

Rodents monitor their error in self-generated duration on a single trial basis

Tadeusz Wladyslaw Kononowicz, Virginie van Wassenhove, Valérie Doyère

▶ **To cite this version:**

Tadeusz Wladyslaw Kononowicz, Virginie van Wassenhove, Valérie Doyère. Rodents monitor their error in self-generated duration on a single trial basis. Proceedings of the National Academy of Sciences of the United States of America, 2022, 119 (9), pp.e2108850119. 10.1073/pnas.2108850119 . hal-03587286

HAL Id: hal-03587286
https://hal.science/hal-03587286

Submitted on 25 Feb 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Rodents monitor their error in self-generated duration on a single trial basis

Authors: T. W. Kononowicz[1,2]*, V. van Wassenhove[3], V. Doyère[2]

Affiliations:
[1]Institute of Psychology, Polish Academy of Sciences, Warsaw, Poland
[2]Université Paris-Saclay, CNRS, Institut des Neurosciences Paris-Saclay (NeuroPSI), 91400 Saclay, France
[3]Cognitive Neuroimaging Unit, NeuroSpin, CEA, INSERM, CNRS, Université Paris-Saclay, 91191 Gif-sur-Yvette, France

*Correspondence to: Tadeusz W. Kononowicz
Email: t.w.kononowicz@icloud.com

Current address:
Instytut Psychologii, Polska Akademia Nauk
Ul. Jaracza 1
00-378 Warszawa
Polska

Author Contributions: TWK and VD conceived and planned the experiments. TWK carried out the experiments. TWK planned and carried out the statistical analyses. TWK, VD, and VvW contributed to the interpretation of the results. TWK prepared the first version of the manuscript. All authors provided critical feedback and helped to shape the research, analysis and manuscript.

Competing Interest Statement: No conflict of interest.
Classification: Social Sciences/Psychological and Cognitive Sciences, Biological Sciences/Neuroscience
Keywords: timing, time perception, error monitoring, metacognition

This PDF file includes:
Main Text
Figures 1 to 4

Abstract

A fundamental question in neuroscience is what type of internal representation leads to complex adaptive behavior. When faced with a deadline, individuals' behavior suggests that they represent the mean and the uncertainty of an internal timer to make near optimal time-dependent decisions. Whether this ability relies on simple trial-and-error adjustments, or whether it involves richer representations, is unknown. Richer representations suggest a possibility of error monitoring, that is, the ability for an individual to assess its internal representation of the world and estimate discrepancy in the absence of external feedback. While rodents show timing behavior, whether they can represent and report temporal errors in their own produced duration on a single trial basis is unknown. We designed a novel paradigm requiring rats to produce a target time interval and subsequently evaluate its error. Rats received a reward in a given location depending on the magnitude of their timing errors. During the test-trials, rats had to choose a port corresponding to the error magnitude of their just-produced duration to receive a reward. High choice accuracy demonstrates that rats kept track of the values of the timing variables on which they based their decision. Additionally, the rats kept a representation of the mapping between those timing values and the target value, as well as the history of the reinforcements. These findings demonstrate for the first time error monitoring abilities in evaluating self-generated timing in rodents. Together these findings suggest an explicit representation of produced duration and the possibility to evaluate its relation to the desired target duration.

Significance

Evaluating the richness of representations mapping animal's experience of time and space is a profound problem in neuroscience. Animals could base their behavior only on the recent rewards, but they also could rely on more complex representations of their own behavior. When faced with a deadline, individuals seem to adjust their decision based on the overall knowledge of their performance. We evaluated whether rats monitor timing of their own actions or adapt only based on recent rewards. Rats accurately reported the magnitude of their timing error in self-generated time intervals. The results for the first time show that rats track their own timing errors, deepening our understanding of error monitoring abilities in rodents and richness of their representation of elapsed time.

Introduction

In neuroscience, a fundamental question is how rich the internal representation of an individual's experience must be to yield adaptive behavior. Let us consider a hungry individual in need of finding food fast: the individual may adopt a trial-and-error foraging strategy to maximize reward, but may also, to maximize its efficiency, represent rich experiential variables, such as how much time it takes to reach a source of food. Both, representing elapsed time and monitoring its inherent uncertainty plays an important role in adaptive behavior, learning, and decision making (Balci et al., 2009). When representing these variables, the sources of uncertainty are both exogenous (stimuli-driven) and endogenous (neural implementation). The mapping of exogenous sources of temporal uncertainty has been well described in timing behavior: for instance, mice can adjust their behaviors to the width of the distribution of temporal intervals provided through external stimuli (Li & Dudman, 2013). On the other hand, the endogenous sources of uncertainty for time perception are less understood and more difficult to address.

Evidence that animals are sensitive and have access to the internal uncertainty of elapsed time comes from a task in which the individual must produce a required target duration using a lever-press or a key-press (Balci et al., 2009; Kheifets et al., 2017; Kheifets & Gallistel, 2012). In a task where individuals must produce an interval of fixed duration to obtain a reward (Figure 1A), a plausible strategy to maximize reward would be to set the produced duration to be longer than the required target duration so as to allow a margin of error (internal target duration, Balci et al., 2009). This is because the larger an individual's representational uncertainty, the larger the margin of error to maximize the reward. Consistent with this, studies have shown that the magnitude of error in produced intervals varies with the magnitude of temporal uncertainties (Çavdaroğlu, Zeki, Balci, 2014; Freestone et al., 2015), and participants with larger temporal uncertainty set larger margins of errors (Fig. 1B; Supp. Fig. 2; Balci et al., 2009; Freestone et al., 2015). The observed optimization of timing behavior begs the question of how rich the representation of elapsed time must be.

A trial-and-error strategy would predict that near optimal behavior can be parsimoniously explained by adaptation, so that timing behavior would fluctuate around the required duration. The representational view would predict that uncertainty and trial-to-trial errors are experiential variables used by the animals to monitor their timing behavior.

To settle the question of whether rodents can monitor their timing errors relative to their target on a trial-by-trial basis, we developed a novel task inspired by human work. Humans required to generate a time interval can also reliably report the magnitude of their errors and their sign (Kononowicz et al., 2019), i.e., they can evaluate by how much (magnitude) their generated duration was too short, or too long (sign), with respect to the target duration. Humans can also report how confident they are in their timing behavior (Akdoğan & Balci 2017). We tested here these temporal cognitive abilities in rats, which were required to produce a time interval and correctly report, in order to obtain a reward, the magnitude of their timing errors on some test-trials. We show that rats correctly reported the magnitude of their timing error, suggesting that their timing behavior uses explicit representations of time intervals together with their uncertainty around the internal target duration.

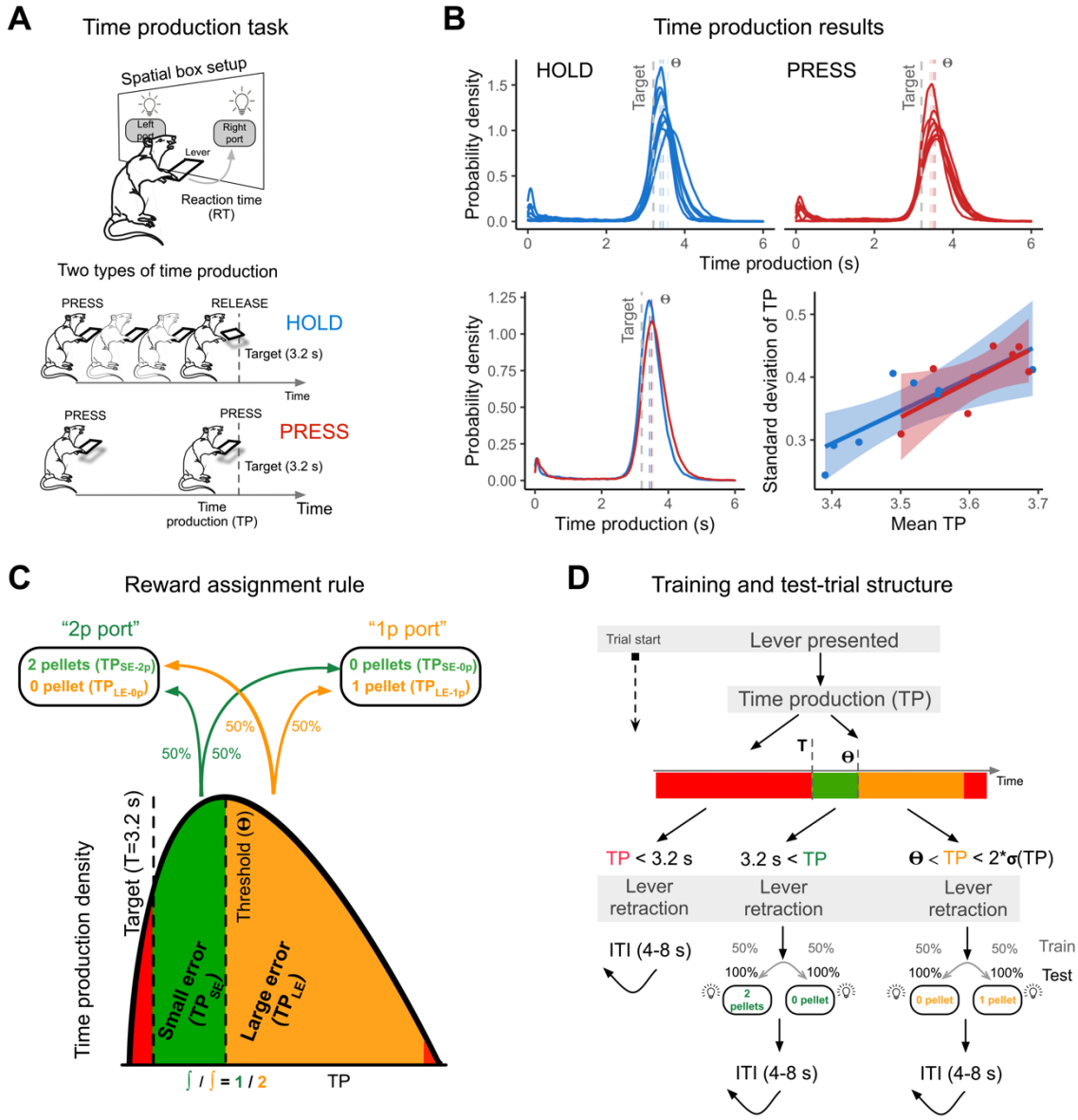


Figure 1. The time production task and error monitoring protocol.

(A) Schematic of a box arrangement with a lever available in the middle of the panel and reward ports on the left and right side of the lever. Reward availability was signaled by the port lit, depicted by the lightbulbs. Reward delivery was triggered by rats' nose poke in the reward port. Depending on the group assignment, rats had to either hold the lever pressed for a minimum of 3.2s (HOLD group), or press the lever twice with a minimal delay (3.2 s) between two presses (PRESS group).

(B) Time production (TP) performance, in error monitoring test sessions, follows Weber's law for both groups, with signatures of optimality. Upper panel depicts probability density functions over TPs for each individual rat in HOLD (blue) and PRESS (red) groups. Thresholds Θ (blue and red dashed lines HOLD and PRESS group, respectively) are plotted for each individual. Bottom left panel depicts average probability density functions over TPs for HOLD and PRESS groups superimposed. Note the distribution shift and width shrinkage for HOLD group. Bottom right panel: for each rat, μ(TP) is plotted against σ(TP). Both at the individual and at the group level the PRESS rats showed larger μ(TP) and σ(TP), visible as an upward right shift of the red curve. This pattern indicates that rats make their choices optimally, taking into account their level of TP variability. The results hold within each rat and across sessions (Supp. Mat. S3).

(C) Schematic depiction of how rewards were assigned to specific parts of TP distribution. Green color is used for 'small error' (SE) trials and orange color for 'large error' trials. Red color indicates TPs that were out of reward range. The arrows indicate probabilistic assignment of TP type (SE or LE) to left and right ports, on training trials. On test trials, the food-port assignments remained, but both ports were available and, thus, the amount of reward was driven by rat's choice.

(D) Schematic of a trial structure. From the top to bottom, the succession of task events is depicted. They alternate along TP-axis (color bar with red, green, and orange) and show different scenarios that are determined by rats' performance on TP in single trials. Inter-trial-interval (ITI) is the last event in a single trial sequence.

Results

Rats produce time intervals optimally

Two groups of rats were trained to produce a minimum Target duration (T = 3.2s) by pressing a lever twice to demarcate T (PRESS group), or by pressing a lever for T amount of time (HOLD group). Producing at least T provided access to a reward (Fig. 1A). All rats produced the required time interval T as all distributions of duration time productions (TP) peaked after T (Fig. 1B). Higher rewards were assigned to more accurate TPs to maximally exploit animal's temporal precision and use this temporal precision efficiently in a reward assignment rule (Fig. 1C). Consistent with Weber's law in duration reproduction (Gallistel & Gibbon, 2000; Gibbon, 1992; Buhusi & Meck, 2005), the standard deviation of TPs (σ(TP)) increased with the mean TP (μ(TP); Fig. 1B bottom panel, Supp. Mat. S3; Supp. Fig. 3; $t(586) = 15.4$, $p < 10^{-10}$). Comparable slopes in PRESS and HOLD groups were found (Fig. 1B, ΔAIC = 0.3, $p = 0.155$). As σ(TP) shrank over sessions (Supp. Fig. 3) and across individuals (Fig. 1B), the distance between T and μ(TP) shrank as well. This covariation strongly suggests a near optimal exploitation of timing uncertainty by the animals. Indeed, the animals displayed behavioral patterns suggesting that they adapt their mean responding on the basis of their timing uncertainty (Balci et al., 2009; Kheifets et al., 2017; Kheifets & Gallistel, 2012). This phenomenon has two plausible explanations: (1) the animals could adapt their TPs by trial-and-error (Simen & Balci, 2011; Luzardo et al., 2013), based exclusively on the memorized association of different rewards with different TP lengths (Jozefowiez et al., 2009); (2) alternatively, their behavior could be driven by model-based, explicit representations of TP and their relation to internal target duration (Tosun et al., 2016), which implies a tracking of their temporal error on a single trial basis. Hence, we explore the two possibilities below.

Rats choose accurately on test-trials, expressing error monitoring

To assess temporal error monitoring, we developed a novel behavioral paradigm in which, on each trial, rats first produced a duration (TP) and subsequently reported its accuracy (i.e., how far the TP is from T). Rats were trained with the reward assignment rule (Fig. 1C) which associated TP error (small error: TP_{SE} trials *vs.* large error: TP_{LE} trials) with reward delivery of different sizes ($TP_{SE\text{-}2p}$ and $TP_{LE\text{-}1p}$ trials), or reward absence ($TP_{SE\text{-}0p}$ and $TP_{LE\text{-}0p}$ trials) in one of two ports. A given port delivered 1p after a large error ($TP_{LE\text{-}1p}$) but 0p after a small error ($TP_{SE\text{-}0p}$, Fig. 1C, '1p port'), whereas the other port delivered 2p after a small error ($TP_{SE\text{-}2p}$) but 0p after a large error ($TP_{LE\text{-}0p}$, Fig. 1C, '2p port'). Before introducing test-trials, the pattern of reaction times (RT) in the control training trials in which only one port was lit showed that rats reached the port faster when expecting a reward delivery based on their TP error (Supp. Mat. S4, Supp. Fig. 4A). Furthermore, when probed on test trials with both reward ports lit up, the rats were rewarded if they chose according to the rule (Fig. 1C), thus maximizing reward rate. Rats increased their choice reaction times on test-trials compared to training-trials, indicating typical choice behavior (Supp Fig. 4B). Test-trial accuracy (proportion of correct choices) is a direct measure of temporal error monitoring, as the reward rule was conditional on the magnitude of the TP error. Note that the reward assignment rule and threshold adjustment were constructed such that overall reward was similar in both ports. Furthermore, if rats exhibited a bias to choose 2p port on all test trials, the accuracy would result in the chance level accuracy scores (0.5). Indeed, we did not observe strong association between accuracy and

pellet ratio on per rat, per session basis ($ß$ = -1.32, SE = 2.06, t(1781) = -0.64, p = 0.53). From Fig. 2A it is evident that accuracy on test trials (collapsed over 1p and 2p ports) was above chance level for all rats (each animal above 50% chance level, $p < 0.05$, two-sided Wilcox test, see Supp. Table 1). Additionally, the PRESS group exhibited higher test-trial accuracy (Fig. 2B; $ß$ = 0.04, SE = 0.007, z(5967) = 5.46, $p < 10^{-7}$). As the PRESS group was also the one in which animals showed poorer TP performance, it suggests better choice performance in animals with larger variability. To test this hypothesis we assessed the relationship between choice performance and standard deviation across sessions by fitting a model where choice accuracy was predicted by standard deviation with inclusion of group factor. The effect of standard deviation (ß = 0.018, SE = 0.022, t(567) = 0.81, p = 0.417) and interaction between std and group (ß = -0.003, SE = 0.029, t(600) = -0.122, p = 0.902) were not significant. That suggested that the difference in choice accuracy between groups was not driven by differences in variability but was related to the different motor sequences required to produce duration. As rodents may use motor routines to support temporal judgments (Gouvêa et al., 2014; Killeen & Fetterman, 1988; Machado, 1997; Safaie et al., 2020), it raises the question to what extent motor routines could contribute to interval timing and error monitoring.

As the test-trial accuracy neglects a continuous link between choice and TP error, we verified that the rats showed an increased probability of 1p choice as a function of TP error (Fig. 2CD), with GLMM predicting the probability of choosing 1p port *vs.* 2p port given the TP error (Fig. 2E; $ß$ = 3.29, SE = 0.30, z(19512) = 11.01, $p < 10^{-15}$). Rats exhibited choice uncertainty for TPs close to their internal threshold (Fig. 2CDE). Of course, the better the internal threshold aligns with Θ threshold, the higher the chances for higher choice accuracy.

Overall, both analyses showed that rats chose between 1p and 2p options with relatively high accuracy, suggesting that subjects have a representation of the relation between the duration of the just-produced interval and the interval, longer than required interval, chosen based on the level of temporal uncertainty. However, the representation of the required interval, the margin of error, and the representation of the time elapsed during an attempt to produce the required interval, all have limited precision. The uncertainty in the port choice must reflect the uncertainties in the first order variables, as listed, on which it depends. Therefore, near perfect accuracy is not to be expected.

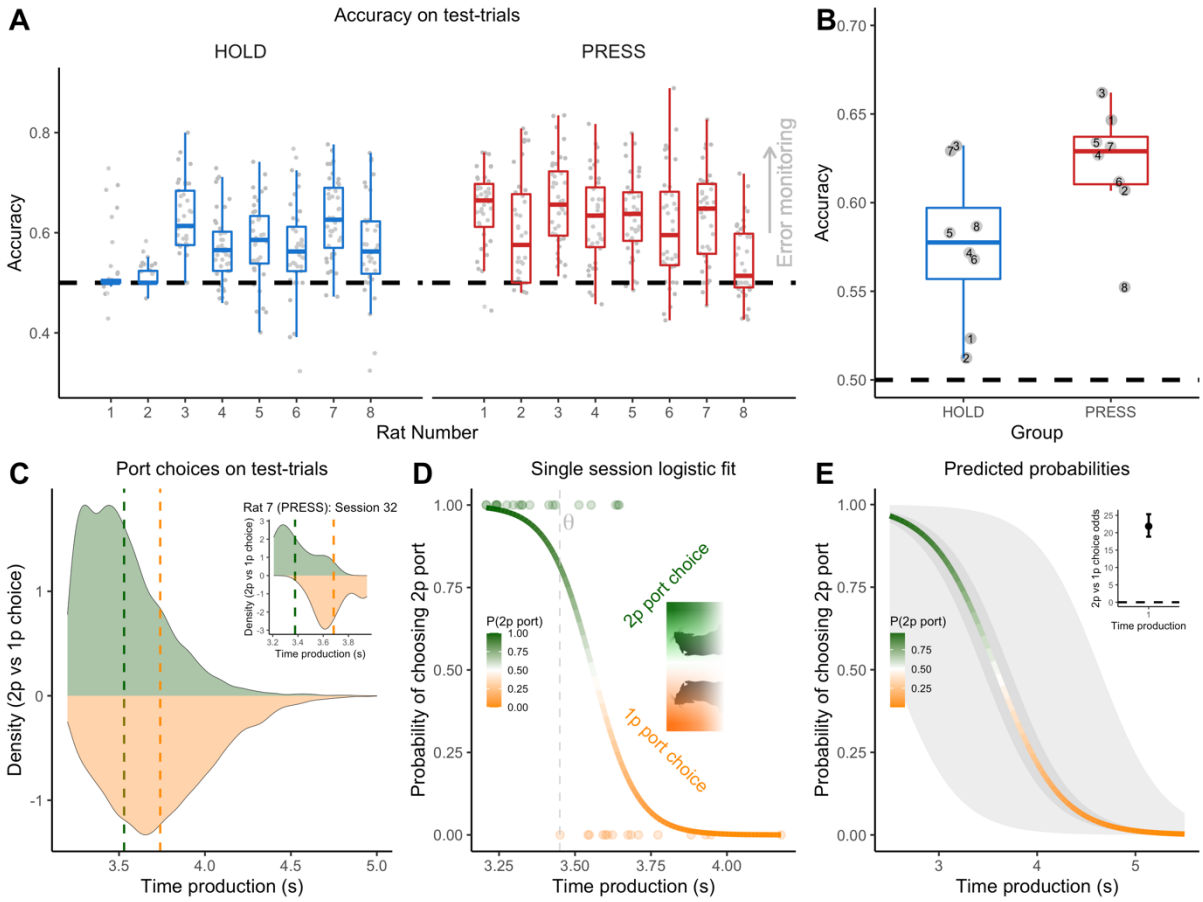


Figure 2. Rats report temporal errors by choosing accurately on test-trials.

(A) Grey data points show session-by-session accuracy on test trials for individual rats, that is, the proportion of test trials in which a rat chose a port at which a reward will effectively be delivered. The 1 pellet and 2 pellets reward trials were averaged together. A strategy in which only one port would be chosen would result in chance level (0.5) accuracy. Each column shows the results for one rat across all experimental sessions for the two groups of rats. All rats chose significantly above the chance level indicating the ability to monitor temporal errors.

(B) Accuracy on test-trials plotted separately for HOLD and PRESS group. Notably PRESS group achieved higher level of accuracy.

(C) Probability density distribution over TPs on test-trials plotted as a function of rats' choices. Time productions reported by rats as 2p trials are plotted in green and trials classified by rats as 1p trials are plotted in orange. The mean of each distribution is indicated by a dashed vertical line of corresponding color. The 1p distribution (orange) is shifted with respect to the 2p distribution (green), clearly indicating that rats' reports are a function of TPs. The main plot shows density across all rats and sessions. The inset plot shows one sample session of one rat.

(D) Logistic function fitted to a sample session of one rat from panel C (inset). As the curve descends, TP duration increases together with the probability of choosing 1p port on test trials and the probability of choosing 2p port decreases. The descent of the logistic function signifies that as TP duration increases the probability of choosing 2p on test trials decreases. The scale ranges from green to orange, corresponding to reward assignment (Fig. 1C). Data points at the top (orange, 1p choice) and at the bottom (green, 2p choice) of the plot show individual TPs, whereas the color code signifies 1p and 2p choices of rat '7' in session 32.

(E) Generalized linear mixed effect model fit including TP as predictor. Data from 14 rats were used to fit the model. Two types of shaded areas display confidence intervals at a population level. The narrow shading displays population level confidence intervals. The wider shading displays population-level predictions with consideration of uncertainty in the variance parameters (Johnson & O'Hara, 2014). Thus, the wider shading can be considered as prediction intervals (Francq et al., 2019). Similarly, as in panel D, color progression from green to orange depicts

increasing probability of 1p port choice as a function of TP. The inset model displays obtained choice odds of TP predictor.

Rats keep track of temporal errors and reward history

We next investigated whether the rats used sources of information other than TP errors to choose accurately on test trials. Recent work showed that local mean or variance of TP is not stable over time in monkeys (Wang & Jazayeri, 2020) or in humans (Kononowicz et al., 2018). This non-stationarity of time production was also observed here, as illustrated in a sample plot (Fig. 3A, Supp. Mat. S7) depicting changes in local mean of TP (we refer to the moving mean of the last n trials as $\mu(TP_n)$ over the course of one sample session). We hypothesized that rats could use fluctuating μ(TP), when faced with test-trials, by keeping track of previous trials. Indeed, rats track reward history of previous trials (Gallistel et al., 2001), also in a situation in which rewards were delivered on the left or on the right side of a main port (Constantinople et al., 2019; Hermoso-Mendizabal et al., 2020). In the current paradigm, the location of the reward delivery port depended on the TP error. Thus, as TPs were non-stationary, rats could have collected more rewards in just one port, in turn contributing to their choices. Indeed, reward on previous trials contributed to test-trial choices (Supp. Mat. S5, Supp. Fig. 5A). We tested the hypothesis that rats based their choices on the *reward history* (RH). For example, if most of the recent trials were TP_{LE}, the rat has been rewarded in the 1p port, therefore increasing a chance to choose that port. To quantify the reward history (RH) we calculated relative amount of reward received in 1p and 2p ports. To test whether rats relied on *reward history*, we used a logistic regression extending the basic model by adding the moving average of n previous trials (μRH_n) to the model, as an expression of reward history. Keeping current TP and adding reward history over a different number of preceding trials (μRH_{10}, μRH_{20}, μRH_{30} for window sizes of 10, 20 or 30 trials) in the same model allowed assessing their relative contributions to choice behavior. Independently of the number of trials considered – namely, independently of the history window size - both current TP and μRH_n contributed to choice behavior, as evidenced by large values of estimated choice odds (Fig. 3B; RH_n: β = 0.60, *SE* = 0.094, $z(16423) = 6.37$, $p < 10^{-9}$; μRH_{10}: β = 0.47, *SE* = 0.041, $z(16423) = 11.34$, $p < 10^{-14}$; μRH_{20}: β = 0.35, *SE* = 0.064, $z(16423) = 0.64$, $p = 10^{-7}$; μRH_{30}: ß = 2.35, $z(16962) = 4.54$, $p < 10^{-5}$).

As the previous TPs can influence current TP (Wiener et al., 2015; Jazayeri & Shadlen, 2010; Roach et al., 2017), we also considered the possibility that reward history due to previous time productions (μTP_n) could influence choices on test-trials (Supp. Mat. S5, Supp. Fig. 5B). As the amount of reward was conditional on TP, mean TP in n previous trials (μTP_n) is a proxy of reward history, incorporating TP history. Indeed, both current TP and μTP_n contributed to choice behavior, as evident in the large values of estimated choice odds (Fig. 3C; TP_n: β = 2.66, $z(16962) = 7.45$, $p < 10^{-13}$; μTP_{10}: β = 1.24, $z(16962) = 3.86$, $p < 10^{-3}$; μTP_{20}: β = 0.50, $z(16962) = 0.91$, $p = 0.36$; μTP_{30}: β = 2.35, $z(16962) = 4.54$, $p < 10^{-5}$). Similar results were obtained for rewarded and unrewarded trials (Supp. Fig. 6). Remarkably, both factors, TP and reward history, remained significant, indicating that rats rely on both sources of information while choosing on test-trials.

Together, these results show that during choice behavior, rats take into account both factors: their temporal error on the current trial (TP_n) and the reward history (μRH_n). Contribution of both factors demonstrates the richness of representation of elapsed time, including error monitoring abilities together with tracking previous behavioral outcomes.

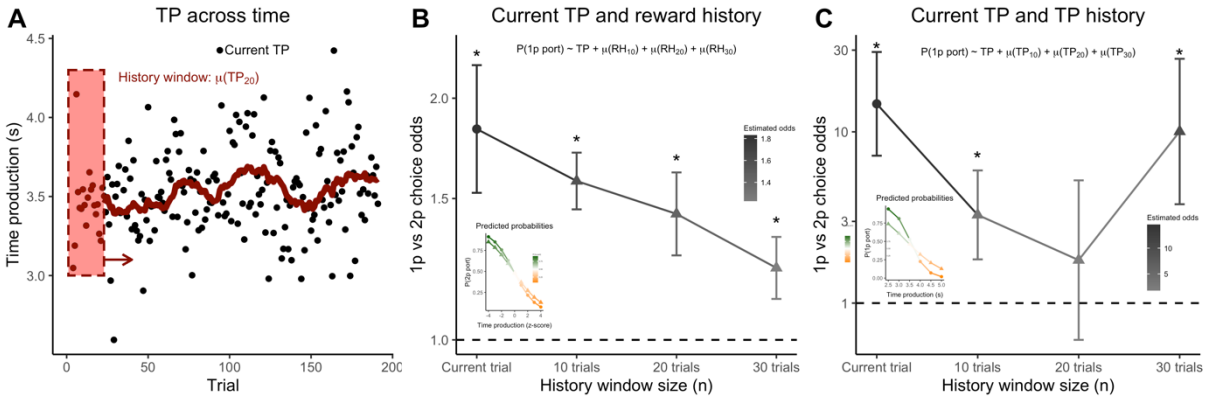


Figure 3. Rats keep track of temporal errors and reward history.

(A) TPs (black) plotted across an experimental session (example). Shaded red square depicts a moving average window. The red trace marks the moving average of 20 TPs. That is, one data point of moving average is a result of averaging the preceding 20 trials. Fluctuation of moving average illustrates non-stationarity of TP.

(B) Results of GLMM fit including current TP and running average of last 10, 20, 30 rewards (Reward History). Data from 14 rats were used to fit the model. The formula in the plot shows specification of fixed terms in the GLMM. Asterisks indicate significant terms. Error bars display confidence intervals of estimated effects. Similar to Fig. 2E, the inset plot depicts logistic fit for Current trial (circle) and reward history ($\mu RW_{10,}$ triangle).

(C) Results of GLMM fit including current TP and running average of last 10, 20, 30 TPs, as a proxy for reward history. Data from 14 rats were used to fit the model. Significant model terms indicate that rats keep track of temporal errors on current TP as well as TP history. Similar to Fig. 2E, the inset plot depicts logistic fit for Current trial (circle) and TP history ($\mu TP_{30,}$ triangle).

Inferred timing errors can guide future behavior.

As rats reported their errors, we asked whether they used this ability to guide future behavior. The first possibility is that temporal errors are only used in the experimental context to obtain rewards. The second possibility is that temporal errors can also be utilized for error adjustment on adjacent trials. Meck (1984) showed that when rats estimated 10 or 20 seconds time intervals, the probability of trial-to-trial alternation between durations shorter, or longer, than the median was above chance level. Alternation of shorter and longer trials in an ordered sequence was a signature of 'threshold hunting', which we investigated in this dataset by analyzing errors on consecutive trials, referred as relative errors (re, Fig. 4A). Focusing on relative errors instead of raw TPs allows highlighting local adjustments of TP, which would otherwise be confounded with other phenomena involving slow TP changes across multiple trials. If there is no representation of relative error, there is no possible adjustment on the next TP. Therefore, no relation between error on *n-th* TP (re_n) and error on *n-th + 1* TP (re_{n+1}) would be expected. In contrast, if temporal errors are utilized, small re_n will be followed by large re_{n+1}, and vice versa, resulting from adjustment of consecutive relative errors. We refer to this behavior as 'threshold hunting' (Meck, 1988; Meck et al., 1984).

We controlled for the effects of reward magnitude on TP (Supp. Mat. 9) by focusing on triplets of trials where the first two trials were accurate trials rewarded with 2 pellets ($TP_{SE\text{-}2p}$). As reward magnitude calibrates performance variability on the next trial (Wang et al., 2019; Dhwale et al., 2019), we verified that the TP variability on the 2 pellets trials was unaffected by reward magnitude (Supp. Mat. 9, Supp. Fig. 9). re_n was calculated as the signed difference of temporal error between the second and the first trial in each triplet (2nd $TP_{SE\text{-}2p}$ - 1st $TP_{SE\text{-}2p}$, Fig. 4A). Likewise, re_{n+1} was calculated as the difference between the third and the second trial in each triplet (3rd TP - 2nd $TP_{SE\text{-}2p}$, Fig. 4A). Each triplet resulted in one pair of re_n and re_{n+1}. We split the re_n trials into short and long bins of the same count. Based on short and long re_n bins, we then evaluated re_{n+1} (Fig. 4A). In line with the error monitoring hypothesis, we found that re_{n+1} in long re_n bin were shorter than re_{n+1} in short re_n bin (Fig. 4B, ß = -0.096, $t(2441) = -7.34$, $p < 10^{-12}$). Thus, after being rewarded on the *n-th* trial with 2 pellets, rats adjusted their relative error on the next trial, suggesting the 'threshold hunting' behavior.

If the 'threshold hunting' behavior is a behavioral proxy for error monitoring, the degree to which individual rats exhibited hunting should be associated with error monitoring metrics: test-trial accuracy and estimated choice odds given current TP and TP history. A hunting index was calculated by subtraction of long bin re_{n+1} from short bin re_{n+1}. It is evident from the left panel of Fig 4C that, the larger the hunting index, the larger the average accuracy (Fig. 4C*; *$t(12) = 3.25$, $p = 0.007$; robust regression). As revealed by previous analyses, the overall test-trial accuracy was influenced by two factors: current TP and TP history. To assess whether they relate differently to the hunting index, we retrieved the estimated parameters of the GLMM for TP and μTP_{30} for each rat. We chose μTP_{30} because this factor showed the highest estimates among TP history factors in the previous analyses. If the degree of hunting is a proxy for error monitoring, the larger the hunting index, the larger the current TP estimates should be; and the opposite prediction holds for TP history (TP_{30}). As suggested by Fig. 4C (right panel), the rats that relied more on the current TP also exhibited the largest hunting index ($t(11) = 7.44$, $p < 10^{-4}$), whereas those who relied more on the TP history exhibited the smallest hunting index ($t(11) = -4.19$, $p = 0.002$; Fig. 4C, right panel inset). The addition of TP_{30} to the model including

TP was statistically justified (Wald test: $W(1) = 17.58$, $p < 10^{-4}$). Together, while controlling for the plausible effects of variance (Supp. Mat. 9), we demonstrated hunting behavior in rats. The association of hunting behavior with indices of choice accuracy suggests that rats can use estimated errors to guide future behavior, supporting a functional role of error monitoring in guiding adaptive behavior.

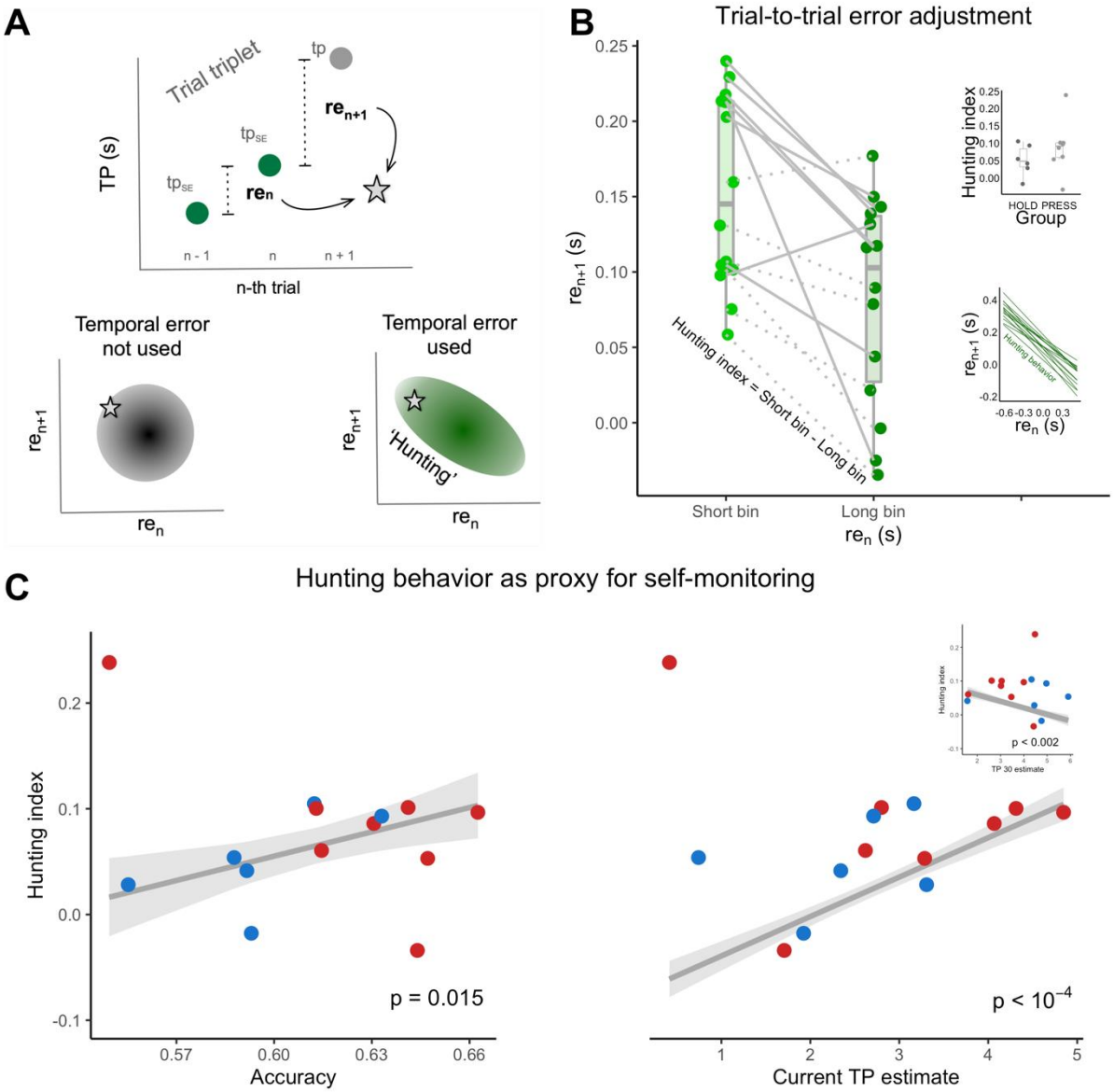


Figure 4. Inferred timing errors guide future behavior.

(A) The error estimation hypothesis predicts adaptation on the n+1 trial and so-called 'hunting' behavior. In top panel, we considered triplets of trials where the two first trials were 2p trials. Relative error in TP on *n* trial is calculated as the subtraction of TP during the second trial from TP during the first in a given triplet (re_n = '*n*' - '*n-1*'). Relative error in TP on *n* + 1 trial is result of the subtraction of TP in the third trial from TP in the second trial in a given triplet (re_{n+1} = '*n+1*' - '*n*'). The bottom panels depict predictions in case when temporal errors are used or not used to guide behavior. If temporal error monitoring does not guide behavior no correlation between re_n and re_{n+1} is predicted. However, the hypothesis of an error monitoring predicts a negative association between re_n and re_{n+1}, which is depicted in the bottom right panel.

(B) re_n were categorized as short and long bins, based on re_n magnitude. Values of re_{n+1} are plotted against short and long re_n bins. Grey lines connect data points for individual rats. Dotted lines represent HOLD group and solid lines represent PRESS group. To compute an index of error adjustment (aka 'hunting index'), we subtracted re_{n+1} of the long bin from that of the short bin. The values of hunting index for each rat are plotted in panel C along y axis.

Top inset plot: The values of hunting index plotted for both groups. No between-group difference was found ($t(12) = 0.40$, $p = 0.70$, t-test).

Bottom inset plot: Results of GLMM fit including current re_n as a predictor of re_{n+1} for each individual rat. Each line depicts a correlation slope for each individual rat. This negative relationship between re_n and re_{n+1} was

confirmed with linear mixed regression (β = -0.36, $t(12)$ = -5.70, $p < 10^{-3}$, GLMM), in line with analysis in the main plot in figure 4B.

(C) The blue (HOLD group) and red data points (PRESS group) display the hunting index for each rat plotted against overall choice accuracy on test-trials (left panel) and values of current TP (right panel) and TP history (inset plot in the right panel) estimated from the same model fit. The green lines depict resulting fits of robust regression model. The shaded grey area represents confidence intervals.

Discussion

Although primates and rodents can report their confidence in perceptual tasks (e.g., Dehaene et al., 2017; Kepecs et al., 2008; Fleming et al., 2010; Miyamoto et al., 2018), whether rodents can monitor the errors of self-generated behavior, such as the production of a time interval, was not known. Our results provide evidence for self-monitoring of temporal errors in rats. During timing behavior and its evaluation, rats displayed behavioral characteristics suggesting an online assessment of their temporal errors in line with the idea that rats represent rich experiential variables. Previous studies suggested that rats may take into account their internal temporal uncertainty during timing behavior, but an alternative explanation for a behavior resembling error monitoring was offered for all the previous studies (for review Balci et al., 2011). For example, Foote and Crystal (2007) trained rats to discriminate short and long stimulus durations; when rats were faced with trials close to discrimination boundary, they chose smaller but certain reward, without taking the risk of choosing short, or long, and receiving no reward. The alternative explanation is that rats learned differential reinforcement for different durations (Jozefowiez et al., 2009), putting in question monitoring of temporal errors. However, if animals were responding according to two different memory representations in our time production task, a bimodal TP distribution would have been expected, which we did not observe in the current study.

The paradigm overcome the previous difficulties by employing time production instead of stimulus duration discrimination. By adapting timing thresholds individually for each rat, we forced the animals to remain close to their individual limits of time discrimination. Additionally, by designing a novel way to report temporal errors, we can avoid relying on the whole distribution of responses, which is an indirect means to infer the plausibility of monitoring of internal uncertainty (Balci et al. 2009; Yi & Dudman, 2013). Instead, rats reported estimated errors in their timing behavior on a trial-by-trial basis to maximize reward. Their behavior was based on both the mean and precision of their representation of the elapsed interval. Remarkably, the rats showed the ability to track their error on the current trial, showing that rats keep track of the internal variables on which decisions were based, and of the relation of those variables to the target value. Along with tracking their error on the current trial, they keep history of recent reinforcements to maximize the amount of reward in a given amount of time. It suggests that, in addition to their capacity to store in memory multiple parameters of their behavior (Kheifets et al., 2017), rats have an exquisite ability to monitor their self-generated behavior. Furthermore, the association of test-trial accuracy and size of the TP correction between consecutive trials suggests that rats can use estimated errors to guide future behavior, supporting a functional role of error monitoring in guiding adaptive behavior.

One limitation of the current study is the usage of just one very specific reward assignment rule. On one hand, it forced the rats to be more precise, as TPs close to the objective threshold were rewarded with two pellets. On the other hand, the distribution had to be titrated in proportion to the reward size. Thus, the question remains of how rats would perform with different reward assignment rules. The question is important as it touches on the relation between levels of temporal uncertainty and ability to report temporal errors. Investigating different levels of precision in the future will allow understanding the properties of representations that rats use to keep track of time. Modification of the reward assignment rule

should also allow to investigate whether rodents possess the ability to monitor the sign of temporal errors as humans (Akdoğan & Balci 2017; Kononowicz et al., 2018).

Our finding of error monitoring of timing behavior in rats is markedly similar to recent findings in humans, where participants were able to report signed magnitude of their temporal errors (Akdoğan & Balci 2017; Kononowicz et al., 2018). Although we do not claim that rats and humans use the same cognitive machinery for monitoring of temporal errors, our results indicate that rats use basic computations allowing error monitoring, whereby the monitoring process actively infers the state of an 'internal clock'. Ultimately, these results deepen our understanding of error monitoring abilities in rodents, and demonstrate a richness of representation of elapsed time that goes beyond trial-and-error strategy in interval timing. Together, thanks to our novel way to investigate animal cognition, the present findings show that monitoring of internal uncertainty is not a privilege of primates. The current study contributes to ongoing discussion on self-monitoring (Dehaene et al., 2017).

References

Akdoğan, B., & Balcı, F. (2017). Are you early or late?: Temporal error monitoring. J Exp Psychol Gen, 146(3), 347-361.

Balci, F., Freestone, D., Simen, P., Desouza, L., Cohen, J. D., & Holmes, P. (2011). Optimal temporal risk assessment. Front Integr Neurosci, 5, 56.

Balci, F., Freestone, D., & Gallistel, C. R. (2009). Risk assessment in man and mouse. Proc Natl Acad Sci U S A, 106(7), 2459-2463.

Barr, D. J. (2013). Random effects structure for testing interactions in linear mixed-effects models. *Frontiers in psychology*, *4*, 328.

Bates D., Maechler M., Bolker B., Walker S. (2015). Fitting Linear Mixed-Effects Models Using lme4. Journal of Statistical Software, 67(1), 1-48. doi:10.18637/jss.v067.i01.

Buhusi, C. V., & Meck, W. H. (2005). What makes us tick? Functional and neural mechanisms of interval timing. *Nature reviews neuroscience*, *6*(10), 755-765.

Çavdaroğlu, B., Zeki, M., & Balcı, F. (2014). Time-based reward maximization. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *369*(1637), 20120461.

Constantinople, C. M., Piet, A. T., & Brody, C. D. (2019). An analysis of decision under risk in rats. *Current Biology*, *29*(12), 2066-2074.

Dhawale, A. K., Miyamoto, Y. R., Smith, M. A., & Ölveczky, B. P. (2019). Adaptive Regulation of Motor Variability. *Current Biology*, *29*(21), 3551-3562.

Dehaene, S., Lau, H., & Kouider, S. (2017). What is consciousness, and could machines have it. Science, 358(6362), 486-492.

Fleming, S. M., Weil, R. S., Nagy, Z., Dolan, R. J., & Rees, G. (2010). Relating introspective accuracy to individual differences in brain structure. *Science*, *329*(5998), 1541-1543.

Foote, A. L., & Crystal, J. D. (2007). Metacognition in the rat. Curr Biol, 17(6), 551-555.

Francq, B. G., Lin, D., & Hoyer, W. (2019). Confidence, prediction, and tolerance in linear mixed models. *Statistics in medicine*, *38*(30), 5603-5622.

Freestone, D. M., Balcı, F., Simen, P., & Church, R. M. (2015). Optimal response rates in humans and rats. J Exp Psychol Anim Learn Cogn, 41(1), 39-51.

Gallistel, C. R., & Gibbon, J. (2000). Time, rate, and conditioning. *Psychological review*, *107*(2), 289.

Gallistel, C. R., Mark, T. A., King, A. P., & Latham, P. E. (2001). The rat approximates an ideal detector of changes in rates of reward: implications for the law of effect. *Journal of experimental psychology: Animal behavior processes*, *27*(4), 354.

Gervini, D., & Yohai, V. J. (2002). A class of robust and fully efficient regression estimators. *The Annals of Statistics*, *30*(2), 583-616.

Gibbon, J. (1992). Ubiquity of scalar timing with a Poisson clock. *Journal of mathematical psychology*, *36*(2), 283-293.

Gouvêa, T. S., Monteiro, T., Soares, S., Atallah, B. V., & Paton, J. J. (2014). Ongoing behavior predicts perceptual report of interval duration. Front Neurorobot, 8, 10.

Hermoso-Mendizabal, A., Hyafil, A., Rueda-Orozco, P. E., Jaramillo, S., Robbe, D., & de la Rocha, J. (2020). Author Correction: Response outcomes gate the impact of expectations on perceptual decisions. Nat Commun, 11(1), 3470.

Jiahui Wang, Ruben Zamar, Alfio Marazzi, Victor Yohai, Matias, Salibian-Barrera, Ricardo Maronna, Eric Zivot, David Rocke, Doug Martin, Martin Maechler and Kjell Konis. (2020). robust: Port of the S+ "Robust Library". R package version 0.5-0.0. https://CRAN.R-project.org/package=robust

Insabato, A., Pannunzi, M., & Deco, G. (2016). Neural correlates of metacognition: A critical perspective on current tasks. *Neuroscience & Biobehavioral Reviews*, *71*, 167-175.
Jazayeri, M., & Shadlen, M. N. (2010). Temporal context calibrates interval timing. *Nature neuroscience*, *13*(8), 1020.
Johnson PC, O'Hara RB. 2014. Extension of Nakagawa & Schielzeth's R2GLMM to random slopes models. Methods Ecol Evol, 5: 944-946.
Jozefowiez, J., Staddon, J. E. R., & Cerutti, D. T. (2009). The behavioral economics of choice and interval timing. *Psychological Review*, *116*(3), 519.
Kepecs, A., Uchida, N., Zariwala, H. A., & Mainen, Z. F. (2008). Neural correlates, computation and behavioural impact of decision confidence. Nature, 455(7210), 227.
Kheifets, A., & Gallistel, C. R. (2012). Mice take calculated risks. *Proceedings of the National Academy of Sciences*, *109*(22), 8776-8779.
Kheifets, A., Freestone, D., & Gallistel, C. R. (2017). Theoretical implications of quantitative properties of interval timing and probability estimation in mouse and rat. *Journal of the Experimental Analysis of Behavior*, *108*(1), 39-72.
Killeen, P. R., & Fetterman, J. G. (1988). A behavioral theory of timing. Psychological review, 95(2), 274.
Kononowicz, T. W., Roger, C., & van Wassenhove, V. (2019). Temporal Metacognition as the Decoding of Self-Generated Brain Dynamics. Cereb Cortex, 29(10), 4366-4380.
Kuznetsova A, Brockhoff PB, Christensen RHB (2017). "lmerTest Package: Tests in Linear Mixed Effects Models." Journal of Statistical Software, *82*(13), 1-26.
Laird, N. M., & Ware, J. H. (1982). Random-effects models for longitudinal data. *Biometrics*, 963-974.
Le Pelley, M. E. (2012). Metacognitive monkeys or associative animals? Simple reinforcement learning explains uncertainty in nonhuman animals. J Exp Psychol Learn Mem Cogn, 38(3), 686-708.
Li, Y., & Dudman, J. T. (2013). Mice infer probabilistic models for timing. *Proceedings of the National Academy of Sciences*, *110*(42), 17154-17159.
Luzardo, A., Ludvig, E. A., & Rivest, F. (2013). An adaptive drift-diffusion model of interval timing dynamics. Behavioural processes, 95, 90–99.
Lüdecke D (2018). *ggeffects: Tidy Data Frames of Marginal Effects from Regression Models.* Journal of Open Source Software, 3(26), 772.
Machado, A. (1997). Learning the temporal dynamics of behavior. Psychological review, 104(2), 241.
Meck, W. H. (1988). Hippocampal function is required for feedback control of an internal clock's criterion. *Behavioral neuroscience*, *102*(1), 54.
Meck, W. H., Komeily-Zadeh, F. N., & Church, R. M. (1984). Two-step acquisition: modification of an internal clock's criterion. Journal of Experimental Psychology: Animal Behavior Processes, 10(3), 297.
Miyamoto, K., Osada, T., Setsuie, R., Takeda, M., Tamura, K., Adachi, Y., & Miyashita, Y. (2017). Causal neural network of metamemory for retrospection in primates. *Science*, *355*(6321), 188-193.
Paton, J. J., & Buonomano, D. V. (2018). The Neural Basis of Timing: Distributed Mechanisms for Diverse Functions. Neuron, 98(4), 687-705.
Platt, J. R., Kuch, D. O., & Bitgood, S. C. (1973). RATS'LEVER-PRESS DURATIONS AS PSYCHOPHYSICAL JUDGMENTS OF TIME 1. *Journal of the experimental analysis of behavior*, *19*(2), 239-250.

Pleskac, T. J., & Busemeyer, J. R. (2010). Two-stage dynamic signal detection: a theory of choice, decision time, and confidence. *Psychological review*, *117*(3), 864.

Roach, N. W., McGraw, P. V., Whitaker, D. J., & Heron, J. (2017). Generalization of prior information for rapid Bayesian time estimation. Proc Natl Acad Sci U S A, 114(2), 412-417.

Safaie, M., Jurado-Parras, M. T., Sarno, S., Louis, J., Karoutchi, C., Petit, L. F., ... & Robbe, D. (2020). Turning the body into a clock: Accurate timing is facilitated by simple stereotyped interactions with the environment. *Proceedings of the National Academy of Sciences*, *117*(23), 13084-13093.

Simen, P., Balci, F., deSouza, L., Cohen, J. D., & Holmes, P. (2011). A model of interval timing by neural integration. Journal of Neuroscience, 31(25), 9238-9253.

Wang, J., Hosseini, E., Meirhaeghe, N., Akkad, A., & Jazayeri, M. (2020). Reinforcement regulates timing variability in thalamus. *eLife*, 9:e55872.

Wiener, M., & Thompson, J. C. (2015). Repetition enhancement and memory effects for duration. *Neuroimage*, *113*, 268-278.

Methods

Subjects

The subjects were male Sprague-Dawley rats (N=16; Envigo, France). They were 3 months old at the beginning of the experiments. The rats were housed in groups of 4, and kept under a 12 h–12 h light/dark cycle. The experiments were performed during the light cycle. After one week of acclimation to the vivarium with ad libitum food and water access, rats were food restricted such that their weight reached and were maintained at 85% of their free-fed weight. All rats were handled, weighed and fed daily until the end of the experiment. For all rats, two experimental sessions were scheduled per day. The first and the second session of the day was separated by approximately 3 hours during which rats returned to home cages with water access. The experiment was carried out in accordance with the guidelines established by the European Communities Council Directive (2010/63/EU) for compliance and use of laboratory animals, and approved by the French Ministry of Research and the French National Ethical Committee (2013/6).

Apparatus

The experimental setup was built using the Habitest Modular System (Coulbourn Instruments, Allentown, PA, USA) with an operant box (Rat Test Cage, H10-11R-TC) placed in an isolation cubicle (H10-24A). Figure 1 depicts the spatial arrangement of the behavioral box. On the left wall, a retractable lever (H23-17RC) was placed between two reward ports (H14-01R), each connected to a feeder (H14-23R) delivering grain-based pellets (45 mg, Phymep, Dustless Precision Pellets, Paris, FR) and with photocell sensors (H20-94) to measure port entries. Lever was separated from feeders by 7 cm and was placed 6.5 cm above the floor grid. Reward ports contained an LED to serve as a reward cue after each successful time production. A red house-light (H11-01R-LED) and a speaker (H12-01R) were placed on the right wall. The system was controlled and data collected using Graphic State 4 software. The data were converted to text files and analyzed using custom written R code.

Procedure: pretraining

The pretraining followed several steps. The first step was the feeder training (2 sessions) which consisted in familiarizing rats with the reward ports, and the pairing of the pellet drop with the onset of the reward-port light. A trial started after an inter-trial interval (ITI), randomly chosen from 10 to 15 s (in steps of 1s), with a pellet drop that was simultaneous with the onset of the light in the reward port. The light stayed on for up to 2 s unless the pellet had been picked-up. Rats had 7 s to pick up the pellet. The next trial started when a pellet was collected or when 7 shad elapsed. The session ended when rats collected 40 pellets. The side of the reward delivery was pseudo-randomized.

In the second pretraining step (port-light training, 10 sessions), the rats were required to nose-poke in the reward port in order to receive the reward. The maximal time for which the port-light cue would stay on was set to 5 s. All other aspects of the task remained unchanged. When all rats successfully collected rewards on more than 80 % of trials in two consecutive sessions, we moved to the third step.

In the third step (lever-press training, 3 sessions), the rats were required to press a lever in order to activate the light cue. The trial started with the lever presentation, that the rats had to press for turning on a port-light cue (randomly chosen to be the left or the right side) waiting for the rat to poke in for reward delivery. The lever was retracted after a lever press was made or 5 s had elapsed, and a new ITI started. Rats could earn 50 pellets in one session, otherwise the session ended after 30 min. All rats reached a criterion of 80 % rewarded trials in two consecutive sessions.

Duration and precision training

Two groups of rats were formed, corresponding to two different ways of producing durations. One group of rats (N=8; HOLD group) was trained to hold the lever down for a required time of target duration. The lever press and the lever release served as onset and offset times for the produced duration, respectively. The second group of rats (N=8; PRESS group) was trained to produce a duration spaced between two lever presses. The first lever press and the second lever press served as the onset and the offset times, respectively. Generation of another press by PRESS group likely involves more degrees of freedom in the rat's motor routine. Slightly more complex motor pattern likely translated to difference in precision of TP (Fig. 1). Four operant boxes were used to train the rats. Two rats of the PRESS and of the HOLD groups were assigned to each operant box. Two rats of each home cage were assigned to PRESS and two were assigned to HOLD to counterbalance for any 'home cage' effects.

We refer to the time interval that the rat produced, on any given trial, as Time Production (TP). To train the rats to produce a duration of 3.2 s, which we refer to as the target duration (T), we followed the procedure used by Platt et al. (1973). We started with T of 0.4 s, for which rats obtained access to the reward only if they produced a duration of at least 0.4 s. To train the rats on longer durations, we increased progressively T following logarithmically spaced targets between 0.4 s and 3.2 s (T: 0.4, 0.61, 0.92, 1.39, 2.11, and 3.2 s). For each rat, and in each session, a density function was calculated for the TP distribution, as well as the associated proportion of the area under the curve that exceeded the criterion T. A longer T was introduced to all rats when 80 % of the area under the curve exceeded the T criterion on two consecutive sessions. However, computation of thresholds in the remaining parts of training was performed individually.

On any given trial of duration and precision training trial (Supp. Fig. S1) the lever was inserted after an ITI of 4 to 8 s (in temporal increments of 1 s). The rats had unlimited time to initiate their interval production by pressing the lever. When the TP (holding or pressing twice, depending on the group assignment) was too short (TP < 3.2 s), the lever was retracted and the trial restarted with the insertion of the lever after a random ITI. After the rewarded trials (3.2 s < TP < $2*\sigma(TP)$), rats had between 4 s to 8 s to collect their reward, after which, the lever was presented again. The ITI was adjusted by the sound duration, resulting in ITI values ranging from 3 s to 7 s. This trial structure remained unchanged until the end of the experiment.

Once all subjects achieved the T criterion for correct performance on the final T (3.2 s), a bound on longer durations was introduced that equated $2*\sigma_{TP}$, meaning that rats were not reinforced in trials in which their TP exceeded two standard deviations away from the mean TP calculated

over the two previous sessions (σ_{TP}). This procedure intended to force the animal to be precise in their TP. To signal to the rats that a TP was too long, a tone (4 kHz, 1s, 75 dB) was played after the TP was terminated. The sound was incorporated to distinguish the too-long no-reward trials from the too-short no-reward trials. During the precision training, every 2 sessions the threshold for too-long TP was recomputed on a per rat basis. We observed that TP precision increased in the course of precision training, however we did not define a strict stop-rule criterion as a changed payoff in the next training step could have caused further increase in TP precision. Within one session, the rats could obtain up to 100 pellets. The session was limited to 45 minutes. Rats adjusted their behavior in response to the sound signal by producing a narrower TP distribution, rendering them more precise in their TP.

During the duration training, the port-light cue always correctly indicated the reward-port in which a reward was available. The reward port was randomly chosen with 50 % chance. All rats chose the cued port first with accuracy above 95 %. The duration training took 34 sessions; the precision training took 12 sessions.

Error monitoring training

We refer to this part of training as error monitoring training, which took 36 sessions. In this phase, the rats learned to assign a reward size as a function of their error size in TP. To assign reward size to TP errors, we estimated a threshold (Θ) based on the individual TP distribution. Figure 1C depicts a hypothetical distribution of TP for a given animal. More accurate TP (which we name TP_{SE} for TP Small Error), between T and Θ, were rewarded with 2 pellets in one port (Fig. 1C, green area). The less accurate TP (which we refer to as TP_{LE} for TP Large Error), between Θ and $2*\sigma$(TP), were rewarded with 1 pellet in the other port (Fig. 1C orange area). Θ was estimated as a point dividing the green and orange areas with 1 to 2 ratio (Fig. 1C). The first threshold (Θ) to separate green and orange areas was computed based on the performance of the animals in the two last sessions of precision training. The optimal ratio of 2p to 1p trials was 0.5, calculated as the proportion of TP_{SE} (2p) and TP_{LE} (1p) trials. That ratio guaranteed equal number of pellets delivered in both ports. Thresholds Θ were computed for each animal individually. Θ was updated when rats deviated from optimal ratio by factor of 3, meaning that when the lower (0.16) or upper (1.5) bound were exceeded. The individual thresholds were recalculated when the update bounds were exceeded on two consecutive sessions.

The contingencies assigning TP length to a reward port are depicted in Figure 1CD. Specifically, on a given trial, after TP generation, if TP fell between T and $2*\sigma$(TP), either the left port or the right port was lit, chosen randomly. For TP_{SE} trials (green area) animals received 2 pellets in the one port ('2p port') or 0 pellet in the other port ('1p port'). We refer to trials rewarded with 2 pellets and 0 pellet as $TP_{SE\text{-}2P}$ and $TP_{SE\text{-}0P}$, respectively. Generation of TP_{SE} on a given trial resulted in a 50% chance of receiving a reward as $TP_{SE\text{-}2P}$ or $TP_{SE\text{-}0P}$, as indicated by the arrows above pointing to the 2p and 1p port. For less accurate TPs that landed in the orange area animals received 1 pellet in the 1p port or 0 pellet in the 2p port. We refer to TP_{LE} trials that were reinforced with 1 pellet and 0 pellet as $TP_{LE\text{-}1P}$ and $TP_{LE\text{-}0p}$, respectively. The assignment of $TP_{LE\text{-}1p}$ and $TP_{SE\text{-}2p}$ to the left and right port was counterbalanced across animals. Subject entered the lit port on 95% of test-trials on average. Entering the unlit port caused trials termination.

Error monitoring test sessions

To test whether animals were able to report temporal errors, we used the same trial structure as in error monitoring training, but added choice trials (Fig. 1C). We used the same paradigm as in the error monitoring training sessions on 80 % of trials. On the remaining 20 % of trials, we introduced test-trials in which both ports were lit up, giving the animal an opportunity to choose according to the rule described above. Thus, in those test-trials, the rats were offered the possibility to get a reward in either port, with a reward size (0 vs 1 vs 2 pellets) depending on their TP. To maximize the reward rate, the animals should choose in line with the accuracy of time production, categorized as TP_{SE} or TP_{LE}. Choosing TP_{SE-2p} on TP_{SE} and TP_{LE-1p} on TP_{LE} trials would be suggestive of error monitoring abilities. Within one session rats could obtain up to 130 pellets. Session time was limited to 65 minutes. In total, the rats performed 40 error monitoring test-sessions.

Data processing

All analyses used custom-written scripts in R (3.6.1; "Action of the Toes"; R Core Team; 2008). Data sets of individual sessions were removed only if equipment malfunctioned, or when a small number of test-trials was obtained in one of the ports (< 5). In total, we removed 50 individual data sets (per rat and per session) out of 640 data sets of all test sessions (equivalent to 8%).

Statistical analysis

To account for sample dependency, we used linear mixed-effects models (e.g., Gelman & Hill, 2007; Pinheiro & Bates, 2000) as implemented in R package *lme4* (v.1.1-21), that is, for multiple per-animal observations across sessions. Linear mixed-effects models (LMM) are regression models that model the data by taking into consideration multiple levels. Overall, linear mixed-effect model can be expressed as (Laird & Ware, 1982):

(1) $y_i = X_i\beta + Z_i b_i + \varepsilon_i,\ i=1,\ldots,M,\ b_i \sim N(0, \Sigma),\ \varepsilon_i \sim N(0, \sigma^2 I)$,

where y_i is a response vector for the i-th grouping, β is a vector of fixed effects, ε_i reflects error term, b_i is a vector of random effects, X_i and Z_i are matrices reflecting fixed and random effect design. Perhaps the most straightforward and intuitive distinction between fixed and random effects posits that fixed effects are assumed to be constant over experiments, while random effects are a random subset of the levels in the underlying population (Barr et al., 2013).

Animals and sessions were random effects in the model and were allowed to vary in their intercept. *P* values were calculated based on a Type 3 ANOVA with Satterthwaite approximation of degrees of freedom, using the lmerTest package in R (Kuznetsova, Brockhoff, & Christensen, 2017). The mixed-effects model approach was combined with model comparisons that allow for selecting the best fitting model in a systematic manner. Model comparisons were performed based on Akaike Information Criterion (AIC).

As we collected 40 test sessions for each rat, we considered the option to add a nested random effect structure, reflecting that each rat could perform slightly differently on each session. To select the random effect structure, we fitted a model including only the intercept term,

comparing different random effect structures using Chi-square test and AIC. We started with a model containing only the intercept term for different rats. Model comparison showed that the addition of an intercept term for each session, nested within a rat, was justified (ΔAIC = 3164, $\chi^2 = 3262$, $P = 10^{-15}$). The addition of a random slope effect for each predictor was also justified when compared against the model with random intercept per rat and per session (ΔAIC = 2290, $\chi^2 = 2300$, $P = 10^{-15}$). The random slope effects reflected the fact that each rat could perform differently. The random effects definition remained unchanged for all assessed models.

Time production analysis (related to Fig. 1B, Supp. Fig. 3)

The mean (μ) and standard deviation (σ) of TP were calculated across sessions and across rats. To remove the left-tail reflecting impulsive responses and motor errors, we analyzed TP longer than 2 s. On average, each rat produced 218 trials per session excluding trials below 2 s, and 223 including trials of all durations. To assess whether the standard deviation of TP scaled with the mean of TP, we performed LMM, which included rats as random effect, of the following form:

(2) $y_{sr} = \beta_0 + \beta_1 \mu TP_{sr} + b_{0r} + \varepsilon_{tsr}$, $b_{0r} \sim N(0, \sigma_0^2)$, $\varepsilon_{sr} \sim N(0, \sigma^2)$.

Indices t, s, and r stand for trial, session, and rat, respectively. μTP_{sr} indicates the mean time production calculated per session and per trial, which was included as a fixed effect as indicated by the β_1 coefficient. b_{0r} stands for the random effect of rat. To compare differences between the two groups of rats, we included session as a random effect (Supp Mat. S3).

Reaction time analyses (related to Fig. S4)

To assess whether rats expressed confidence in their choices, differences in RT for rewarded and unrewarded trials, for the 1 pellet and 2 pellet ports (Supp. Fig. 4A; RT $TP_{SE\text{-}1p}$ - RT $TP_{LE\text{-}0p}$ (dark red)); RT $TP_{LE\text{-}2p}$ - RT $TP_{SE\text{-}0p}$ (green)) were calculated. Differences in mean RT were calculated for each session and each rat. Therefore, the linear mixed model included a nested intercept term of session, which was allowed to vary within each rat.

To test whether rats expressed choice behavior on test-trials in which both ports were lit, RT were also averaged on a per session and per rat basis (Supp. Fig. 4B). Test-trials were contrasted against training-trials. LMM was formulated similarly to formula 2, including a random effect of rat.

Choice accuracy on test-trials (related to Fig. 2)

Overall we collected 21129 test trials. Average number of test-trial per rat, per session was 37.9 (Supp. Fig. 10). Individual accuracies were tested against chance level using Wilcoxon test, separately for each rat. Samples were mean accuracy in a session. The exact results are presented in the Supplementary Table 1. The last column of Supplementary Table 1 provides information in how many session accuracy of an individual rat was higher that 55%. Additionally, Supp. Fig. 11 displays 2p/1p ratio obtained per session, per rat. The data displayed in Supp. Fig. 10 and Supp. Fig. 11 were used to analyze the association between accuracy and pellet ratio in paragraph: “Rats choose accurately on test-trials, expressing error monitoring”.

To assess the difference in accuracy between PRESS and HOLD groups, the average accuracy was computed on a per session and per rat basis. The data were tested using LMM with a random intercept of session.

Predicting choices based on TP (related to Fig. 2CDE)

Accuracies for rat 1 and 2 in HOLD group were marginally significant, and they did not exhibit choice behavior on at least half of the sessions (Fig. 2A). Therefore, they were excluded from the single-trial analyses.

A simplified generalized linear model (GLM) was fit to illustrate the principle of performed analyses. Rat (nr 7, PRESS group, session 32) choices from a single session (2p port vs 1p port choices) were predicted using TP from the same trial. That GLM had the following form:

(3) $\log\left(\frac{P(1p)}{P(2p)}\right)_t = \beta_0 + \beta_1 TP_t + \varepsilon_t,\ \varepsilon_t \sim N(0, \sigma^2)$,

where TP_{nt} stands for TP on single trial.

Modified version LMM, that is generalized linear mixed model (GLMM), was used to model binary outcome variable (2p port vs 1p port choices) to determine whether TP predicts rats' choices. The most basic GLMM had the following form:

(4) $\log\left(\frac{P(1p)}{P(2p)}\right)_{tsr} = \beta_0 + \beta_1 TP_{tsr} + b_{0r} + b_{1sr} + \varepsilon_{tsr}$,

$b_{0r} \sim N(0, \sigma_0^2)$, $b_{1sr} \sim N(0, \sigma_1^2)$, $\varepsilon_{tsr} \sim N(0, \sigma^2)$

The log-odds of rats' choices were modeled as a function of time production (TP), expressed by TP rat. $b_{0r} + b_{1sr}$ express random intercept of rat and session, with session nested within a rat. Nested random intercepts of session within rat was kept for remaining model fits. The model was fit using 19,512 samples of test-trials.

Predicting choices based on TP and reward history (related to Fig. 3, Supp Fig. 5, 6, 7)

To investigate whether rats used sources of information other than current trial TP error to choose accurately on test-trials, we first fitted a model incorporating lagged rewards and TP on previous trials (Supp. Fig. 5A, 5B, respectively). We used current TP and fifteen other predictors (different lags) which were reward size (Reward History, RH) or TP vectors lagged by n trials (n = {1:15}). We used a similar approach, as introduced in the previous paragraph with several extensions. The model of the following form:

(5) $\log\left(\frac{P(1p)}{P(2p)}\right) = \beta_0 + \beta_1 TP_{tsr} + \beta_2 RH_{1tsr} + \beta_3 RH_{2tsr} + \beta_4 RH_{3tsr} + \ldots + \beta_{16} RH_{17tsr} + b_{0r} + b_{1sr} + \varepsilon_{tsr},\ b_{0r} \sim N(0, \sigma_0^2)$, $b_{1sr} \sim N(0, \sigma_1^2)$, $\varepsilon_{tsr} \sim N(0, \sigma^2)$,

was enhanced with a fixed effect of reward history (RH_n), where n stands for the trial lag. The model where TP was included instead of RH (Supp. Fig. 5B) was of the same form. Both models were fit using 19,309 samples of test-trials.

To investigate cumulative effects of reward history, we calculated mean μRH_n across n trials. One individual trace of nonstationary TP was visualized in Fig. 3A. All RH traces that were used in this analysis were visualized in Supp. Fig. 7. Using these predictors, we fitted a model of the following form:

(6) $\log\left(\frac{P(1p)}{P(2p)}\right) = \beta_0 + \beta_1 TP_{tsr} + \beta_2 \mu RW_{10tsr} + \beta_3 \mu RW_{20tsr} + \beta_4 \mu RW_{30tsr} +$

$b_{0r} + b_{1sr} + b_{2sr}\, TP_r + b_{3r}\, \mu RW_{10r} + b_{4r}\, \mu RW_{20r} + b_{5r}\, \mu RW_{30r} + \varepsilon_{tsr}$,

$b_{0r} \sim N(0, \sigma_0^2)$, $b_{1sr} \sim N(0, \sigma_1^2)$, $b_{2sr} \sim N(0, \sigma_2^2)$, $b_{3sr} \sim N(0, \sigma_3^2)$, $b_{4sr} \sim N(0, \sigma_4^2)$,

$b_{5sr} \sim N(0, \sigma_5^2)$, $\varepsilon_{tsr} \sim N(0, \sigma^2)$,

where $\beta_1 TP_{tsr}$ express a fixed effect of TP, and $\beta \mu RW_{ntsr}$ express a fixed effect of cumulative reward in 1p and 2p ports in n last trials. In addition to nested session by rat random intercepts, as in the previous models, we added a random effect of slope for each rat ($b_r \, \mu RW_{nr}$). A model of the same structure was fit for μTP_n cumulative TP history. Both models were fit using 16,973 samples of test-trials. The difference of 2,539 trials, as compared to initial model, was due to averaging of previous trial, causing removal of initial 30 trials of each session.

Predicting relative errors based on error monitoring accuracy (related to Fig. 4)

To investigate whether rats adjust their errors between trials we identified triplets of trials where the two first trials were accurate trials rewarded with 2 pellets. In total, we identified 2750 such triplets from all rats (Fig. 4B).

We observed outliers in Fig. 4C, which violate the normality assumption of a linear regression. Therefore, we used a robust linear model that attempts to minimize the influence of outlier observations without removing them from a dataset. Using Robust and Efficient Weighted Least Squares estimator (Gervini & Yohai, 2002), the least square weights were adaptively estimated using the empirical distribution of residuals. As such, robust methods are less sensitive to a non-normal distribution. All models, in this section, were implemented using 'robust' R package (v0.5, Wang et al., 2020). The fitted model predicted hunting index with test-trial accuracy (Fig. 4B). To disentangle the effect of different contributions of test-trial accuracy to hunting index, we used individual coefficient for each rat, expressing current trial contributions and previous trials contributions. Thus, we used parameters of fitted GLMM as described in the Equation 6, however just with two terms: $\beta_1 TP_{tsr}$ and $\beta_2 \mu TP_{30tsr}$. We retrieved the coefficient for each rat and used them as predictors of hunting index in the second fitted robust regression (Fig. 4C).

Acknowledgments

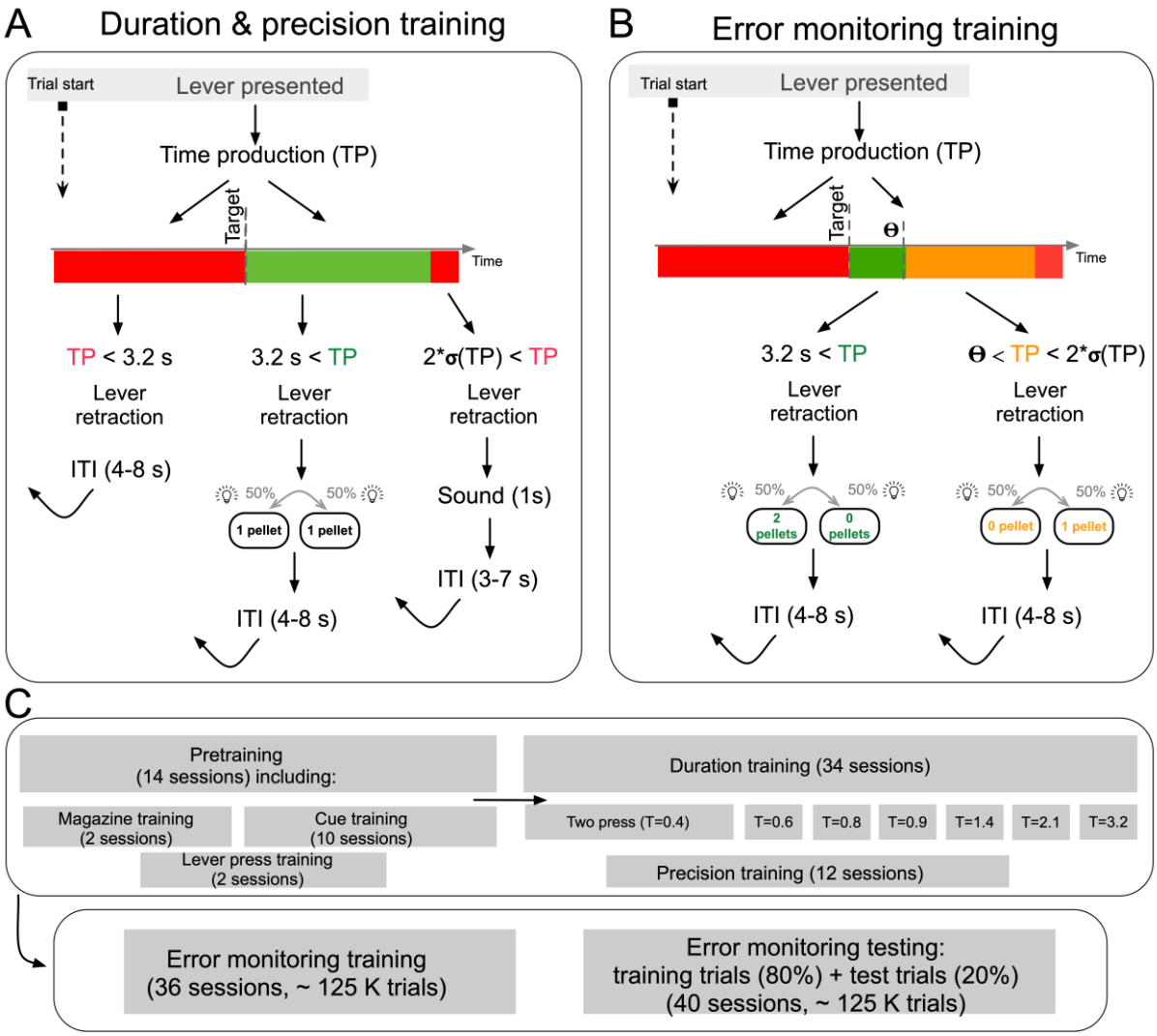
We thank Sloane Paulcan, Lisa Salleille, and Dr. Lucille Tallot, from Memory, Emotion & Time group, for their help with troubleshooting, and the zootechnicians of NeuroPSI for taking care of the animals. We thank Jasmina Saadi and Jonathan Schmidt for their help with data collection. We thank Dr. Bruce L. Brown for his helpful comments on the initial draft the manuscript, and , Dr. Sylvia Wirth, and Dr. Rene Duhamel for their helpful feedback on the work. We thank the editor and two reviewers for their extremely helpful comments. Funding: ANR-AutoTime & ANR-TimeMemory.

Funding

This work was supported by an OPUS grant 2019/35/B/HS6/04389 to TWK, ANR-AutoTime & ANR-TimeMemory to VD and VvW.

Supplementary materials and figures

S1. Schematic of a trial structure during training

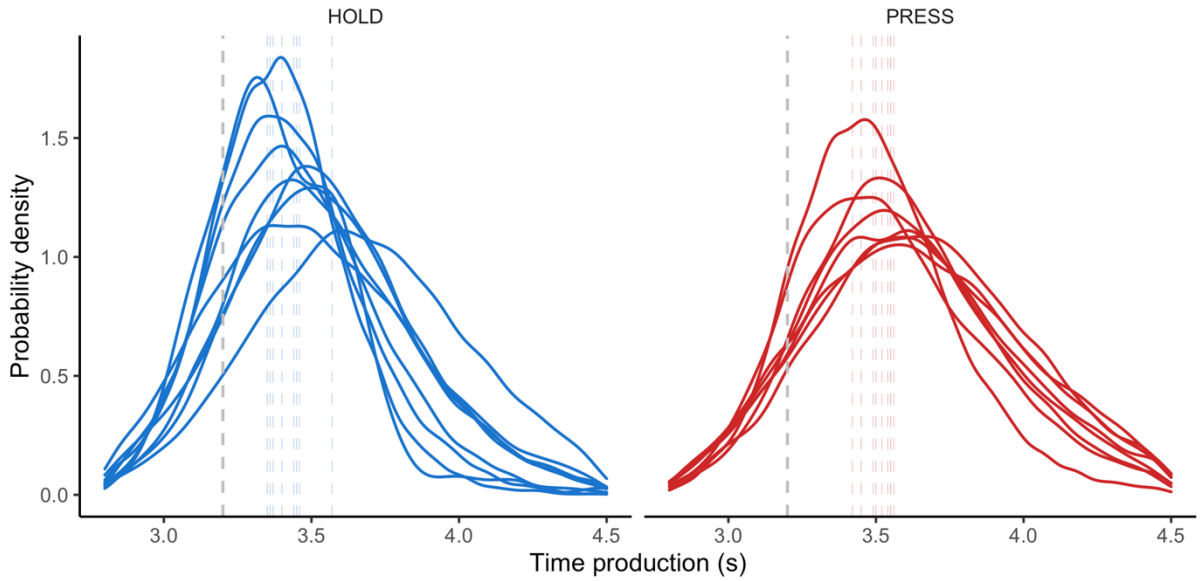


Supplementary Figure 1. Trial structure during training.
(A) Schematic of a trial structure on duration and precision training trials. Task events are depicted in sequence from the top to bottom. They differ along the TP-axis (color bar with red, green, and orange colors) and show different scenarios, in single trials, that were determined by rats' performance on TP. An inter-trial-interval (ITI) was the last event in a single trial sequence. See Methods: *Duration and precision* for further explanation.

(B) Schematic of a trial structure on error monitoring training trials. See Methods*: Error monitoring training* for further explanation.

(C) Schematic of pretraining and training procedure.

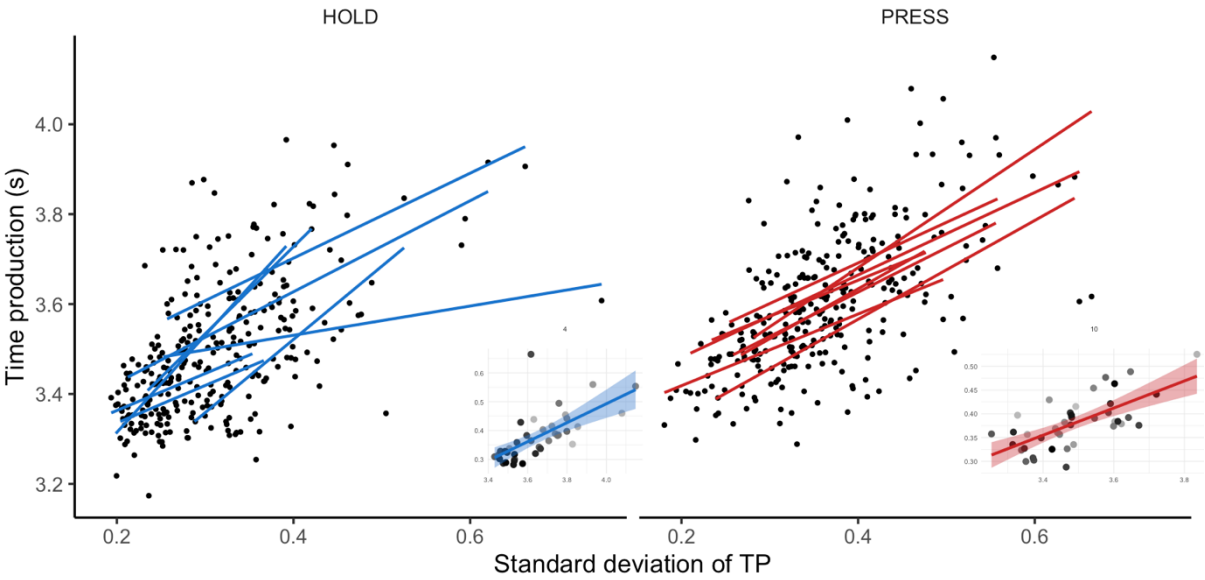
S2. Time Production distributions for HOLD and PRESS groups



Supplementary Figure 2. *Time Production distributions for HOLD and PRESS* groups displayed in Fig. 1. Premature, impulsive responses were removed from the plot to accentuate performance of interest around 3.2 s target duration.

S3. Performance in Time Production differs between HOLD and PRESS groups *(related to figure 1B)*

In the time production task (Fig. 1), animals had to produce time intervals by making two presses spaced by the minimum amount of time (PRESS) or hold pressed a lever (HOLD) for a minimum amount of time T (T = 3.2s) in order to get access to a reward. All rats were able to produce the required time interval T, as TP distributions for all rats peaked after T (Fig. 2a). The HOLD group was more accurate and precise than the PRESS group (Fig. 2B), which was evident in a model comparing mean and standard deviation of TP between the two groups across all sessions (β = 0.11, $t(551)$ = 8.86, SE = 0.012 $p < 10^{-9}$; β = 0.05, $t(553)$ = 7.2, SE = 0.007, $p < 10^{-9}$, respectively). Despite this difference, but consistent with scalar variability in duration reproduction tasks (Jazayeri & Shadlen 2010), both groups showed increased variability with increasing mean TP (Supp. Fig. 3). This was confirmed in a regression model predicting standard deviation of TP with mean TP (β = 0.28, $t(586)$ = 15.4, SE = 0.18, $p < 10^{-14}$). The addition of a group factor to the model was not justified (ΔAIC = 0.3, p = 0.155), suggesting that scalar variability was preserved in both groups in a similar form (Supp. Fig. 3).

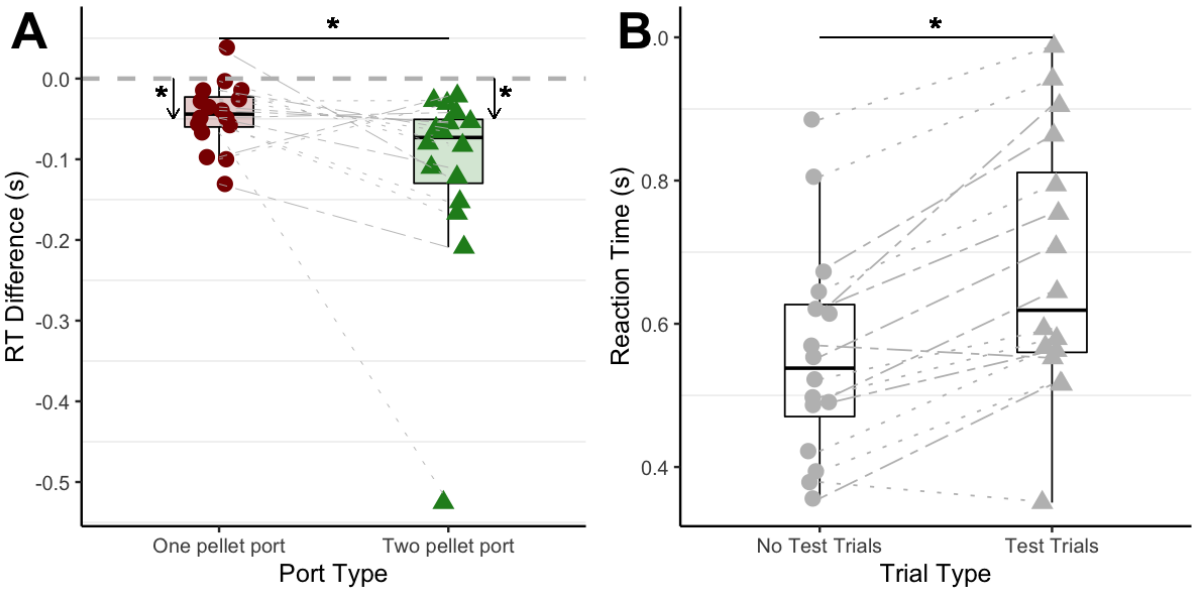


Supplementary Figure 3. For each session and each rat, μ(TP) is plotted against σ(TP) across sessions. Colored lines depict the regression fits for individual rats. Both panels contain inset plots showing a fit of one example rat.

S4. RT reflects the prospect of larger reward

As an indirect measure of temporal error monitoring, we analyzed reaction times (RT). We hypothesized that RTs should be sensitive to reward anticipation based on reward prospect.

To assess the possibility of error monitoring in rats, the animals were trained to associate TP accuracy (small error: TP_{SE} trials vs large error: TP_{LE} trials) with reward delivery ($TP_{SE\text{-}2p}$ and $TP_{LE\text{-}1p}$ trials) or its absence ($TP_{SE\text{-}0p}$ and $TP_{LE\text{-}0p}$ trials) in one of the two ports. The prospect of reward was indicated to animals, after TP was terminated, by a light cue in one of the two ports. To check if reward will indeed be delivered on any given trial, animals were required to first visit a port. Only upon the port visit did reward delivery or its absence become apparent to the animal. From the animal's perspective, light cue to a particular port only indicates the delivery side. However, if rats are sensitive to temporal errors, RTs should be related to the accuracy of TPs, as classical work has shown that rats respond faster when expecting larger rewards (Rosenbaum, 1951). Indeed, if rats produced accurate TP_{SE}, their RT to the light cue in a port associated with 2 pellets should be more rapid than RTs following the production of less accurate TP_{LE}, as 0 pellet would then be expected in that port. The opposite dependence should hold for the other port (1 pellet port). This was evident in the relative RTs comparing rewarded to unrewarded trials in a given port (Supp. Fig. 4A; RT $TP_{LE\text{-}1p}$ - RT $TP_{SE\text{-}0p}$ (dark red); RT $TP_{SE\text{-}2p}$ - RT $TP_{LE\text{-}0p}$ (green), 1-pellet and 2-pellets port, respectively). The majority of rats responded faster to a light cue on rewarded trials than on the non-rewarded trials in each port. This effect was confirmed with a mixed model testing differences between the rewarded and unrewarded trials. A significant intercept in the model indicated that the RT difference was smaller than zero (β = -0.018 s) showing that RTs were faster on trials with the reward prospect (SE = 0.008, $t(34)$ = -2.11, p = 0.043). In line with the reward prospect interpretation, rats responded also faster when reward prospect was larger (2p vs 1p), which was evident as a significant factor of port type in the model (1p vs 2p; β = -0.051, SE = 0.009, $t(1127)$ = -5.35, $p < 10^{-6}$). The shortening of RT in anticipation of a larger reward prospect strongly suggests that rats may monitor their temporal errors.



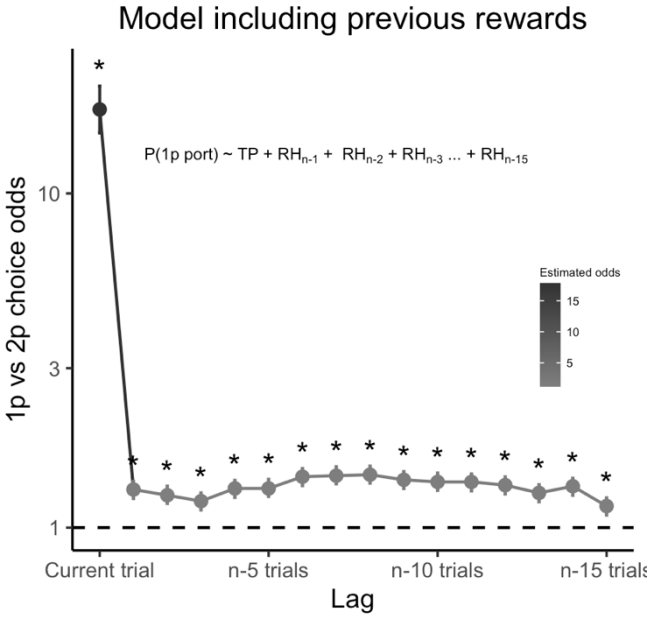
Supplementary Figure 4. Rats reacted faster to larger subjective internally-based reward prospect.

(A) Data points from one column correspond to one rat. Each data point shows a reaction time difference between rewarded and unrewarded trials for 1p port (dark red; $TP_{LE\text{-}1p} - TP_{SE\text{-}0p}$) and 2p port (green; $TP_{SE\text{-}2p}$ - $TP_{LE\text{-}0p}$). A majority of rats reacted faster to rewarded trials as evident in negative values. Additionally, reaction time difference was larger for 2p trials.

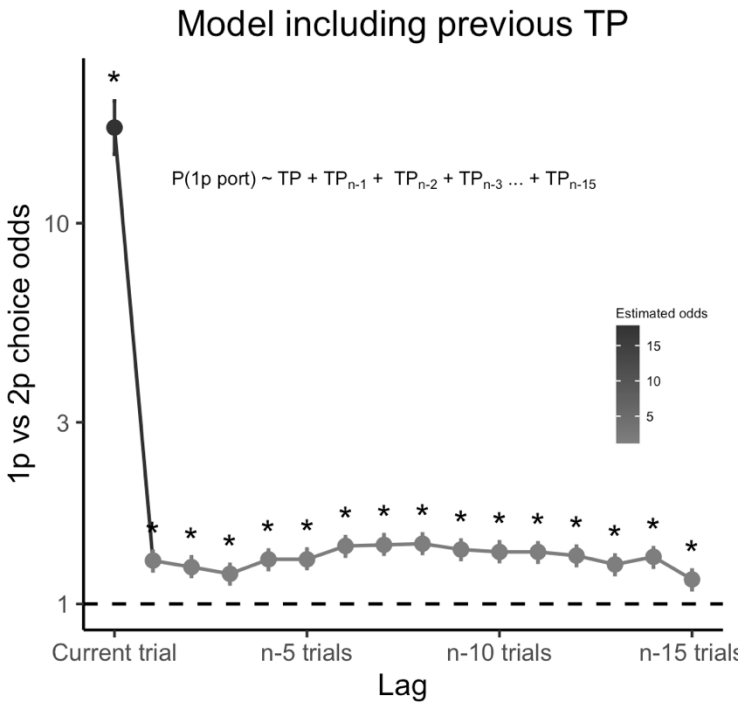
(B) RT on test-trials indicated choice behavior. Each individual data point shows a reaction time in training-trials (no choice) and test-trials (choice available). Rats showed longer RT during test-trials than during the training-trials (*ß* = 0.14, *SE* = 0.011, $t(1163) = 9.54$, $p < 10^{-15}$), indicating expression of choice behavior as opposed to random responding.

S5. RT reflects the prospect of larger reward

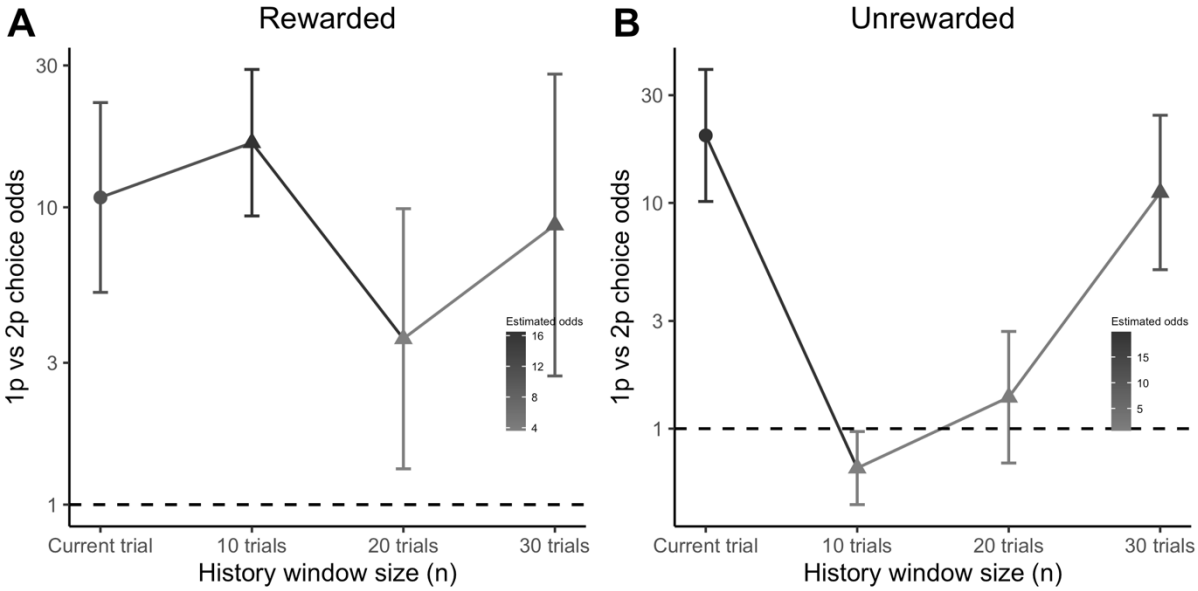
To assess whether rats used previous trial outcomes, we fitted the model that included current TP and previous TPs, or previous rewards (RH) up to 15 trials back (n-15). Indeed, the estimated odds showed that previous rewards contributed to rats' choices (Supp Fig. 5a; n=0: *ß* = 2.87, *SE* = 0.083, $z(18224) = 34.27$, $p < 10^{-15}$; n-1: ß = 0.25, SE = 0.031, $z(18224) = 8.02$, $p < 10^{-14}$; see Supp Table 2), but with clear evidence for the strongest contribution of current TP. A similar pattern of results was obtained when current TPs were fitted together with previous TPs (Supp Fig. 4; n=0: ß = 2.82, $z(18205) = 33.71$, $p < 10^{-15}$; n-1: ß = 0.51, $z(18205) = 8.59$, $p < 10^{-15}$; see Supp Table 3).



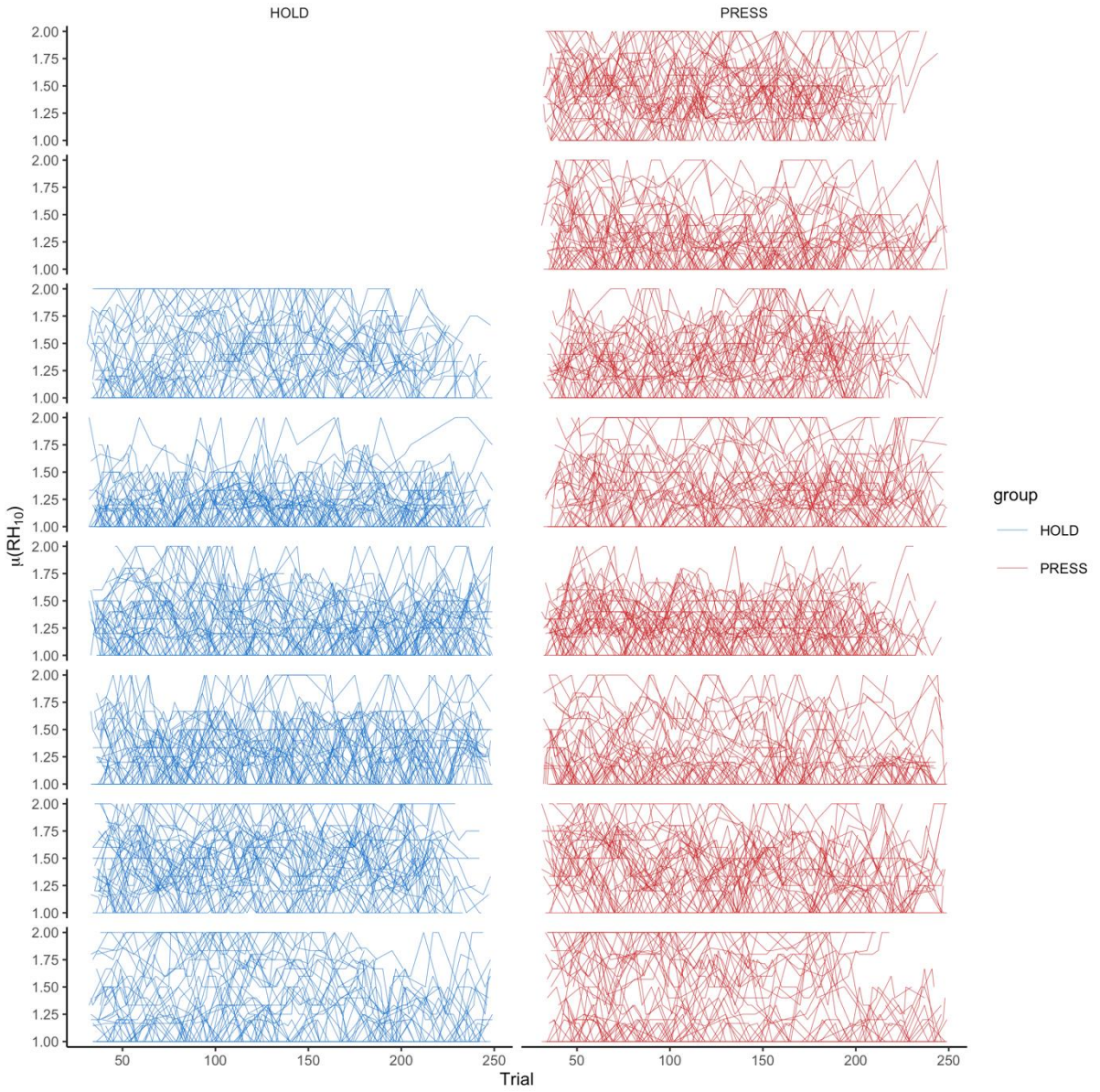
Supplementary Figure 5A. Results of GLMM fit including current TP and reward history (RH) incorporated as rewards received on up to 15 previous trials. The formula in the plot shows specification of fixed terms in the GLMM. Asterisks indicate significant terms. Error bars display confidence intervals of estimated effects.



Supplementary Figure 5B. Results of GLMM fit including current TP and 15 previous trials as a proxy for reward history. As opposed to Supp. Fig. 4A where amount of reward was used, here TP were used. The formula in the plot shows specification of fixed terms in the GLMM. Asterisks indicate significant terms. Error bars display confidence intervals of estimated effects.



Supplementary Figure 6. Results of GLMM fit including current TP and running average of last 10, 20, 30 TPs. TP history predictors were computed separately for rewarded (A) and unrewarded trials (B). Significant model terms indicate that rats keep track of temporal errors on current TP as well as TP history.

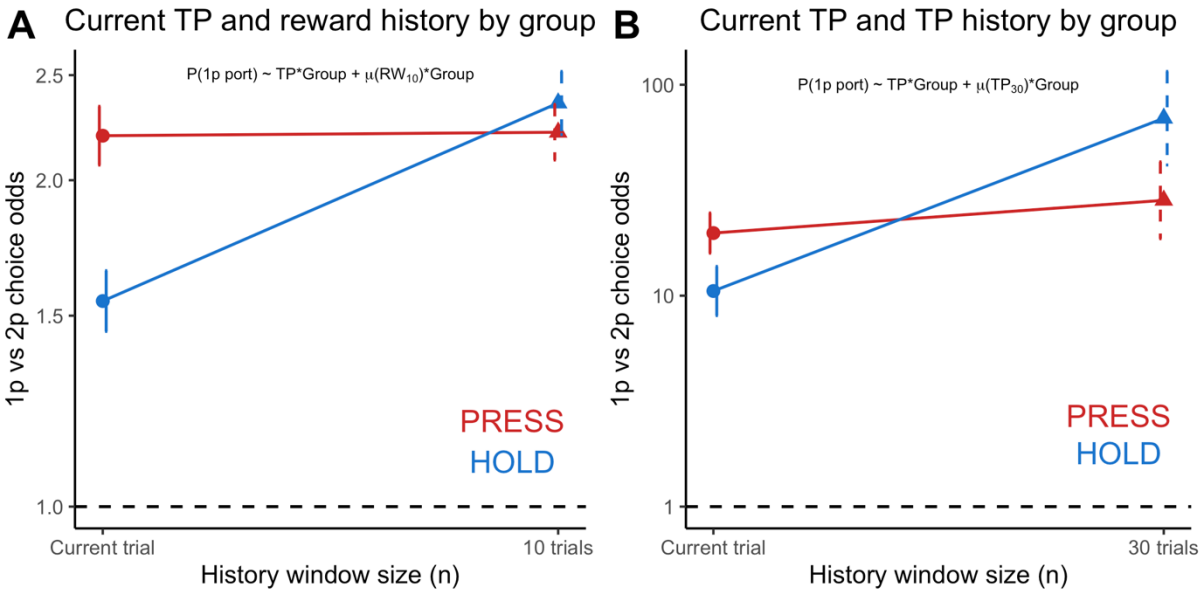


Supplementary Figure 7. Each panel depicts reward history (RH) traces for individual rats. Each line originates from a single experimental session.

S8*. HOLD and PRESS groups keep track of temporal errors and reward history differently*

To understand what behavioral features are related to monitoring of temporal errors, we hypothesized that HOLD and PRESS groups may differently utilize current temporal errors and previous trial information such as reward history. We hypothesized that owing to its wider TP distribution (Fig. 1B), the PRESS group could achieve better choice performance (Fig. 2). To test whether PRESS and HOLD groups achieved their choice behavior differently, we re-fitted the model including TP_n and the strongest factors reflecting previous trial information, that is μRH_0 and μTP_{30} (see Fig. 3BC). We added an interaction with group factor to each predictor (Supp Fig. 7). As hypothesized, we found that HOLD and PRESS groups relied differently on current trial TP (Fig S6A, TP_n * Group: β = 0.33, *SE* = 0.046, $z(16426) = 7.21$, $p < 10^{-12}$; Fig S6B, TP_n * Group: β = 0.53, *SE* = 0.18, $z(16965) = 2.97$, $p = 0.003$) and previous trial information (Fig S6A, μRW_{10} * Group: β = -0.09, *SE* = 0.047, $z(16426) = -1.97$, $p = 0.049$; Fig S6B, μTP_{30} * Group: β = -1.08, *SE* = 0.34, $z(16965) = -3.12$, $p = 0.002$). Notably the PRESS group tended to rely more on

the current TP, whereas the HOLD group relied more on previous trial information (TP and reward history), suggesting that the nature of motor sequence and associated variability contributed differently to choice behavior. Nevertheless, both PRESS and HOLD groups still relied on the current trial TP.

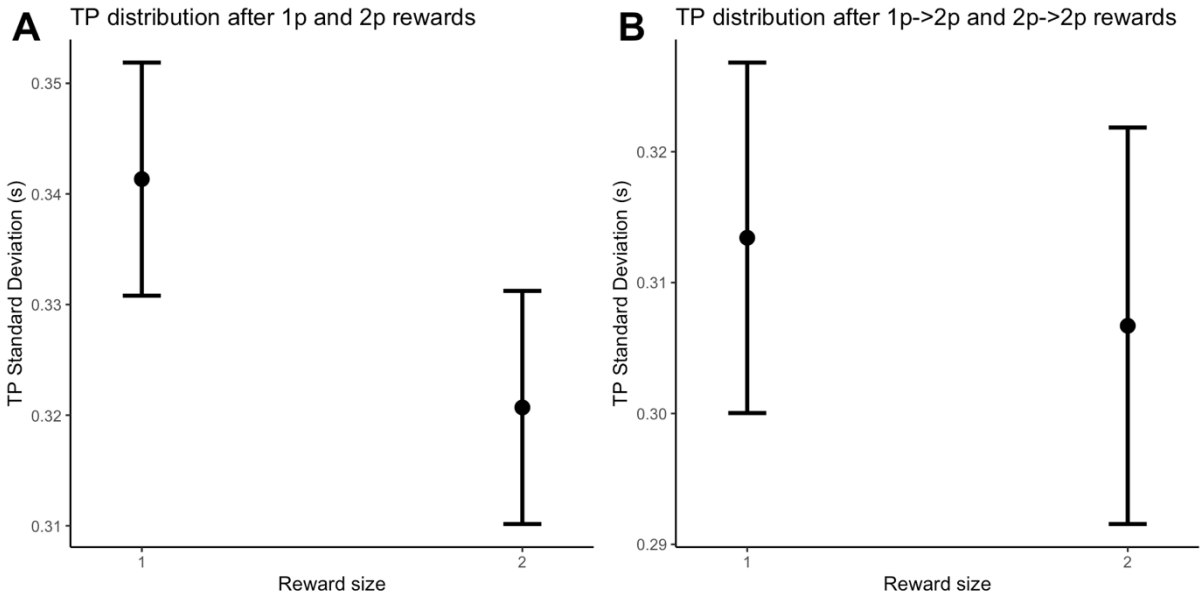


Supplementary Figure 8. Results of GLMM fit including current TP and previous trial history. In conjunction with current TP, we tested μRW_{10} (A) and μTP_{30} (B).

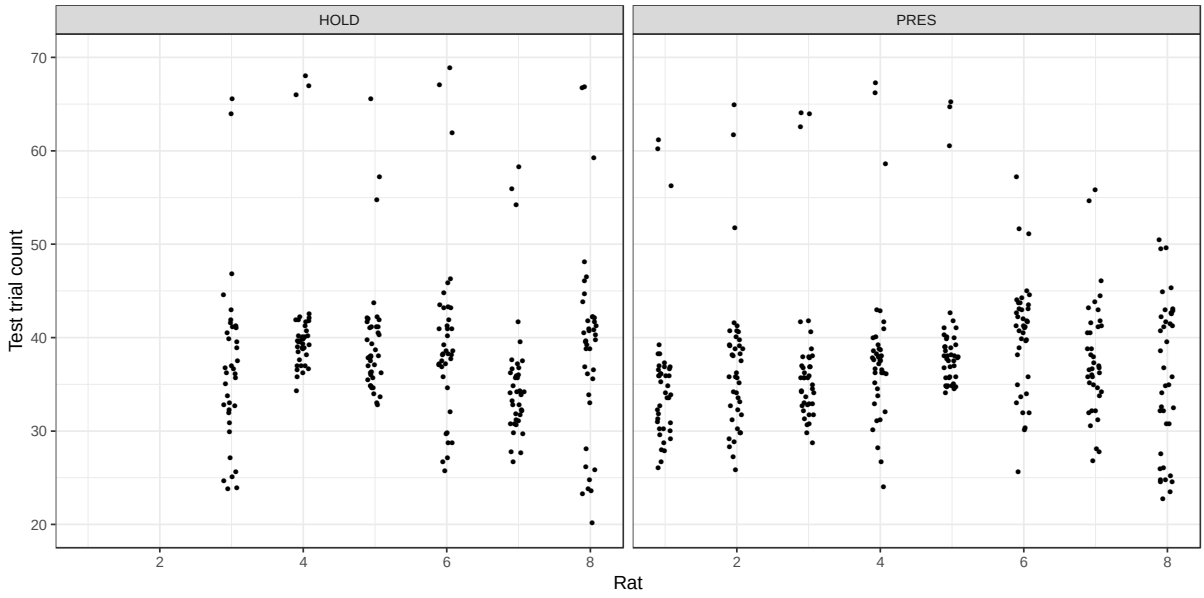
S9. Behavioral variability and hunting behavior (related to figure 4C)

We considered a possibility that the hunting index could be affected by the degree of variability of a given rat, and that, as a result, it could affect the association between hunting and error monitoring. We addressed that concern in two ways. We verified that the addition of standard deviation of TP for each rat to the model including TP was not justified (Wald test: $W(1) = 0.03$, $p = 0.866$).

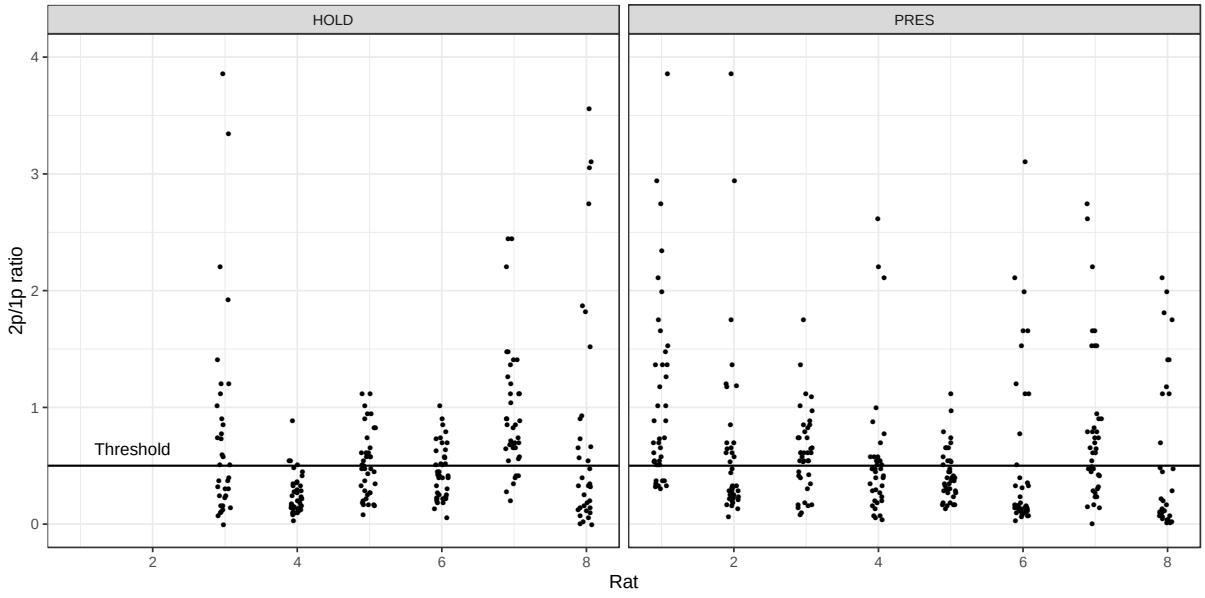
The other verification was based on the previous work in rat, monkey and humans indicating that reward magnitude calibrates performance variability on the next trial (Wang et al., 2019; Dhwale et al., 2019). We confirmed that TPs that followed $TP_{SE\text{-}2p}$ had smaller variability than TPs that followed $TP_{LE\text{-}1p}$ ($t(1009) = -4.13$, $p < 10^{-4}$; Supp. Fig. S9A). Given that we replicated the impact of variability in our data set, we compared TPs that followed a $TP_{LE\text{-}1p}$ - $TP_{SE\text{-}2p}$ sequence with those that followed a $TP_{SE\text{-}2p}$ - $TP_{SE\text{-}2p}$ sequence. TPs that followed these two types of sequences did not differ ($t(882) = -0.867$, $p = 0.39$; Supp. Fig. S9B). Together, these two analyses suggest that TP variance did not mediate the relationship between propensity of individual rats to monitor temporal error and to hunt for minimizing it.



Supplementary Figure 9. Behavioral variability and hunting behavior (related to figure 4C).
(A) Standard deviation of TP preceded (n-1 trial) by 2p trials is smaller than those preceded by 1p trials.
(B) Standard deviation of TP preceded by '1p -> 2p' and '2p -> 2p' trial sequences (n-2 and n-1) did not differ.



Supplementary Figure 10. The number of test-trials obtained per session per rat.



Supplementary Figure 11. 2p/1p ratios obtained per session per rat. The horizontal black line displays Θ threshold.

Supplementary Tables

Group	Rat number	Wilcoxon test	p value	Session count > 55%
HOLD	1	96	4.4e-02	6
PRES	1	810.5	7.6e-08	35
HOLD	2	100	3.2e-03	3
PRES	2	519	1.9e-06	21
HOLD	3	595	3.8e-07	34
PRES	3	820	3.7e-08	35

Group	Rat number	Wilcoxon test	p value	Session count > 55%
HOLD	4	774	1.0e-06	24
PRES	4	805	1.1e-07	31
HOLD	5	754	3.9e-06	26
PRES	5	775	8.1e-08	33
HOLD	6	583	8.9e-05	23
PRES	6	727	2.7e-06	25
HOLD	7	811	7.3e-08	32
PRES	7	695	2.3e-07	31
HOLD	8	669.5	9.9e-05	24
PRES	8	455	1.9e-03	18

Supplementary Table 1. Results of two-sided Wilcox test performed on test-trial accuracy for each rat separately. Test-trial accuracy on a single session served as a unit of observation for statistical testing. The last column displays the number of session where test-trial accuracy was above 55%.

TP Lag (n)	Beta estimate	z value	p value
0	16.94	34.4	3.0e-259
1	1.64	8.51	1.7e-17
2	1.43	6.25	4.0e-10
3	1.14	2.36	1.8e-02
4	1.22	3.47	5.1e-04
5	1.23	3.68	2.3e-04
6	1.25	3.8	1.4e-04
7	1.18	2.84	4.5e-03
8	1.3	4.59	4.5e-06
9	1.26	4.04	5.4e-05
10	1.1	1.65	9.8e-02
11	1.28	4.3	1.7e-05
12	1.24	3.88	1.0e-04
13	1.14	2.36	1.8e-02
14	1.3	4.61	4.0e-06
15	1.05	0.81	4.2e-01

Supplementary Table 2. Detailed results of GLMM fit including current TP and rewards received on 15 previous trials. The values correspond to the data visualized in Supp. Fig. 4A.

TP Lag (n)	Beta estimate	z value	p value
0	17.69	34.27	2.4e-257
1	1.28	8.02	1.1e-15
2	1.24	6.96	3.5e-12
3	1.19	5.53	3.3e-08
4	1.29	8.35	6.6e-17
5	1.29	8.25	1.5e-16
6	1.39	10.88	1.5e-27
7	1.42	11.4	4.1e-30
8	1.44	11.81	3.4e-32
9	1.38	10.54	5.9e-26
10	1.36	9.91	3.7e-23
11	1.35	9.86	6.3e-23
12	1.32	9.06	1.4e-19
13	1.26	7.52	5.5e-14
14	1.31	8.87	7.5e-19
15	1.14	4.16	3.1e-05

Supplementary Table 3. Detailed results of GLMM fit including current TP and 15 previous TPs as a proxy for reward history and produced intervals. The values correspond to the data visualized in Supp. Fig. 4B.