

# Mice and rats fail to integrate exogenous timing noise into their time-based decisions

Dilara Berkay<sup>1</sup> · David Freestone<sup>2,3</sup> · Fuat Balci<sup>1</sup>

Received: 23 May 2016/Revised: 27 August 2016/Accepted: 3 September 2016/Published online: 19 September 2016  
© Springer-Verlag Berlin Heidelberg 2016

**Abstract** Endogenous timing uncertainty results in variability in time-based judgments. In many timing tasks, animals need to incorporate their level of endogenous timing uncertainty into their decisions in order to maximize the reward rate. Although animals have been shown to adopt such optimal behavioral strategies in time-based decisions, whether they can optimize their behavior under exogenous noise is an open question. In this study, we tested mice and rats in a task that required them to space their responses for a minimum duration (DRL task) in different task conditions. In one condition, the minimum wait time was fixed, whereas in other conditions minimum wait time was a Gaussian random variable. Although reward maximization entailed waiting longer with added exogenous timing variability, results indicated that both mice and rats became more impulsive and deviated from optimality with increasing levels of exogenous noise. We introduce a reward-rate-dependent sampling function to SET to account for optimal performance in noiseless and suboptimal performance in noisy environments.

**Keywords** DRL · Exogenous noise · Optimality · Temporal decision-making

## Introduction

Both human and nonhuman animals are capable of using temporal information adaptively in decision-making (Balci et al. 2011; Buhusi and Meck 2005). Although accurate on average, interval timing has limited precision reflected in the trial-to-trial variability in timed responses. Since in several scenarios, the optimal decision strategy depends on the level of variability in the resultant timed response distributions, it becomes important to monitor the level of this imprecision. This variability observed in behavioral responses is referred to as endogenous timing uncertainty since it mainly originates from the processing dynamics of the endogenous timing mechanism (e.g., Balci and Simen 2016; Gibbon et al. 1984; Killeen et al. 1997; Matell and Meck 2004; Simen et al. 2011, 2013).

Constituting one of the prominent psychophysical signatures of interval timing, response time variability has a well-defined relationship to the target durations; the standard deviation of timed responses is a constant proportion of the average timed responses ( $\sigma/\mu = CV$ ), which is referred to as the scalar property (e.g., Buhusi and Meck 2005; Gibbon 1977). Thus, the scalar property of interval timing results in a scale-invariant measure of timing uncertainty for each individual. As the reward-maximizing decision strategies often depend on the level of endogenous uncertainty, estimating the temporal variability characteristics of decisions is crucial for their adaptiveness (Balci et al. 2009, 2011).

The degree to which humans and other animals can incorporate their timing uncertainty into their time-based

---

**Electronic supplementary material** The online version of this article (doi:10.1007/s10071-016-1033-y) contains supplementary material, which is available to authorized users.

---

✉ Fuat Balci  
fbalci@ku.edu.tr;  
<http://mysite.ku.edu.tr/fbalci/>

<sup>1</sup> Department of Psychology, Koç University, Rumelifeneri Yolu, 34450 Sarıyer, Istanbul, Turkey

<sup>2</sup> Department of Psychology, Brown University, Providence, RI, USA

<sup>3</sup> Present Address: Department of Psychology, Bucknell University, Lewisburg, PA, USA

decisions is an emergent research theme in the interval timing and decision-making literatures. For instance, a recent line of evidence has exhibited that humans, rats, and mice can incorporate their timing uncertainty into their time-based decisions in a nearly normative fashion (for review see Balci et al. 2011) and can adopt optimal strategies in variants of temporal reproduction paradigms (e.g., Çavdaroglu et al. 2014; Jazayeri and Shadlen 2010; Freestone et al. 2015) as well as prospective and retrospective temporal discrimination paradigms (e.g., Akdoğan and Balci 2016; Balci et al. 2009; Coşkun et al. 2015).

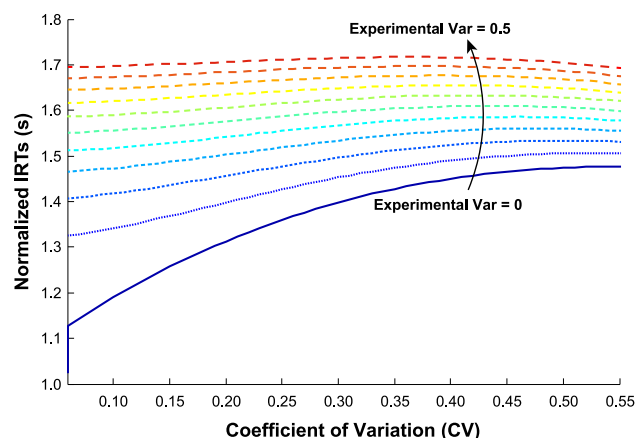
One of the interval timing paradigms that most directly captures the relationship between reward-rate-maximizing temporal decisions and the level of endogenous timing uncertainty is the differential reinforcement of low rates of responding (DRL). In this task, a subject waits a minimum amount of time since its previous response before emitting the next response. The response is reinforced every time the inter-response time (IRT) is greater than this minimum wait time (DRL schedule).

Importantly, due to variability in timed responses, even when a subject aims to respond later than the minimum wait time, sometimes it will respond earlier. This leads to a trade-off between reward probability and speed (IRT). As the IRT gets longer, the probability of receiving a reward increases, but so does the average time between rewards. Conversely, as the IRT gets shorter, the time between rewards decreases, but so does the probability of receiving a reward for each response. The reward-rate-maximizing (optimal) balance between these two time-variant factors depends on the subject's level of timing uncertainty because it helps determine the fraction of responses that are rewarded. Assuming inverse Gaussian-distributed IRTs, the reward rate for a  $T$  second interval, given average IRT  $t$  and shape parameter  $\lambda$  ( $t/\text{CV}^2$ ) is

$$\text{RR}(\hat{t}) = \hat{t}^{-1} \left( 1 - \text{IG}_{cdf} \left( T, \hat{t}, \hat{\lambda} \right) \right) \quad (1)$$

The value for  $t$  that maximizes this equation is the optimal mean IRT. We assume that timing uncertainty is a constraint. That is, an animal cannot choose what timing uncertainty to have. This results in an optimal performance curve (OPC) that maps different levels of timing uncertainty to optimal average IRTs (Fig. 1; solid blue line).

Other studies have used similar approaches to formulate an OPC based on timing variability (e.g., Akdoğan and Balci 2016; Balci et al. 2009; Coşkun et al. 2015; Jazayeri and Shadlen 2010). But one can see indications of optimally timed behavior even in studies that did not directly investigate this topic. For instance, in a previous study (Mita et al. 2009) monkeys were tested on a timing task where they needed to hold a response key for at least one of the target durations to receive reward in each trial. We



**Fig. 1** Optimal performance curve defined as a function of the level of timing uncertainty for different levels of experimentally added (exogenous) timing variability in the DRL task. The y axis shows the empirical IRTs normalized by the DRL schedule. The solid line is the OPC for the conventional DRL task, in which there is no variability in the minimum wait time

conducted the optimality analyses of the data of this study based on the descriptive statistics provided and the task parameters (see Fig. 1b of Mita et al. 2009). Our results indicated an almost perfect correspondence between optimal and empirical wait times for different target durations (see Online Source 1).

Overall, studies conducted in this area had a relatively limited scope in the sense that they, at the most, integrated the exogenous point probabilities of different temporal options into the computation of the optimal decision strategies. Due to the stochastic nature of the physical world, the durations of events also vary between different occasions. For instance, the replenishment time of a given food patch can exhibit variability with some degree of predictability (e.g., Gaussian-distributed replenishment times). A memory estimate of the duration of red traffic lights depends both on how well the brain can time (endogenous uncertainty) and the fact that the duration of red lights, themselves, are variable between intersections or times of day (exogenous uncertainty). In cases where temporal characteristics of different options exhibit trial-to-trial variability, reward maximization requires the agents to combine both sources of uncertainty into their decisions. In the DRL task, adding experimental noise to the schedule changes the optimal response times such that as the added variability increases so does the time point at which a subject should aim in order to maximize its reward rate (see Fig. 1). More specifically, exogenous noise leads to an upshifted and flatter relationship between optimal wait times and the level of endogenous timing uncertainty.

Hudson et al. (2008) asked whether humans can optimize their timed motor movements with experimentally added exogenous noise. In the task, participants were

required to touch a spatial target on a screen in a specific time window to attain reward. Responses made either earlier or later than the temporal window were penalized. Reward maximization in this task required participants to incorporate their level of timing uncertainty into their decisions. The experimenters also added exogenous noise to the response times by perturbing them with Gaussian noise. Their results revealed that participants could incorporate both endogenous and exogenous uncertainty levels in order to maximize reward, suggesting a remarkable level of adaptability in temporal risk assessment (at least in the domain of motor timing in the range of subsecond intervals).

In the current study, we investigated whether mice and rats can also integrate experimentally added exogenous noise with their own endogenous timing variability. For this purpose, we tested mice (Experiment 1) and rats (Experiment 2) at two different sites in the DRL task with and without experimentally added noise either to the DRL schedule (for mice) or the IRTs (for rats). Specifically, for the mouse experiment, constant noise was added to the DRL schedule, whereas for the rat experiment scalar noise was added to the IRTs. In both experiments we examined whether subjects increased their inter-response times with added exogenous noise as prescribed by optimality. These two experiments were conducted independently at Koç University (in mice) and Brown University (in rats); however, as the predictions and the findings of both experiments were equivalent at steady state, we decided to present the results together.

## Methods

### Experiment 1

#### *Subjects*

Eighteen C57BL/6J male mice (Koç University Animal Research Facility, Istanbul) were tested in the experiment. One mouse was discarded from the experiment due to health problems. The mice were approximately 10 weeks old by the time the experiment began. They were kept in a humidity and temperature-controlled colony room lit on a 12:12 light/dark cycle (lights off at 6:00 pm). The mice were tested during their light cycle, and they were maintained at 85 % of their free-feeding weight through caloric restriction. 0.01 cc of diluted liquid formula (Nestlé Nutrition Isosource Standard) was used as reinforcement. The mice had ad libitum access to water in their home cages. All experimental procedures were conducted in

accordance with Koç University Animal Research Local Ethics Committee guidelines.

#### *Apparatus*

In the experiment, nine operant chambers (MedAssociates, ENV-307W, dimensions 21.6 cm × 17.8 cm × 12.7 cm) were used. All operant chambers were placed in ventilated sound-attenuating cubicles. The ceiling and two opposing walls of the operant chambers were made of polycarbonate. The remaining two walls were made of metal. There were two retractable levers (ENV-312-2W) mounted on one side wall and three feeding hoppers (ENV-313W) located on the opposite wall. Only one lever (counterbalanced across subjects) and only one feeding hopper (i.e., middle) was kept active in the experiment. A dipper with a switchable arm delivered 0.01 cc of diluted liquid formula through the middle hopper as reinforcement. MED-PC IV software was used to control the experiment. All events were recorded with a 10-ms temporal resolution.

## Procedure

#### *Lever press training*

Mice were initially trained to lever press using the fixed ratio (FR) 1–fixed time (FT) 60-s schedule. Each mouse was trained with only one lever, either left or right, which was determined randomly prior to the training sessions. In each trial, the lever determined for the given subject was inserted into the box and the reward was delivered either after the subject pressed the lever or after the 60-s interval elapsed, whichever came first. Mice were trained on this schedule for at least three sessions. As none of the mice reached the learning criterion (40 lever presses in two consecutive sessions) in the last two sessions, the schedule was changed to FR-1 where the delivery of reward was contingent on a lever press. The learning criterion was held the same (i.e., emitting at least 40 lever presses in two successive sessions). FR-1 training lasted for on average three sessions.

#### *Testing phase 1*

Mice were tested in the DRL 12-s schedule. Each session started with the insertion of one lever into the operant chamber. Mice were required to space their responses by at least 12 s to receive reward. Each lever press reset the wait time and started a new trial. Early responses emitted before 12 s since the last response were not rewarded. Mice were tested in Phase 1 for 50 sessions to establish task learning.

### Testing phase 2

In Phase 2, we added constant noise ( $\gamma_+$ ) to the DRL schedule. Mice were divided into two groups which differed in terms of the level of noise added to the DRL schedule, which was  $\gamma_+ = 0.10$  for Group 1, and  $\gamma_+ = 0.20$  for Group 2. On each trial, a new DRL schedule,  $T_+$ , was calculated as follows:

$$T_+ \sim T + N(0, T\gamma_+) \quad (2)$$

where  $T$  is the DRL schedule (i.e., 12 s). If the IRT in a given trial was greater than the value of the  $T_+$  reward was presented and if it was lower than  $T_+$ , then the mouse received no reward. This phase lasted for 15 sessions.

### Data analysis

The data obtained from the last five sessions were used for the analyses. In the data, there were timed responses peaking near the DRL schedule, and seemingly untimed responses with very short (or very long) latencies. In order to capture these bimodal response distributions, we fit both the exponential-inverse Gaussian (Eq. 3) and the exponential-Gaussian mixture distribution functions (Eq. 4) to the obtained data using the maximum likelihood estimation method (Freestone et al. 2015).

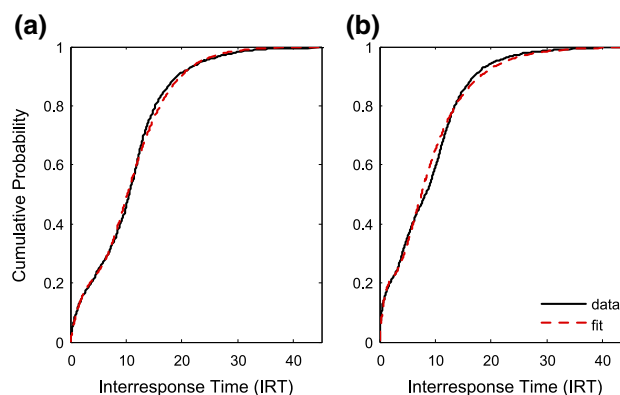
$$p(x|p, \mu, t, \lambda) = p(\mu e^{-\mu x}) + (1 - p) \sqrt{\frac{\lambda}{2\pi x^3}} e^{\left(\frac{-\lambda(x-t)^2}{2x^2}\right)} \quad (3)$$

$$p(x|p, \mu, t, \sigma) = p(\mu e^{-\mu x}) + (1 - p) \frac{1}{\sqrt{2\pi\sigma^2}} e^{\left(\frac{-(x-t)^2}{2\sigma^2}\right)} \quad (4)$$

The likelihood ratio was calculated for each subject to determine which mixture distribution provided a better fit. The average likelihood ratio of 65.19 indicated that inverse Gaussian fits better accounted for the data consistent with the human data presented in Çavdaroglu et al. (2014). Therefore, we used parameter estimates obtained from the exponential-inverse Gaussian distribution fits in all analyses.

A sample cumulative response distribution is presented along with the exponential-inverse Gaussian fit to it in Fig. 2 for one randomly chosen subject (see Online Source 2 for all individual fits). The exponential portion of these distributions represented untimed responses, whereas the inverse Gaussian portion represented the cases where the subjects made task-representative timed decisions. For each subject, the mean obtained from the inverse Gaussian fit was treated as the target IRT and the CV of the distribution was calculated to estimate the level of endogenous timing uncertainty.

The effect of introducing noise (i.e., Phase 1 vs. Phase 2) and the level of experimentally added noise (i.e., 0.10 vs.



**Fig. 2** Cumulative inter-response time distributions of a randomly chosen mouse for Phase 1 (a) and Phase 2 (b) along with the best-fitting exponential-inverse Gaussian functions. Smooth (black) lines represent empirical data and dashed (red) lines represent the best-fitting functions (color figure online)

0.20) on IRTs were investigated using a mixed-design ANOVA. Greenhouse–Geisser correction was applied when the sphericity assumption was violated. When needed, follow-up pairwise comparisons were conducted using  $t$  tests. To adjust  $P$  values for multiple comparisons, the Holm–Bonferroni correction (Holm 1979) was applied. Alpha was set to 0.05 (two-tailed) for all of the statistical analyses.

In this task, the optimal target IRT that maximizes the reward rate depends on both the endogenous and exogenous timing variability. We obtained each individual animal's optimal mean IRT by numerical simulation, calculating the reward rate for each possible target IRT given the total variability in that condition (i.e., by combining both endogenous and exogenous variability into a single estimate), and then finding the target IRT that maximized this reward rate. When the endogenous and exogenous variability measures are combined by adding variances, the pooled standard deviation equals  $\sqrt{(\gamma t)^2 + (\gamma_+ t)^2}$ , which can be rewritten as  $t\sqrt{\gamma^2 + \gamma_+^2}$ . Thus, the total coefficient on variation for a given noise level is

$$\gamma_{\text{tot}} = \sqrt{\gamma^2 + \gamma_+^2} \quad (5)$$

## Experiment 2

### Subjects

Twelve male Sprague–Dawley rats (Taconic Laboratories, Germantown, NY) were used in the experiment. A power supply problem caused unreliable results for six rats half-way through the experiment; thus only the data from the first six rats are reported. By the time the experiment

started, the rats were 8 weeks of age. The rats were kept in a colony room on a 12:12 light–dark cycle (lights off at 8:30 am). They were tested in their light cycle. During the experimental sessions, the rats were fed 45-mg Noyes pellets (Improved Formula A) as reinforcement. Water was available ad libitum in both the home cage and the testing chamber. All experimental procedures were conducted in accordance with Brown University Institutional Animal Care and Use Committee guidelines.

### Apparatus

Twelve operant chambers (Med Associates, dimensions  $25 \times 30 \times 30$  cm) were used in the experiment. Each chamber was contained in a sound-attenuating box with a fan for ventilation. Each operant chamber was equipped with a pellet dispenser (Med Associates, ENV-203) on the front wall that delivered the reward into a food cup. A head entry into this cup interrupted a photobeam (Med Associates, ENV-254). On both sides of the food cup, there were two retractable levers. On the opposite wall, a water bottle protruded into the chamber allowing ad libitum access to water during the session. A lick on the spout of the water bottle completed an electric circuit. MED-PC IV software controlled the experiments and recorded the data. The interruption of the photobeam and the completion of the lick and lever circuits were recorded in time event format with 10-ms accuracy.

### Procedure

#### Lever press training

Rats were trained to press the right lever in three 1-h sessions for two nights. The rats were placed in the box and the right lever was inserted. On the first day, reinforcement was presented on an FT 60-s schedule until the first lever

press and on an FR-1 schedule thereafter. On the following days, the rats were rewarded only on an FR-1 schedule.

#### Testing phase 1

The rats were tested in DRL 10-s schedule in three 1-h sessions per night, each separated by a 3-h period. At the beginning of each session, the right lever was inserted into the operant chamber and the house light was turned on. Responses separated by a minimum 10-s interval reset the time waited and delivered a reward, whereas responses given before the specified duration elapsed since the last response reset the wait time without reward delivery. Phase 1 lasted for six nights, for a total of 18 1-h sessions.

#### Testing phase 2

In the Phase 2, scalar noise was introduced to the IRTs. On each trial, the IRT that was used to determine reward was jittered to be

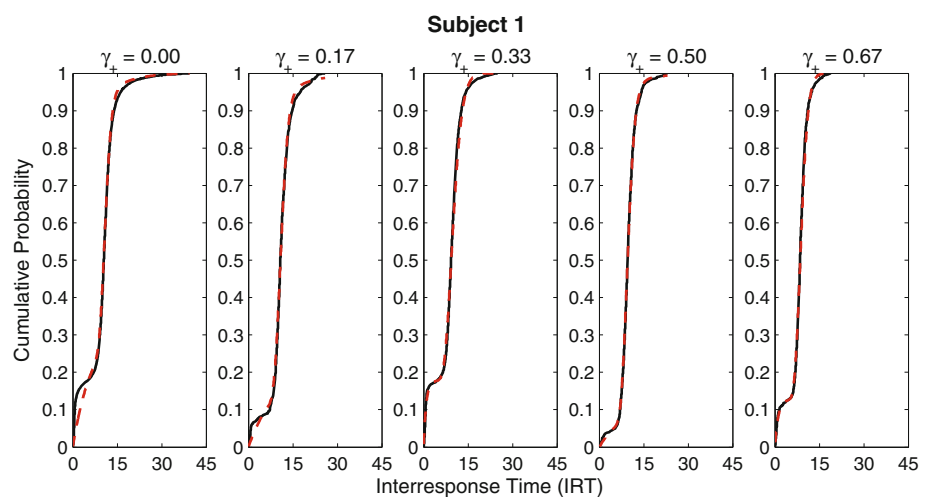
$$t_+ \sim t + N(0, t\gamma_+) \quad (6)$$

where  $t$  is the empirical IRT emitted on that trial and  $\gamma_+$  is the added noise level. The rat was rewarded if this new value was greater than the DRL schedule. Note that the noise was introduced symmetrically around the IRTs and thus it could either decrease or increase the IRT effective in a given trial. Rats were tested in four different conditions differing in the level of experimentally added noise. These phases were  $\gamma_+ = 0.33, 0.67, 0.17, 0.50$  in order of presentation. Each condition lasted at least six 1-h sessions.

### Data analysis

The last three nights of each phase were used in the analysis (nine 1-h sessions). Like observed in the mouse data, there were both timed and untimed responses in the rat

**Fig. 3** Cumulative inter-response time distributions of a randomly chosen rat for different noise level conditions and the best-fitting exponential–inverse Gaussian functions. Smooth (black) lines represent empirical data and dashed (red) lines represent the best-fitting functions (color figure online)





data. We fit the exponential–inverse Gaussian and exponential–Gaussian mixture distribution functions to the data using maximum likelihood estimation and compared the fit qualities. The average likelihood ratio was 42.62, favoring the exponential–inverse Gaussian distribution. The cumulative exponential–inverse Gaussian density function fit to the response distribution of a randomly chosen subject is presented in Fig. 3 (see Online Source 2 for all individual fits). For each subject, the mean obtained from the inverse Gaussian fit was treated as the target IRT, and the CV of the distribution was calculated to estimate the level of endogenous timing uncertainty.

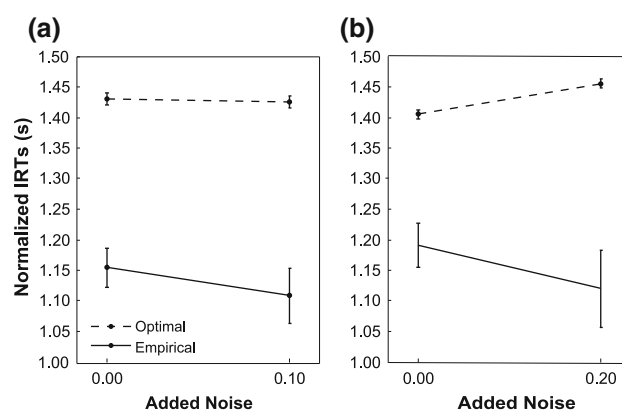
Changes in IRTs across different noise levels were investigated using repeated-measures ANOVA. Greenhouse–Geisser correction was applied when the sphericity assumption was violated. When needed, follow-up pairwise comparisons were conducted using *t* tests. To adjust *P* values for multiple comparisons, the Holm–Bonferroni correction was applied. Alpha level was set to 0.05 (two-tailed) for all of the statistical analyses. As we did in the mouse data, we calculated the optimal target IRT that would maximize the reward rate for each subject and phase (see Eq. 1).

## Results

### Experiment 1

In order to determine whether subjects reached steady state performance, we evaluated the change in the IRTs that are included in the analyses using linear regression. We then compared the individually obtained slope estimates with the value of 0. The results indicated that subjects' performance was stable across trials included in the analyses, [Phase 1:  $t(16) = 0.45$ ,  $P = 0.66$ ; Phase 2:  $t(16) = 0.02$ ,  $P = 0.99$ ].

Figure 4 shows the normalized mean optimal and empirical IRTs of the mice in different groups for Phase 1 and Phase 2. We examined the normalized empirical IRTs and asked whether there was a significant difference between phases before and after adding exogenous noise, and whether the introduction of noise had a differential effect on behavioral performance in different noise level groups. Our results showed that, in both groups, the mice decreased their mean inter-response times when exogenous variability was added [ $F(1,15) = 4.98$ ,  $P = 0.04$ ; Phase 1 (mean  $\pm$  SE =  $1.17 \pm 0.02$ ) to Phase 2 (mean  $\pm$  SE =  $1.11 \pm 0.04$ )]. This runs counter to the predictions of optimality because mice should have either had the same inter-response time (in Group 1; although they were still suboptimal), or increased their mean inter-response time (Group 2). This result is not due to the mice changing their own endogenous variability as the CV



**Fig. 4** Average optimal and empirical IRTs normalized by the DRL schedule (i.e., 12 s) obtained before (Phase 1) and after (Phase 2) introducing exogenous noise for Group 1 (a) and Group 2 (b) (Mice). Error bars represent the standard errors

estimates obtained before (mean  $\pm$  SE =  $.34 \pm .01$ ) and after (mean  $\pm$  SE =  $.34 \pm .02$ ) introducing exogenous noise were not different [ $F(1,15) = 0.06$ ,  $P = 0.81$ ].

Our analyses comparing the optimal and the empirical IRTs revealed that subjects' empirical IRTs were shorter than the corresponding optimal IRTs in both Phase 1 [ $F(1,15) = 138.56$ ] and Phase 2 [ $F(1,15) = 72.35$ ], both  $P$ s < 0.001. Independent of the added noise level, the deviation from optimality in the direction of impulsivity increased with the introduction of exogenous noise [ $F(1,15) = 9.26$ ,  $P = 0.008$ ].

Even though the subjects deviated from optimality in both phases, it is still possible that they took account of their endogenous uncertainty level and tracked the optimal response curve with a bias (Freestone et al. 2015). In order to investigate whether this was the case, we then regressed empirical IRTs on optimal IRTs across subjects separately for Phase 1 and Phase 2 and examined whether optimal IRTs predicted empirical IRTs. The results indicated that optimal IRTs was a significant predictor of empirical IRTs in Phase 1 where no experimental noise was introduced,  $\beta \pm$  SE =  $3.64 \pm 1.64$ ,  $P = 0.03$ . This is a very robust finding shown in multiple studies (Balci et al. 2011; Çavdaroglu et al. 2014). However, no significant relationship between the optimal and the empirical IRTs was found in Phase 2 for either the low noise level or the high noise level group (min  $P > 0.22$ ). These results suggest that the decision process that leads to optimality in temporally deterministic scenarios does not incorporate exogenous sources of uncertainty.

It is possible that, even in the phases where exogenous noise was introduced, subjects took account of only their endogenous uncertainty levels and not the level of the experimentally added exogenous noise. If this is the case, one would expect that the optimal IRTs calculated based on

only the endogenous uncertainty level of subjects to predict the empirical IRTs in the phases with experimentally added noise. However, when we tested this possibility, we did not observe any evidence favoring this claim as the optimal IRTs calculated based on the endogenous uncertainty levels did not predict the empirical IRTs (min  $P > 0.29$ ).

We also examined whether the exponential portion of the responses, which correspond to untimed responses that may be attributed to exploratory or other untimed behavioral tendencies, changed depending on the presence of experimentally added noise. Our results indicated a differential effect of introducing noise for different noise level groups [ $F(1,15) = 5.18$ ,  $P = 0.04$ ]. But this was only true for our  $\gamma_+ = 0.20$  group, who significantly increased the proportion of their untimed responses after the exogenous noise was presented [Phase 1: mean  $\pm$  SE =  $0.18 \pm 0.02$ , Phase 2: mean  $\pm$  SE =  $0.22 \pm 0.03$ ,  $P = 0.02$ ]. We did not observe such a change in  $\gamma_+ = 0.10$  group [Phase 1: mean  $\pm$  SE =  $0.25 \pm 0.03$ , Phase 2: mean  $\pm$  SE =  $0.27 \pm 0.03$ ,  $P = 0.44$ ]. We did not have a prior hypothesis about how added noise would affect these untimed responses, and our results do not support any particular theoretical viewpoint.

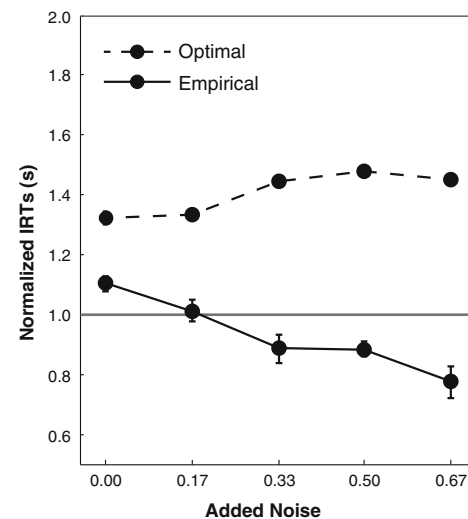
## Experiment 2

In order to evaluate whether subjects' performance stabilized in the sessions included in the analyses, we regressed the IRTs on the trial numbers and investigated whether there was a change in the IRTs as a function of trials. The comparison of individually obtained slope estimates did not differ from the value of 0, indicating that the IRTs were stabilized, all  $P$ s  $> 0.20$ .

The optimal and empirical IRTs for different noise levels are presented in Fig. 5.

The results of our analyses indicated that the subjects changed their empirical IRTs as function of the exogenous noise level [ $F(4,20) = 30.44$ ,  $P < 0.001$ ]. The results of CV comparison across conditions indicated that there was no linear or quadratic change across different noise level conditions [mean  $\pm$  SE =  $0.20 \pm 0.02$ ,  $0.15 \pm 0.01$ ,  $0.20 \pm 0.01$ ,  $0.17 \pm 0.01$ , and  $0.20 \pm 0.01$ , for increasing noise level conditions, respectively], and thus the change in empirical IRTs could not be accounted for by the change in the level of endogenous uncertainty (see Table 1) as a function of added noise level [linear:  $F(1,5) = 0.32$ ,  $P = 0.60$ ; quadratic:  $F(1,5) = 2.37$ ,  $P = 0.18$ ].

Next, we investigated the differences between optimal and empirical IRTs in different noise level conditions. Our results indicated that the empirical IRTs were significantly shorter than the optimal IRTs in all conditions, all  $P$ s  $< 0.001$ . Thus, the rats waited significantly shorter than



**Fig. 5** Average empirical and optimal IRTs normalized by the DRL schedule (i.e., 10 s) obtained from different exogenous noise level conditions (Rats). Error bars represent the standard errors

**Table 1** Mean and standard errors for the proportion of exponential distributions, the mean and lambda values obtained from the best-fitting exponential-inverse Gaussian mixture distribution functions and corresponding CV values for different noise level conditions

	$P(\text{Exp})$		$t$		$\lambda$		CV	
	$M$	$SE$	$M$	$SE$	$M$	$SE$	$M$	$SE$
$\gamma_+ = 0.00$	0.26	0.04	1.07	0.02	30.17	5.30	0.20	0.02
$\gamma_+ = 0.17$	0.29	0.08	1.01	0.03	47.33	3.64	0.15	0.01
$\gamma_+ = 0.33$	0.23	0.04	0.89	0.05	23.87	2.47	0.20	0.01
$\gamma_+ = 0.50$	0.38	0.11	0.88	0.02	33.57	3.67	0.17	0.01
$\gamma_+ = 0.67$	0.18	0.03	0.72	0.05	19.75	1.94	0.20	0.01

optimal in all noise level conditions. Additionally, as the noise level increased, so did the ratio of the normalized empirical IRTs to the normalized optimal IRTs pointing at a greater deviation from optimality [ $F(4,20) = 91.44$ ,  $P < 0.001$ ].

We regressed the empirical IRTs on the optimal IRTs obtained from each noise level ( $\gamma_+ = 0, 0.33, 0.66, 0.16, 0.50$ ) separately in order to examine whether the rats monitored their endogenous timing uncertainty and whether their optimal IRTs predicted their empirical IRTs. We found that optimal IRTs did not predict the empirical IRTs in any of the noise level conditions, all  $P$ s  $> 0.20$ .

Next, in order to examine whether rats simply disregarded experimentally added noise, we calculated the optimal IRTs for all conditions without taking account of exogenous noise level and investigated whether these optimal IRTs predicted empirical IRTs. However, the optimal IRTs calculated based on the endogenous

uncertainty levels did not predict the empirical IRTs (all  $P$ s  $> 0.05$ ).

Table 1 presents the proportion of the exponential portion of the best-fitting mixture distributions for each subject at each noise level. The exponential portion of the responses was compared across different conditions to reveal any possible changes in proportion of the untimed responses depending on the noise level. The results showed no significant difference,  $F(4,20) = 1.439$ ,  $P = 0.26$ .

## Discussion

In this study we investigated whether mice and rats can incorporate the exogenous timing noise into their temporal decisions in a normative fashion. In the first phase of each experiment, animals were tested without being exposed to exogenous noise, whereas in subsequent phase(s) exogenous Gaussian timing noise was introduced. For mice, we added constant noise to the DRL schedule, and for rats we added scalar noise to the IRTs. These two manipulations are empirically comparable, since they make the same predictions for the steady state data. In both cases, the optimal strategy dictates increasing target IRT as a function of both endogenous and exogenous uncertainty for reward-rate maximization. We examined whether animals changed their target wait times in the direction prescribed by the optimal strategy. Our results indicated that both mice and rats were sensitive to added noise, as they changed their target IRTs when external noise was introduced. However, this change was in the opposite direction to what was predicted by the optimal model; subjects decreased their mean wait time rather than increasing it.

Optimal performance in these conditions requires subjects first to learn the exogenous noise level and integrate this knowledge into their statistical model of the temporal structure of the environment. Specifically, animals should estimate  $\gamma_{\text{tot}}$  (Eq. 5) to build an accurate model of the temporal statistics of the task. Our behavioral results suggest that animals either cannot form these estimates, or cannot integrate them into their timing behavior.

This implication is theoretically important. All current models of timing of which we are aware do not separate endogenous from exogenous noise. In these models, the remembered intervals are the times of events in the world, like the time of the reward. And these events are retrieved in some form to guide behavior. When we added exogenous noise, the total variability of the reinforcement times is  $\gamma_{\text{tot}} = \sqrt{\gamma^2 + \gamma_+^2}$ . If animals cannot separate endogenous from exogenous sources of variability—and the current models suggest they cannot—then we should see a

corresponding rise in the variability of their own inter-response times. But we did not, even over a fourfold change in the variability of the reinforcement times. The result that animals do not optimally integrate endogenous and exogenous sources of variability, coupled with the result that their mean inter-response times do not increase, could imply that they do not confound different sources of variability and can discriminate them.

A possible explanation for the divergence from the predictions of the optimal model is that subjects are actually maximizing a different function. For instance, it might be the case that subjects are actually incorporating an urgency signal into their judgments and maximizing their gains given this parameter. Alternatively, the function that is being maximized by the subjects might contain temporal parameters that increase as a power function of objective time with an exponent higher than 1. Both of these cases would predict OPCs that are left-shifted compared to the OPC formulated in this paper. However, in these cases, too, the main conclusion that mice and rats deviate from optimality would hold as the model presented here is not intended to explain the data; it is intended to prescribe the reward-maximizing decision strategies.

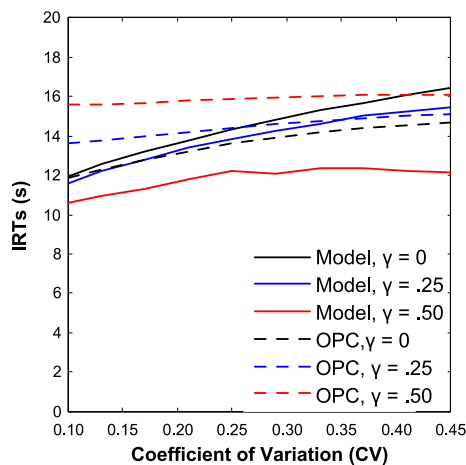
Another possible reason for earlier than optimal inter-response times could be the reduction in the overall reinforcement rate with the introduction of exogenous noise. This decrement in reinforcement rate could potentially jeopardize original response pattern/task representation such that subjects favor earlier responses resulting in deviations from optimality. However, even if this was the case, the main conclusion derived from the study still holds as the subjects clearly deviated from optimality with added noise. In the same vein, experimentally increased uncertainty might have facilitated an alternative strategy altogether. Consistent with this possibility, the ability of animals to integrate previously acquired knowledge regarding task parameters into choice behavior has also been found to differ between variable and fixed interval schedules which correspond to noisy and noise-free conditions, respectively. For instance, Tosun et al. (2016) showed that mice can integrate previously learned information regarding location, probabilities and fixed time intervals into timed decisions immediately and abruptly when the task conditions require them to do so. But when animals are trained independently on two different variable interval (VI) schedules and then two VI schedules are presented concurrently, instead of exhibiting traditional matching behavior (by allocating their responses between different options), they exhibit an exclusive preference for the richer option (Cerutti and Staddon 2004; Donahoe and Palmer 1994; Gallistel and Gibbon 2000).



Previous studies that evaluated the DRL performance within the framework of optimality offered primarily a descriptive account of optimal performance. Having observed clear deviations from optimality (particularly with the inclusion of exogenous timing noise), we offer a memory-based approach that could begin to explain both close to optimal and clearly suboptimal performance in the DRL task. The basic idea behind this account is aligned with the assumptions of the information processing variant of scalar expectancy theory (SET; Gibbon and Church 1981, 1992; Gibbon et al. 1984). SET assumes that the representations of time intervals are transferred to the memory when they signal a biologically critical outcome like reinforcement. In each subsequent trial, the subject draws a random sample from memory and uses this value

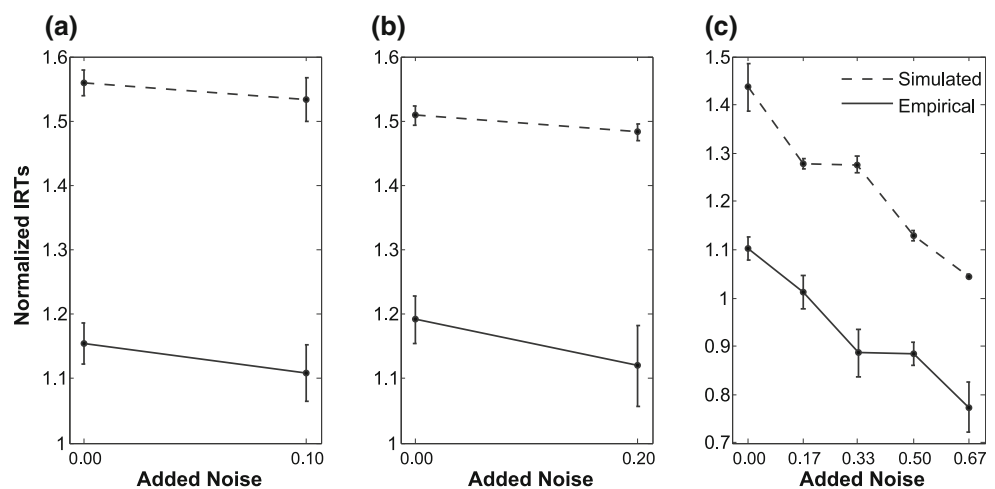
as a reference against which elapsing time is constantly compared.

According to this model, all remembered reinforcement times are equally likely to be sampled. In a simple variant on this idea, we suppose that the remembered intervals are sampled in proportion to their associated reinforcement rates (i.e.,  $1/t$ ). When exogenous noise is added, the subjects can be reinforced earlier than the DRL schedule. And since these intervals are associated with a higher reinforcement rate, they get sampled more frequently. When this idea is applied to DRL, it predicts the leftward shift in the data when exogenous noise is added, and it can do so in a way that mirrors the discrepancy between the data and the optimal model. Figure 6 shows simulations run for three different exogenous noise levels ( $\gamma_+ = 0, 0.25, 0.50$ ) along with the corresponding optimal performance curves. Note that the predictive power of this model with respect to DRL responding under different conditions is primarily related to the direction of the alterations in the behavior (i.e., leftward shift in IRTs as a function of exogenous noise) rather than in terms of the degree of overlap between the predicted and empirical data. Thus, as also seen in Fig. 6, the model predicts a larger leftward shift in IRTs as a function of added exogenous noise level, which fits well with the empirical output (see Fig. 7). The sampling method based on the reward rate is crucial for explaining this behavioral observation as other sampling methods such as uniform sampling (although it approximates the optimal performance curve in noiseless conditions) does not predict a leftward shift in IRTs with added exogenous noise. The model's prediction also tracks the optimal performance curve at all noise level conditions albeit with a bias. When we regressed the IRTs predicted by the model on those prescribed by the optimal performance curve for each exogenous noise level, we observed a significant relationship for each condition [all  $P$ s < 0.01]. Thus, the model



**Fig. 6** Model simulations (*dashed lines*) and optimal performance curves (*solid lines*) for  $\gamma_+ = 0, 0.25$ , and  $0.50$  noise level conditions. With the memory sampling function of  $1/t$ , the simulation output indicates a leftward shift in IRTs as a function of experimentally introduced exogenous noise

**Fig. 7** Empirical (*solid lines*) and simulated (*dashed lines*) IRTs for mice in Group 1 ( $\gamma_+ = 0.10$ ; **(a)**), mice in Group 2 ( $\gamma_+ = 0.20$ ; **(b)**) and rats **(c)**. IRTs are normalized by the schedule. *Error bars* represent the standard errors. Simulations are run for each subject's CV level in the corresponding condition



predicts a left-shifted performance curve that tracks the curvature of optimal performance.

Note that in Fig. 6, the estimated mean IRTs as a function of CVs are right-shifted in relation to the optimal performance curve in the condition where no exogenous noise is introduced. This is rarely observed in the data. There are many simple additions to the model that could account for this result (like including a term that accounts for the time spent consuming the reward), but our goal is to focus on what we consider to be the heart of the model: that the remembered time intervals are sampled in proportion to the reinforcement rate associated with them.

At the neurobiological level, this might relate to the role of dopaminergic function in signaling rewards as well as in memory formation and retention through synaptic modulation (Bach et al. 1999; Bernabeu et al. 1997; Gasbarri et al. 1994; Lisman and Grace 2005; Shohamy and Adcock 2010). For a fixed reward magnitude, the reinforcement signaling by dopaminergic system appears to dampen the utility of the reward hyperbolically as a function of time (Kobayashi and Schultz 2008), which captures the sampling function that is determined by  $1/t$  (the numerator becomes 0 when the IRT is not reinforced). Note that subject could still end up sampling values that are shorter than the schedule because the sampled value would serve as a target and its behavioral manifestation would still be subject to endogenous scalar timing noise.

A more model-based approach is that what animals learn on this task is the minimum reinforcement time. A Bayesian timer, for example, would learn an estimate of the minimum reinforcement time, and the distribution surrounding this estimate (that is, the endogenous uncertainty) would be by definition independent of the other reinforcement times (see Freestone and Church 2016). The minimum reinforcement time is of course a function of exogenously added noise: the higher the noise, the more likely a short inter-response time is to be reinforced. This may also account for a leftward shift in the mean inter-response times seen in our data. Future studies may be able to distinguish between the possibilities by changing the distributional characteristics of the exogenous noise. Additionally, in the current study subjects were always tested in noise-free condition first and noise was introduced only after they learned the task. Having established a steady performance might have disabled the subjects from adjusting to the experimentally manipulated noise conditions, which might have accounted for the obtained results. Future studies could investigate this issue by assigning subjects to different conditions in a counterbalanced order.

The datasets generated during and/or analyzed during the current study are available from the corresponding author on request.

**Acknowledgments** We would like to thank Filiz Çoşkun for her help in data collection.

**Funding** The mouse experiment was supported by the Scientific and Technological Research Council of Turkey (TÜBİTAK 1001 Grant No 111K402).

**Compliance with ethical standards**

**Conflict of interest** The authors declare that they have no conflict of interest.

**Ethical approval** All applicable international, national, and/or institutional guidelines for the care and use of animals were followed.

## References

- Akdogan B, Balci F (2016) Stimulus probability effects on temporal bisection performance of mice (*Mus musculus*). *Anim Cogn* 19:15–30. doi:[10.1007/s10071-015-0909-6](https://doi.org/10.1007/s10071-015-0909-6)
- Bach ME, Barad M, Son H, Zhuo M, Lu YF, Shih R, Mansuy I, Hawkins RD, Kandel ER (1999) Age-related defects in spatial memory are correlated with defects in the late phase of hippocampal long-term potentiation in vitro and are attenuated by drugs that enhance the cAMP signaling pathway. *Proc Natl Acad Sci USA* 96(9):5280–5285. doi:[10.1073/pnas.96.9.5280](https://doi.org/10.1073/pnas.96.9.5280)
- Balci F, Simen P (2016) A decision model of timing. *Curr Opin Behav Sci* 8:94–101. doi:[10.1016/j.cobeha.2016.02.002](https://doi.org/10.1016/j.cobeha.2016.02.002)
- Balci F, Freestone D, Gallistel CR (2009) Risk assessment in man and mouse. *Proc Natl Acad Sci USA* 106:2459–2463. doi:[10.1073/pnas.0812709106](https://doi.org/10.1073/pnas.0812709106)
- Balci F, Freestone D, Simen P, deSouza L, Cohen JD, Holmes P (2011) Optimal temporal risk assessment. *Front Integr Neurosci* 5:1–15. doi:[10.3389/fnint.2011.00056](https://doi.org/10.3389/fnint.2011.00056)
- Bernabeu R, Bevilacqua L, Ardenghi P, Bromberg E, Schmitz P, Bianchin M, Izquierdo I, Medina JH (1997) Involvement of hippocampal cAMP/cAMP-dependent protein kinase signaling pathways in a late memory consolidation phase of aversively motivated learning in rats. *Proc Natl Acad Sci USA* 94:7041–7046. doi:[10.1073/pnas.94.13.7041](https://doi.org/10.1073/pnas.94.13.7041)
- Buhusi CV, Meck WH (2005) What makes us tick? Functional and neural mechanisms of interval timing. *Nat Rev Neurosci* 6(10):755–765. doi:[10.1038/nrn1764](https://doi.org/10.1038/nrn1764)
- Çavdaroglu B, Zeki M, Balci F (2014) Time-based reward maximization. *Philos Trans R Soc B Biol Sci* 369:20120461. doi:[10.1098/rstb.2012.0461](https://doi.org/10.1098/rstb.2012.0461)
- Cerutti DT, Staddon JER (2004) Immediacy versus anticipated delay in the time-left experiment: a test of the cognitive hypothesis. *J Exp Psychol Anim Behav Process* 30(1):45–57. doi:[10.1037/0097-7403.30.1.45](https://doi.org/10.1037/0097-7403.30.1.45)
- Çoşkun F, Sayalı ZC, Gürbüz E, Balci F (2015) Optimal time discrimination. *Q J Exp Psychol* 68(2):381–401. doi:[10.1080/17470218.2014.944921](https://doi.org/10.1080/17470218.2014.944921)
- Donahoe JW, Palmer DC (1994) Selection in the experienced learner. In: Dorsel VP (ed) *Learning and complex behavior*. Allyn & Bacon, Boston, pp 112–113
- Freestone DM, Church RM (2016) Optimal timing. *Curr Opin Behav Sci* 8:276–281
- Freestone DM, Balci F, Simen P, Church RM (2015) Optimal response rates in humans and rats. *J Exp Psychol Anim Learn Cogn* 41(1):39–51. doi:[10.1037/xan0000049](https://doi.org/10.1037/xan0000049)
- Gallistel CR, Gibbon J (2000) Time, rate, and conditioning. *Psychol Rev* 107(2):289–344. doi:[10.1037//0033-295X.107.2.289](https://doi.org/10.1037//0033-295X.107.2.289)

- Gasbarri A, Packard MG, Campana E, Pacitti C (1994) Anterograde and retrograde tracing of projections from the ventral tegmental area to the hippocampal formation in the rat. *Brain Res Bull* 33(4):445–452. doi:[10.1016/0361-9230\(94\)90288-7](https://doi.org/10.1016/0361-9230(94)90288-7)
- Gibbon J (1977) Scalar expectancy theory and Weber's law in animal timing. *Psychol Rev* 84(3):279–325. doi:[10.1037/0033-295X.84.3.279](https://doi.org/10.1037/0033-295X.84.3.279)
- Gibbon J, Church RM (1981) Time left: linear versus logarithmic subjective time. *J Exp Psychol Anim Behav Process* 7(2):87–108. doi:[10.1037//0097-7403.7.2.87](https://doi.org/10.1037//0097-7403.7.2.87)
- Gibbon J, Church RM (1992) Comparison of variance and covariance patterns in parallel and serial theories of timing. *J Exp Anal Behav* 57:393–406. doi:[10.1901/jeab.1992.57-393](https://doi.org/10.1901/jeab.1992.57-393)
- Gibbon J, Church RM, Meck WH (1984) Scalar timing in memory. *Ann N Y Acad Sci* 423:52–77
- Holm S (1979) A simple sequentially rejective multiple test procedure. *Scand J Stat* 6(2):65–70. doi:[10.2307/4615733](https://doi.org/10.2307/4615733)
- Hudson TE, Maloney LT, Landy MS (2008) Optimal compensation for temporal uncertainty in movement planning. *PLoS Comput Biol* 4(7):e1000130. doi:[10.1371/journal.pcbi.1000130](https://doi.org/10.1371/journal.pcbi.1000130)
- Jazayeri M, Shadlen MN (2010) Temporal context calibrates interval timing. *Nat Neurosci* 13(8):1020–1026. doi:[10.1038/nn.2590](https://doi.org/10.1038/nn.2590)
- Killeen PR, Fetterman JG, Bizo LA (1997) Time's causes. *Adv Psychol* 120:79–132 (**Amsterdam**)
- Kobayashi S, Schultz W (2008) Influence of reward delays on responses of dopamine neurons. *J Neurosci* 28(31):7837–7846
- Lisman JE, Grace AA (2005) The hippocampal-VTA loop: controlling the entry of information into long-term memory. *Neuron* 46(5):703–713. doi:[10.1016/j.neuron.2005.05.002](https://doi.org/10.1016/j.neuron.2005.05.002)
- Matell MS, Meck WH (2004) Cortico-striatal circuits and interval timing: coincidence detection of oscillatory processes. *Brain Res Cogn Brain Res* 21(2):139–170. doi:[10.1016/j.cogbrainres.2004.06.012](https://doi.org/10.1016/j.cogbrainres.2004.06.012)
- Mita A, Mushiaki H, Shima K, Matsuzaka Y, Tanji J (2009) Interval time coding by neurons in the presupplementary and supplementary motor areas. *Nat Neurosci* 12(4):502–507. doi:[10.1038/nn.2272](https://doi.org/10.1038/nn.2272)
- Shohamy D, Adcock RA (2010) Dopamine and adaptive memory. *Trends Cogn Sci* 14(10):464–472. doi:[10.1016/j.tics.2010.08.002](https://doi.org/10.1016/j.tics.2010.08.002)
- Simen P, Balci F, Cohen JD, Holmes P (2011) A model of interval timing by neural integration. *J Neurosci* 31(25):9238–9253. doi:[10.1523/JNEUROSCI.3121-10.2011](https://doi.org/10.1523/JNEUROSCI.3121-10.2011)
- Simen P, Rivest F, Ludvig EA, Balci F, Killeen P (2013) Timescale invariance in the pacemaker-accumulator family of timing models. *Timing Time Percept* 1(2):159–188. doi:[10.1163/22134468-00002018](https://doi.org/10.1163/22134468-00002018)
- Tosun T, Gür E, Balci F (2016) Mice plan decision strategies based on previously learned time intervals, locations, and probabilities. *Proc Natl Acad Sci USA* 113(3):787–792. doi:[10.1073/pnas.1518316113](https://doi.org/10.1073/pnas.1518316113)