right panel shows network activity for long-standard case. Cyan plots represent activity of the output unit.

**Figure 3.** Four example RNN units. For each panel, the top plot shows activity for trials in which the standard interval (200 ms) is shown first, and the bottom plot shows activity for trials in which the standard is second. (A) A sensory unit that responds consistently to each input pulse. (B) An interval-tuned unit that exhibits progressively stronger responses to longer intervals. The unit also encodes long intervals during the WM epoch. (C) WM unit that exhibits approximately categorical response to short-first intervals. (D) Interval-tuned unit that preferentially responds to long intervals during the sensory epoch.

**Figure 4** (A) Interval tuning to first interval (sensory epoch). Each row represents a single unit's tuning function across all nine possible first interval conditions during the sensory epoch. The units are sorted in order of preferred interval tuning. Only the top quartile of active units is shown. (B) Interval tuning during the WM epoch. Note that in contrast to interval tuning during the sensory epoch, most units categorically encode short (<200 ms) or long (>200 ms) intervals.

1  
2  
3  
4 550 **Figure 5** (A) Scatter plot of the mutual information (in bits) that each active units contains  
5 551 about the nine intervals during the sensory and WM epochs. Note that the  
6 552 overrepresentation of WM MI values that approximate 1 reflect the categorical encoding  
7 553 of short versus long intervals. (B) Mean mutual information for all data points shown in A.  
8 554 There is significantly greater mutual information during the sensory epoch compared to  
9 555 the WM epoch ( $p < 0.01$ , Wilcoxon rank-sum test).  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65

1  
2  
3  
4 556 **REFERENCES**  
5 557  
6  
7 558 Ahissar M, Hochstein S (2004) The reverse hierarchy theory of visual perceptual learning.  
8 559 Trends in Cognitive Sciences 8:457-464.  
9 560 Anwyl-Irvine A, Dalmajer ES, Hodges N, Evershed JK (2021) Realistic precision and accuracy of  
10 561 online experiment platforms, web browsers, and devices. Behavior Research Methods  
11 562 53:1407-1425.  
12  
13 563 Bakhurin KI, Goudar V, Shobe JL, Claar LD, Buonomano DV, Masmanidis SC (2017) Differential  
14 564 Encoding of Time by Prefrontal and Striatal Network Dynamics. The Journal of  
15 565 Neuroscience 37:854-870.  
16  
17 566 Balci F, Simen P (2016) A decision model of timing. Current Opinion in Behavioral Sciences 8:94-  
18 567 101.  
19  
20 568 Bueti D, Buonomano DV (2014) Temporal Perceptual Learning. Timing and Time Perception  
21 569 2:261-289.  
22 570 Buhusi CV, Meck WH (2005) What makes us tick? Functional and neural mechanisms of interval  
23 571 timing. Nat Rev Neurosci 6:755-765.  
24  
25 572 Buhusi CV, Oprisan SA, Buhusi M (2016) Clocks within clocks: timing by coincidence detection.  
26 573 Current Opinion in Behavioral Sciences 8:207-213.  
27  
28 574 Buonomano DV (2000) Decoding temporal information: a model based on short-term synaptic  
29 575 plasticity. J Neurosci 20:1129-1141.  
30 576 Buonomano DV (2005) A learning rule for the emergence of stable dynamics and timing in  
31 577 recurrent networks. J Neurophysiol 94:2275-2283.  
32  
33 578 Buonomano DV, Merzenich MM (1995) Temporal information transformed into a spatial code  
34 579 by a neural network with realistic properties. Science 267:1028-1030.  
35  
36 580 Buonomano DV, Maass W (2009) State-dependent Computations: Spatiotemporal Processing in  
37 581 Cortical Networks. Nat Rev Neurosci 10:113-125.  
38  
39 582 Buonomano DV, Bramen J, Khodadadifard M (2009) Influence of the interstimulus interval on  
40 583 temporal processing and learning: testing the state-dependent network model. Philos  
41 584 Trans R Soc Lond B Biol Sci 364:1865-1873.  
42  
43 585 Carnevale F, de Lafuente V, Romo R, Barak O, Parga N (2015) Dynamic Control of Response  
44 586 Criterion in Premotor Cortex during Perceptual Detection under Temporal Uncertainty.  
45 587 Neuron 86:1067-1077.  
46  
47 588 Chaisangmongkon W, Swaminathan SK, Freedman DJ, Wang X-J (2017) Computing by Robust  
48 589 Transience: How the Fronto-Parietal Network Performs Sequential, Category-Based  
49 590 Decisions. Neuron 93:1504-1517.e1504.  
50  
51 591 Chubykin AA, Roach EB, Bear MF, Shuler MGH (2013) A Cholinergic Mechanism for Reward  
52 592 Timing within Primary Visual Cortex. Neuron 77:723-735.  
53  
54 593 Covey E, Casseday JH (1999) Timing in the auditory system of the bat. Annu Rev Physiol 61:457-  
55 594 476.  
56  
57 595 Cravo AM, Rohenkohl G, Santos KM, Nobre AC (2017) Temporal Anticipation Based on Memory.  
58 596 Journal of Cognitive Neuroscience 29:2081-2089.  
59  
60 597 Creelman CD (1962) Human discrimination of auditory duration. J Acoust Soc Am 34:582-593.  
61 598 D'Esposito M, Postle BR (2015) The Cognitive Neuroscience of Working Memory. Annu Rev  
62 599 Psychol 66:115-142.  
63  
64  
65

- 1  
2  
3
- 4 600 Doupe AJ, Kuhl PK (1999) Birdsong and human speech: common themes and mechanisms. *Annu  
5 601 Rev Neurosci* 22:567-631.
- 6 602 Emmons EB, De Corte BJ, Kim Y, Parker KL, Matell MS, Narayanan NS (2017) Rodent Medial  
7 603 Frontal Control of Temporal Processing in the Dorsomedial Striatum. *The Journal of  
8 604 Neuroscience* 37:8718-8733.
- 9 605 Fornaciai M, Markouli E, Di Luca M (2018) Modality-specific temporal constraints for state-  
10 606 dependent interval timing. *Scientific Reports* 8:10043.
- 11 607 Funahashi S, Bruce CJ, Goldman-Rakic PS (1989) Mnemonic coding of visual space in the  
12 608 monkey's dorsolateral prefrontal cortex. *J Neurophysiol* 61:331-349.
- 13 609 Fusi S, Miller EK, Rigotti M (2016) Why neurons mix: high dimensionality for higher cognition.  
14 610 *Current Opinion in Neurobiology* 37:66-74.
- 15 611 Gibbon J (1991) Origins of scalar timing. *Learning and Motivation* 22:3-38.
- 16 612 Gibbon J, Church RM, Meck WH (1984) Scalar timing in memory. *Ann N Y Acad Sci* 423:52-77.
- 17 613 Gibbon J, Malapani C, Dale CL, Gallistel CR (1997) Toward a neurobiology of temporal cognition:  
18 614 advances and challenges. *Curr Opinion Neurobiol* 7:170-184.
- 19 615 Goudar V, Buonomano DV (2018) Encoding sensory and motor patterns as time-invariant  
20 616 trajectories in recurrent neural networks. *eLife* 7:e31134.
- 21 617 Gupta A, Bansal R, Alashwal H, Kacar AS, Balci F, Moustafa AA (2022) Neural Substrates of the  
22 618 Drift-Diffusion Model in Brain Disorders. *Frontiers in computational neuroscience*  
23 619 15:678232-678232.
- 24 620 Jin W, Nobre AC, van Ede F (2020) Temporal Expectations Prepare Visual Working Memory for  
25 621 Behavior. *J Cogn Neurosci* 32:2320-2332.
- 26 622 Karmarkar UR, Buonomano DV (2007) Timing in the absence of clocks: encoding time in neural  
27 623 network states. *Neuron* 53:427-438.
- 28 624 Kim J, Ghim J-W, Lee JH, Jung MW (2013) Neural Correlates of Interval Timing in Rodent  
29 625 Prefrontal Cortex. *The Journal of Neuroscience* 33:13834-13847.
- 30 626 Kim R, Sejnowski TJ (2021) Strong inhibitory signaling underlies stable temporal dynamics and  
31 627 working memory in spiking neural networks. *Nature Neuroscience* 24:129-139.
- 32 628 Laje R, Buonomano DV (2013) Robust timing and motor patterns by taming chaos in recurrent  
33 629 neural networks. *Nat Neurosci* 16:925-933.
- 34 630 Lapid E, Ulrich R, Rammsayer T (2008) On estimating the difference limen in duration  
35 631 discrimination tasks: A comparison of the 2AFC and the reminder task. *Perception &*

36 632 *Psychophysics* 70:291-305.

37 633 Mante V, Sussillo D, Shenoy KV, Newsome WT (2013) Context-dependent computation by  
38 634 recurrent dynamics in prefrontal cortex. *Nature* 503:78-84.

39 635 Matell MS, Meck WH (2000) Neuropsychological mechanisms of interval timing behavior.  
40 636 *BioEssays* 22:94-103.

41 637 Matthews WJ, Meck WH (2016) Temporal Cognition: Connecting Subjective Time to Perception,  
42 638 Attention, and Memory. *Psychol Bull* 142:865-890.

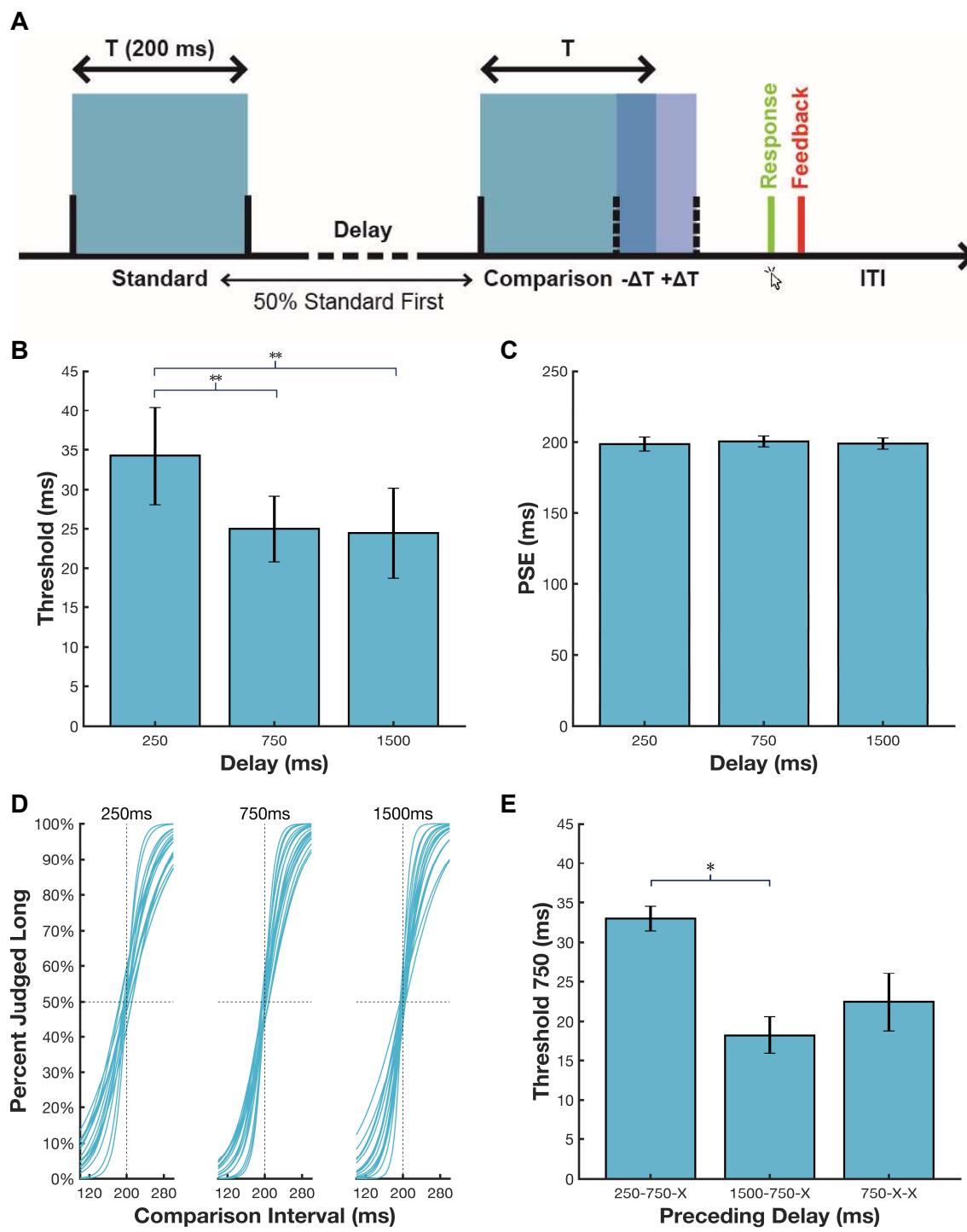
43 639 Mauk MD, Buonomano DV (2004) The Neural Basis of Temporal Processing. *Ann Rev Neurosci*  
44 640 27:307-340.

45 641 Meck WH (1996) Neuropharmacology of timing and time perception. *Cog Brain Res* 3:227-242.

46 642 Meck WH, Ivry RB (2016) Editorial overview: Time in perception and action. *Current Opinion in  
47 643 Behavioral Sciences* 8:vi-x.

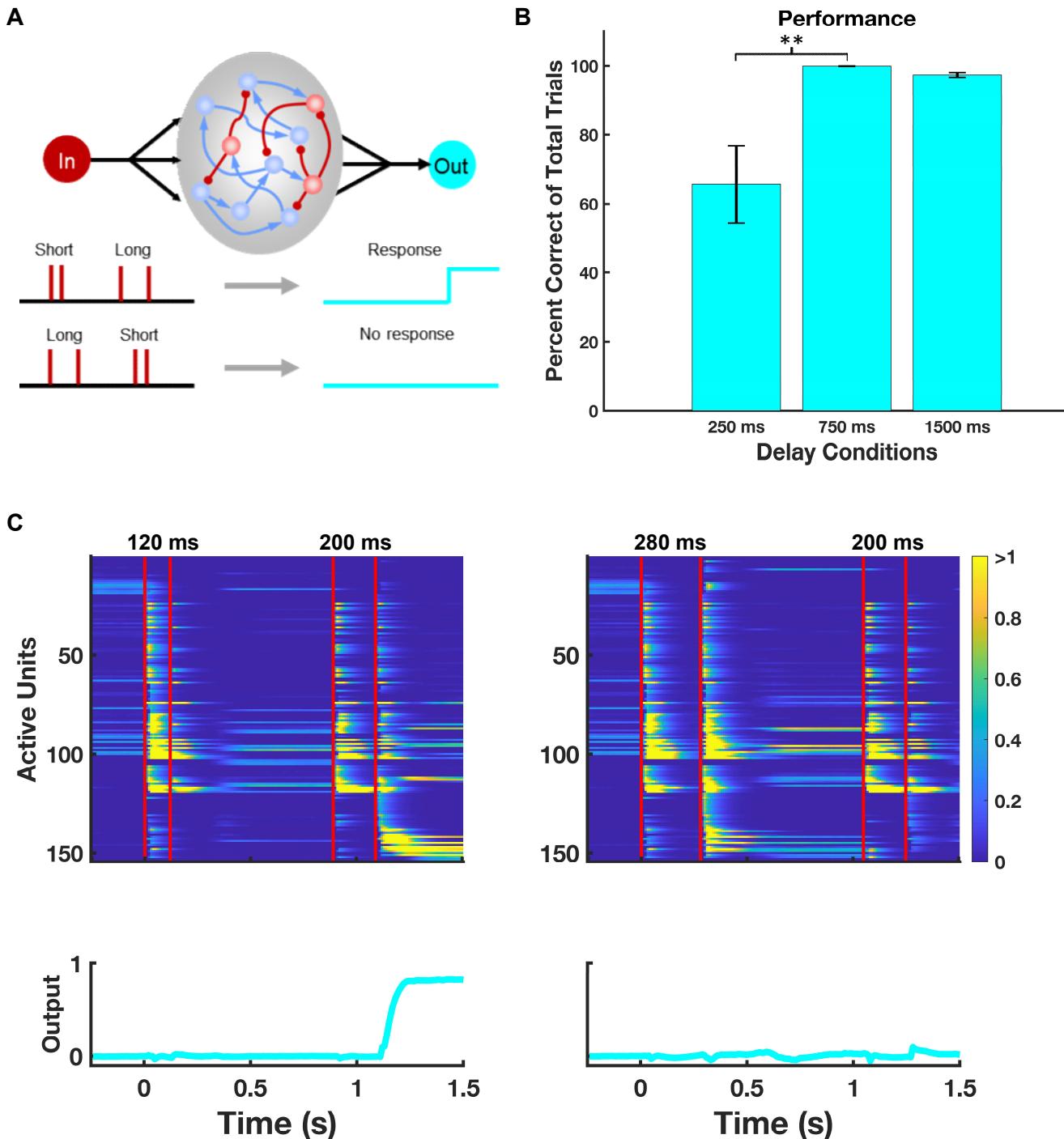
- 1  
2  
3
- 4 644 Merchant H, Harrington DL, Meck WH (2013) Neural Basis of the Perception and Estimation of  
5 645 Time. Annual Review of Neuroscience 36:313-336.
- 6 646 Miller EK, Erickson CA, Desimone R (1996) Neural mechanisms of visual working memory in  
7 647 prefrontal cortex of the macaque. J Neurosci 16:5154-5167.
- 8 648 Monk KJ, Allard S, Hussain Shuler MG (2020) Reward Timing and Its Expression by Inhibitory  
9 649 Interneurons in the Mouse Primary Visual Cortex. Cerebral Cortex 30:4662-4676.
- 10 650 Motanis H, Seay MJ, Buonomano DV (2018) Short-Term Synaptic Plasticity as a Mechanism for  
11 651 Sensory Timing. Trends in Neurosciences 41:701-711.
- 12 652 Namboodiri VM, Mihalas S, Marton T, Hussain Shuler M (2014) A general theory of  
13 653 intertemporal decision-making and the perception of time. Frontiers in Behavioral  
14 654 Neuroscience 8.
- 15 655 Nobre AC, van Ede F (2018) Anticipated moments: temporal structure in attention. Nature  
16 656 Reviews Neuroscience 19:34.
- 17 657 Orhan AE, Ma WJ (2019) A diverse range of factors affect the nature of neural representations  
18 658 underlying short-term memory. Nature Neuroscience 22:275-283.
- 19 659 Oshio K, Chiba A, Inase M (2008) Temporal filtering by prefrontal neurons in duration  
20 660 discrimination. Eur J Neurosci 28:2333-2343.
- 21 661 Paton JJ, Buonomano DV (2018) The Neural Basis of Timing: Distributed Mechanisms for  
22 662 Diverse Functions. Neuron 98:687-705.
- 23 663 Quintana J, Fuster JM (1992) Mnemonic and predictive functions of cortical neurons in a  
24 664 memory task. Neuroreport 3:721-724.
- 25 665 Rainer G, Rao SC, Miller EK (1999) Prospective Coding for Objects in Primate Prefrontal Cortex. J  
26 666 Neurosci 19:5493-5505.
- 27 667 Rigotti M, Barak O, Warden MR, Wang X-J, Daw ND, Miller EK, Fusi S (2013) The importance of  
28 668 mixed selectivity in complex cognitive tasks. Nature 497:585-590.
- 29 669 Rose GJ (2014) Time computations in anuran auditory systems. Frontiers in Physiology 5.
- 30 670 Sadibolova R, Sun S, Terhune DB (2021) Using adaptive psychophysics to identify the neural  
31 671 network reset time in subsecond interval timing. Experimental Brain Research.
- 32 672 Shuler MG, Bear MF (2006) Reward timing in the primary visual cortex. Science 311:1606-1609.
- 33 673 Spencer RM, Karmarkar U, Ivry RB (2009) Evaluating dedicated and intrinsic models of temporal  
34 674 encoding by varying context. Philos Trans R Soc Lond B Biol Sci 364:1853-1863.
- 35 675 Stokes MG (2015) 'Activity-silent' working memory in prefrontal cortex: a dynamic coding  
36 676 framework. Trends in Cognitive Sciences 19:394-405.
- 37 677 Treisman M (1963) Temporal discrimination and the indifference interval: implications for a  
38 678 model of the 'internal clock'. Psychol Monogr 77:1-31.
- 39 679 van Ede F, Niklaus M, Nobre AC (2017) Temporal Expectations Guide Dynamic Prioritization in  
40 680 Visual Working Memory through Attenuated  $\alpha$  Oscillations. The Journal of Neuroscience  
41 681 37:437-445.
- 42 682 Wang X-J (2001) Synaptic reverberation underlying mnemonic persistent activity. Trends  
43 683 Neurosci 24:455-463.
- 44 684 Xu M, Zhang S-y, Dan Y, Poo M-m (2014) Representation of interval timing by temporally  
45 685 scalable firing patterns in rat prefrontal cortex. Proceedings of the National Academy of  
46 686 Sciences 111:480-485.
- 47 687  
48 688  
49 689  
50 690  
51 691  
52 692  
53 693  
54 694  
55 695  
56 696  
57 697  
58 698  
59 699  
60 700  
61 701  
62 702  
63 703  
64 704  
65 705

- 1  
2  
3  
4 687 Yang GR, Wang X-J (2020) Artificial Neural Networks for Neuroscientists: A Primer. *Neuron*  
5 688 107:1048-1070.  
6 689 Yang GR, Joglekar MR, Song HF, Newsome WT, Wang X-J (2019) Task representations in neural  
7 690 networks trained to perform many cognitive tasks. *Nature Neuroscience*.  
8 691  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65

**FIGURE 1**

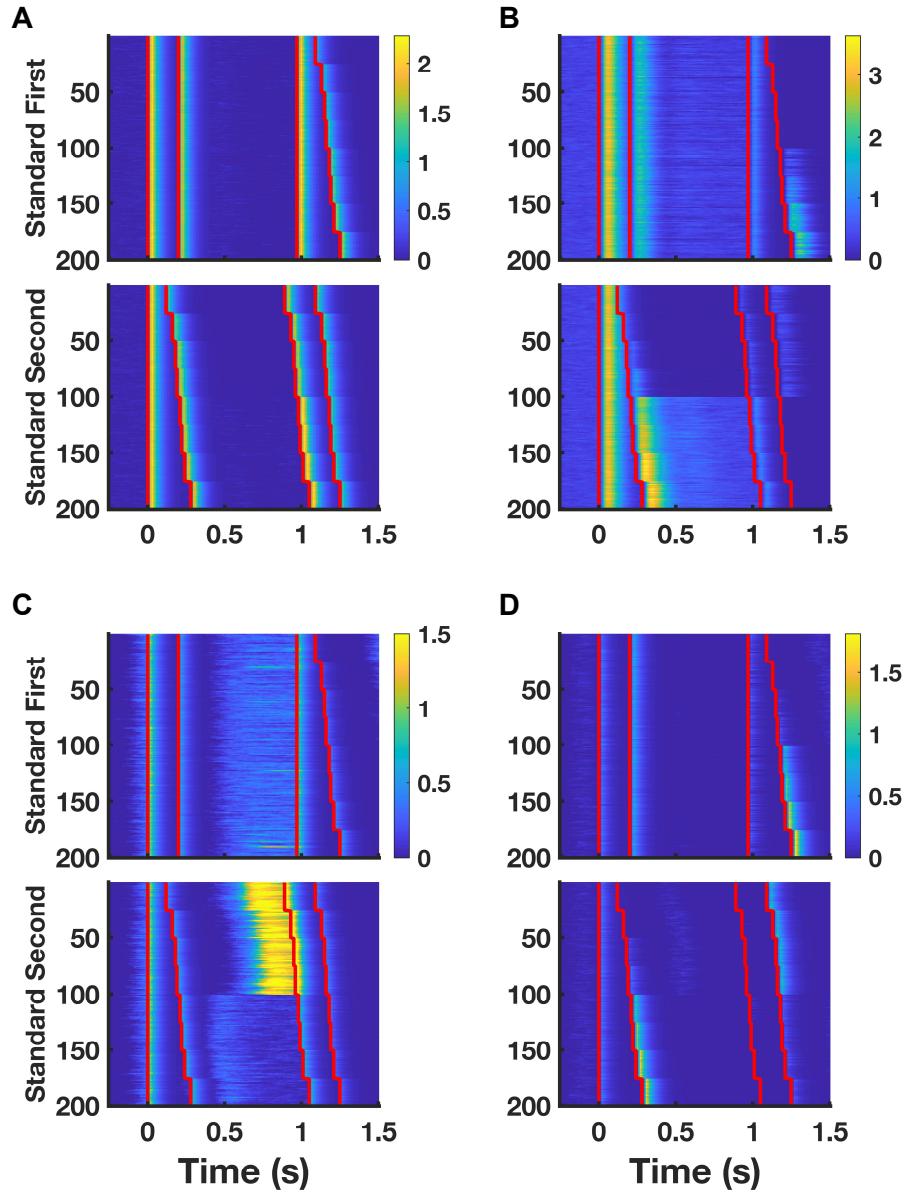
**Figure 1.** Psychophysics. (A) Schematic illustration of the task. Each trial consisted of two auditory intervals bounded by 15 ms tones (standard interval = 200 ms). The inter-stimulus delay was either 250 ms, 750 ms, or 1500 ms. Participants responded by indicating whether the second interval was shorter or longer than the first. (B) Threshold data for all participants by delay condition. The interval discrimination threshold was significantly higher for the 250 ms interstimulus delay condition compared to the 750 and 1500 ms conditions ( $p < 0.01$ ). There was no significant difference between the 750 and 1500 ms condition. (C) Point of subjective equality (PSE) data for all participants. There was no significant difference between the average PSE between the three delay types. (D) Fitted psychometric functions for all 18 subjects on the 250 ms, 750 ms, and 1500 ms delay conditions. (E) Threshold data for the 750 ms delay condition separated by whether it was preceded by a 250 ms block, a 1500 ms block, or no blocks. A significant increase in threshold for the 750 ms delay condition was observed when it was preceded by the 250 ms delay condition ( $p < 0.05$ ).

**FIGURE 2**



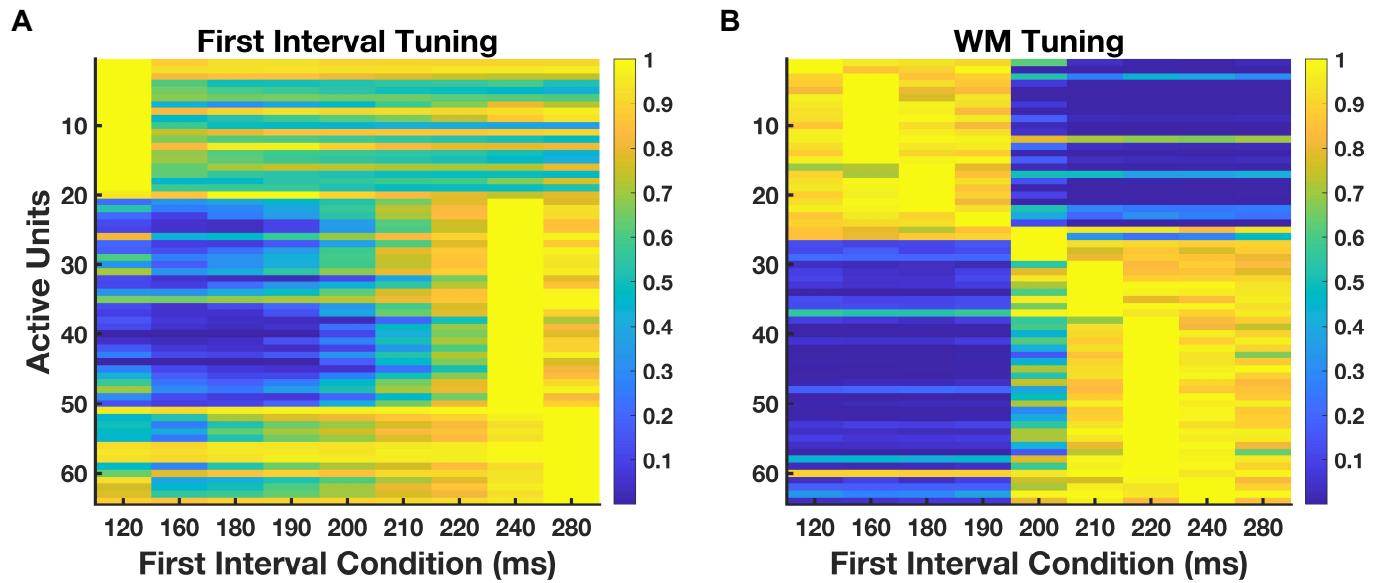
**Figure 2.** (A) Schematic of RNN model. The input unit projects to the RNN composed of 256 units (204 excitatory, 52 inhibitory). The recurrent units project to the output unit. The network is trained to produce a motor response for trials in which the second interval is longer than the first interval following the offset of the second interval. (B) Performance. Percent correct trials for five networks on three different delay conditions. Networks were trained on the 750 ms condition. There was a significant decrease in performance in the 250 compared to the 750 ms condition. ( $p < 0.01$ , Wilcoxon rank-sum test). (C) Sample population activity of a trained RNN tested on two different trial conditions. Red bars indicate input pulses. Left panel shows network activity for the short interval (120 ms) followed by the standard interval (200 ms) case, and right panel shows network activity for long-standard case. Cyan plots represent activity of the output unit.

**FIGURE 3**



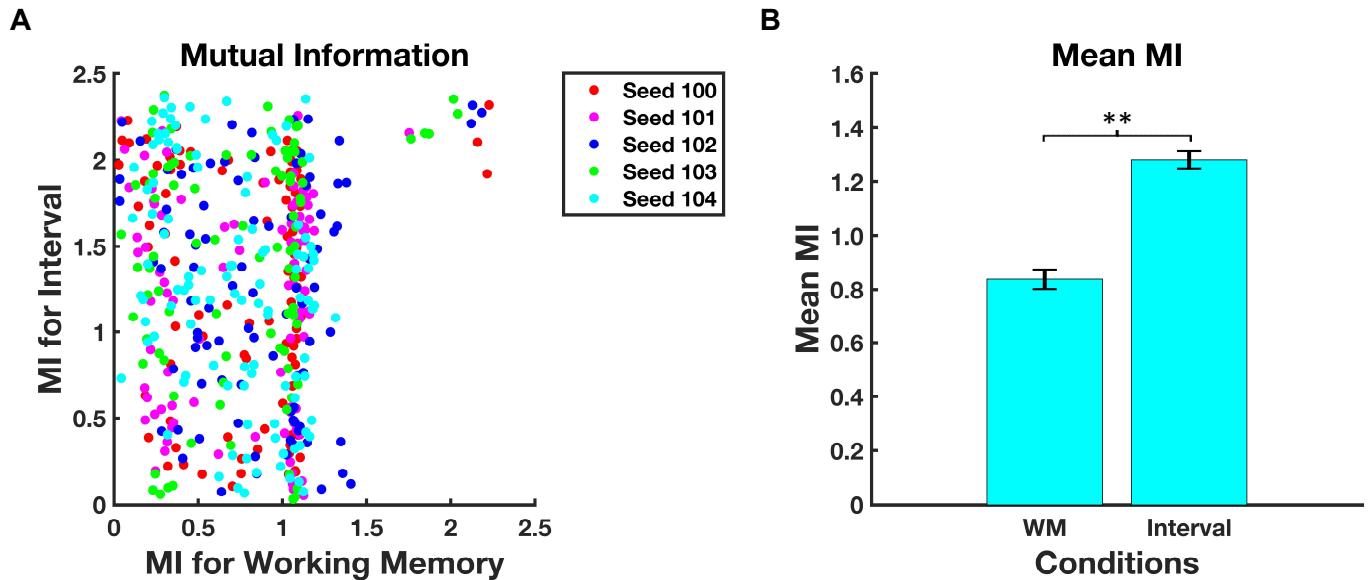
**Figure 3.** Four example RNN units. For each panel, the top plot shows activity for trials in which the standard interval (200 ms) is shown first, and the bottom plot shows activity for trials in which the standard is second. (A) A sensory unit that responds consistently to each input pulse. (B) An interval-tuned unit that exhibits progressively stronger responses to longer intervals. The unit also encodes long intervals during the WM epoch. (C) WM unit that exhibits approximately categorical response to short-first intervals. (D) Interval-tuned unit that preferentially responds to long intervals during the sensory epoch.

**FIGURE 4**



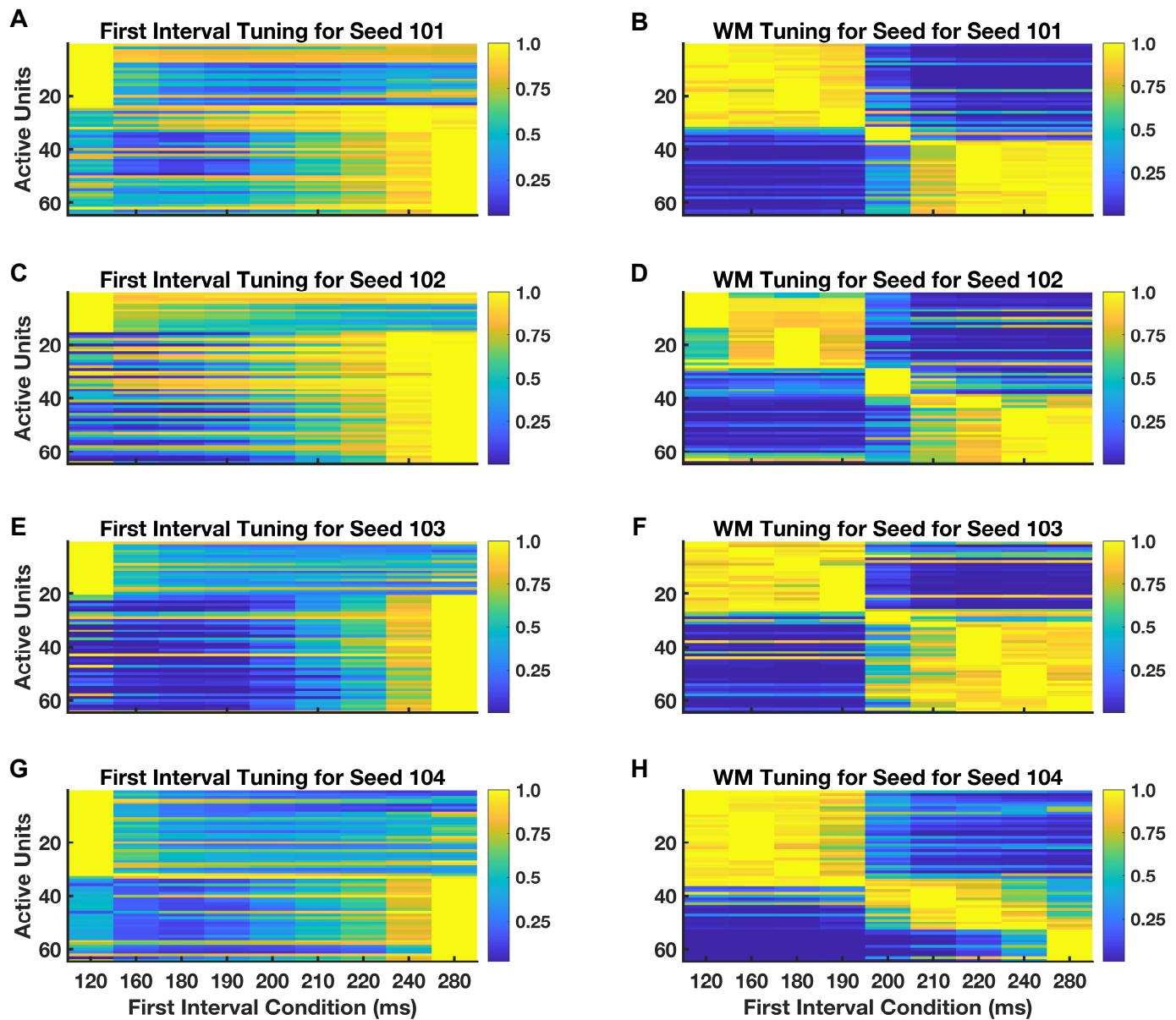
**Figure 4** (A) Interval tuning to first interval (sensory epoch). Each row represents a single unit's tuning function across all nine possible first interval conditions during the sensory epoch. The units are sorted in order of preferred interval tuning. Only the top quartile of active units is shown. (B) Interval tuning during the WM epoch. Note that in contrast to interval tuning during the sensory epoch, most units categorically encode short (<200 ms) or long (>200 ms) intervals.

**FIGURE 5**



**Figure 5** (A) Scatter plot of the mutual information (in bits) that each active units contains about the nine intervals during the sensory and WM epochs. Note that the overrepresentation of WM MI values that approximate 1 reflect the categorical encoding of short versus long intervals. (B) Mean mutual information for all data points shown in A. There is significantly greater mutual information during the sensory epoch compared to the WM epoch ( $p < 0.01$ , Wilcoxon rank-sum test).

## SUPPLEMENTARY FIGURE 1



**Supplementary Figure 1.** Plots in the style of Figure 4, shown for the four other trained networks.