Duration-specific effects of outcome devaluation in temporal control are differentially sensitive to amount of training

Sho Araiba,[1] Nicole El Massioui,[2] Bruce L. Brown,[1,3] and Valérie Doyère[2]

[1]Graduate Center of the City University of New York, New York, New York 10016, USA; [2]Institut des Neurosciences Paris-Saclay (Neuro-PSI), UMR 9197, Université Paris Sud, CNRS, Université Paris Saclay, Orsay 91405, France; [3]Queens College, Queens, New York 11367, USA

This study demonstrates that overtraining in temporal discrimination modifies temporal stimulus control in a bisection task and produces habitual responding, as evidenced through insensitivity to food devaluation. Rats were trained or overtrained in a 2- versus 8-sec temporal discrimination task, with each duration associated with a lever (left or right) and food (grain or sucrose). Overtraining produced a leftward shift in the bisection point. Devaluation treatment induced a differential loss of responding depending on stimulus duration (short versus long) and the level of training (training versus overtraining). The relationships between timing behavior and habitual behavior are discussed.

[Supplemental material is available for this article.]

Interval timing is a research area that investigates the mechanisms permitting the organism to perceive time, and extract durations or intervals between stimuli or events. The ability to discriminate time intervals has been widely demonstrated through the temporal bisection procedure (Church and Deluty 1977) in which rats learned to discriminate between two reinforced durations (2- versus 8-sec), followed by temporal bisection sessions with five added geometrically spaced intermediate nonreinforced durations. Rats bisected pairs of durations at their geometric mean (GM), which means that the point of subjective equality (PSE) (the duration corresponding to p[long] = 0.50) fell at the GM of each set of anchor durations. In contrast to the ratio similarity rule (Gibbon 1981) postulating that the location of the PSE is determined by the values of the short and long anchor durations, some recent studies suggested that the location of the PSE might rather depend on a comparison of the short anchor with any longer duration. That is, only one of the anchor stimuli acquired control over the behavior with a prolonged exposure to the discrimination (Platt and Davis 1983; Machado and Keen 2003; Callu et al. 2009; Brown et al. 2011; Araiba and Brown 2017).

One possible explanation for the gain of control by one discriminative stimulus with prolonged exposure to training and/or testing is the development of a habit. Habit formation refers to the process by which a goal-driven performance (action-outcome association) developed during training becomes a habit-based performance solely controlled by antecedent stimuli (stimulus-response association) that progressively gain exclusive domination over behavior with extensive training (e.g., Adams and Dickinson 1981; Dickinson 1985). Recent neurological studies indicated that both timing behavior and habit formation share the same neural networks (for review, see Doyère and El Massioui 2016). The present study examined the potential development of habit-based timing performance using a classic devaluation procedure of the instrumental outcome by a specific satiety treatment in a temporal

discrimination task. After moderate amounts of training, reward devaluation is expected to reduce responding on trials with the stimulus associated with that reward, but not on trials with the nondevalued reward, in accord with the devaluation literature (e.g., Dickinson 1985; Faure et al. 2005). After overtraining, however, the instrumental response becomes habitual, driven by the stimulus rather than by the reward properties of the outcome, and thus is less sensitive to reward devaluation.

During training and overtraining (see experimental design, Fig. 1A) of a temporal discrimination (2 versus 8 sec), correct responding to the short and long durations were associated with different actions (pressing the left or right levers) and different rewards (grain versus sucrose pellets) (Fig. 1B). The temporal discrimination training procedure was adapted from Callu et al. (2009). In brief, rats learned the discrimination in three phases. First (100% Forced choice), a tone stimulus was presented for either 2 or 8 sec and its termination coincided with the presentation of one lever, associated with the correct response and delivery of the reward for a response (40 trials for each duration per session). Second (50% Forced/50% Free choice), 40 trials were identical to the 100% Forced choice phase, and for the other 40 Free-choice trials, the 2- or 8-sec tone stimulus termination coincided with the presentation of both levers. Only a correct lever press produced the associated food. In the third phase (0% Forced/100% Free choice), all 80 trials were free-choice. All animals received discrimination training until accuracy reached 85% correct responding for two consecutive sessions. Animals of the overtraining group received 600 additional trials for each stimulus/action/reward pairing after they met the mastery criterion. The average number of training sessions across the three training phases was similar for the training and the overtraining groups (10.77 and 10.26, respectively, $t_{(43)} = 0.71$, $P = 0.48$). Similarly, the mean accuracy of discrimination performance for the training and the overtraining

Corresponding authors: valerie.doyere@u-psud.fr; sho.araiba@gmail.com
Article is online at http://www.learnmem.org/cgi/doi/10.1101/lm.047878.118.
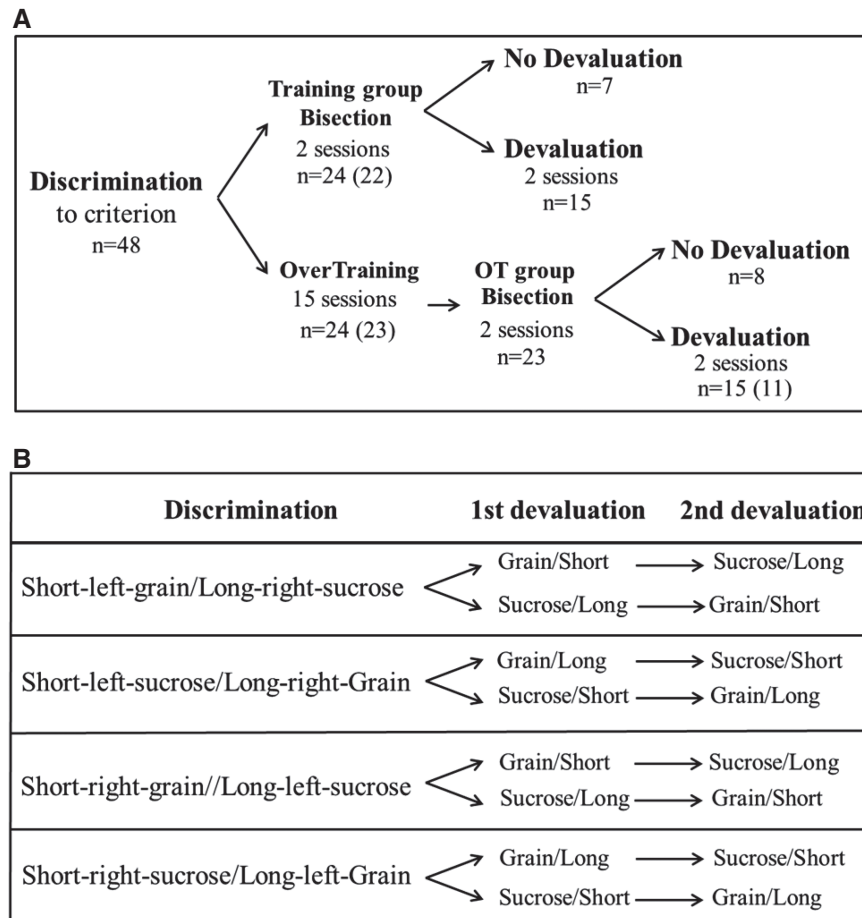
A

B

Figure 1. (*A*) General design of the experiment. The number under each group is the initial number of trained or overtrained animals; the number in parentheses is the final number of animals included in statistics, as three rats were eliminated because of experimental disturbances; (OT) overtraining. (*B*) Description of the experimental groups (within each training and overtraining group). For the discrimination, duration-lever-food assignments (duration: short versus long, position of the lever: *left* versus *right* and reward: grain versus sucrose pellets) were counterbalanced between rats, yielding four discrimination groups. During devaluation tests, two sessions were run to counterbalance food/anchor devaluation (for example, devaluating the food associated with the short anchor in the first test and the food associated with the long anchor in the second test), thus creating eight subgroups. The performance on the lever associated with the duration whose outcome was devalued was then compared with the performance on the lever associated with the other duration whose outcome was not devalued within the same session.

$F_{(2.135,91.786)} = 10.15$, $P < 0.001$, $\eta_p^2 = 0.191$ (Fig. 2). At both anchor durations, there were no differences between groups in $p(\text{long})$, whereas a maximum group mean difference was observed at the 4 sec intermediate duration, $t_{(43)} = 4.340$, $P < 0.001$, Cohen's $d = 1.294$. The difference in the bisection function between the two groups was reflected by a smaller PSE for the overtraining group than for the training group (group mean PSE for training: 3.63 (SEM = ±0.05), and for overtraining: 3.31 (SEM = ±0.06); $t_{(43)} = -4.11$, $P < 0.001$, Cohen's $d = -1.227$). There was no significant difference in gamma (index of precision, inversely related to the slope) between the groups (training: 0.21 (SEM = ±0.01) and overtraining: 0.18 (SEM = ±0.02; $t_{(43)} = 1.61$, $P = 0.11$). Thus, the bisection curve was shifted to the left after overtraining with no significant change in the slope of the function, which suggests that overtraining produces a change in bias toward long duration judgments while leaving temporal precision stable.

After bisection testing, each group was divided into two subgroups, the nondevaluation and the devaluation groups (Fig. 1A). Two devaluation sessions were given, separated by two retraining sessions of discrimination "100% Free choice." Devaluation sessions consisted of three phases (see Faure et al. 2005 for full details): (1) the prefeeding during which rats had access to 50 g of either grain or sucrose pellets for 1 h in a rectangle box. Rats from the nondevaluation group were subjected to the same procedure without food; (2) a devaluation test session in extinction, similar to the last discrimination training sessions, with 20 trials for each duration; (3) a satiety test during which rats were presented 50 pellets of grain and sucrose in two separate dishes for 5 min in a shoebox cage. A 2
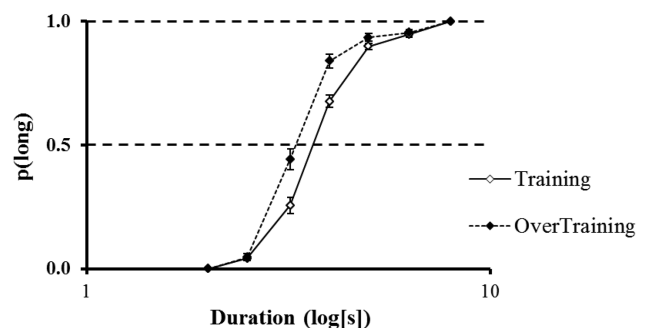
groups on the last two sessions where they met the criterion was similar (94.29% and 95.43%, respectively; $t_{(43)} = 1.774$, $P = 0.083$). The accuracy of discrimination performance during the overtraining phase was stable, averaging 98.18% across 15 sessions.

Twenty-four hours after mastering the discrimination or after overtraining, the extent to which temporal discrimination was affected by extended training was assessed with bisection test sessions. Rats were tested in a psychophysical choice (bisection) procedure in two daily sessions with five intermediate durations (2.5, 3.2, 4, 5, and 6.3 sec) on nonreinforced trials (12 trials each duration), in addition to the two training anchor durations (2 and 8 sec, 60 trials each) with reinforcement available. The bisection procedure is extensively described in Callu et al. (2009) and Es-seddiqi et al. (2016), and its detailed analysis is explained in Supplemental Material. A 2 (group) × 7 (duration) ANOVA of $p(\text{long})$ yielded a main effect of group, $F_{(1,43)} = 16.57$, $P < 0.001$, $\eta_p^2 = 0.278$, a main effect of duration, $F_{(2.135,91.786)} = 1149.85$, $P < 0.001$, $\eta_p^2 = 0.964$, and a group × duration interaction,

Figure 2. Number of "long" responses divided by the sum of the "long" responses and the "short" responses for each stimulus across the two sessions of testing (group mean $p(\text{long})$) as a function of duration (log scale) for the training ($n = 22$) and overtraining ($n = 23$) groups. Error bars represent ±SEM.

(group) × 2 (food) mixed ANOVA analysis of the amount of food in grams consumed by each group during prefeeding for sucrose and grain food pellets (Devaluation 1 and 2) shows no effect of group, $F_{(1,24)} = 2.84$, $P = 0.105$, a main effect of food-type (35.9 g grain versus 30.6 g sucrose), $F_{(1,24)} = 20.75$, $P < 0.001$, $\eta_p^2 = 0.464$, and no group × food-type interaction, $F < 1$. During the satiety test, there was less consumption of the devalued food (44.94 pellets remaining) compared to the nondevalued food (8.92 pellets remaining), with no effect of group, $F < 1$, a main effect of food devaluation, $F_{(1,24)} = 274.7$, $P < 0.001$, $\eta_p^2 = 0.920$, and no group × food devaluation interaction, $F < 1$, thus showing selectivity of the devaluation procedure similarly for both groups.

During the devaluation test in extinction, response tendency was calculated based on the first 18 trials (nine for each duration), corresponding to the minimum number of trials for which short and long duration presentations were balanced before a loss of responding due to extinction. The specific effect of devaluation was analyzed in the devalued animals by comparing responding to the duration associated with the devalued food (i.e., devalued duration) to responding to the other, nondevalued, anchor duration, and analyzing them in three categories, i.e., correct, incorrect or no response, in separate 2 (devaluation) × 2 (anchor duration) within-subject ANOVAs.

For the training group, the proportion of "no response," which may reflect a general decrease in motivation, as reported in previous studies (Ward and Odum 2006, 2007; Galtress and Kirkpatrick 2009, 2010; McClure et al. 2009), was not differentially affected by duration (no main effect of devaluation, $F_{(1,14)} = 1.67$, $P = 0.22$, no main effect of anchor duration, $F_{(1,14)} = 1.76$, $P = 0.21$, and no devaluation × anchor duration interaction, $F_{(1,14)} = 2.38$, $P = 0.15$, see Fig. 3 left lower panel). In contrast, correct responses were differentially modulated, with no main effect of devaluation, $F_{(1,14)} = 1.93$, $P = 0.19$, or anchor duration, $F < 1$, but with a devaluation × anchor duration interaction, $F_{(1,14)} = 18.01$, $P < 0.001$, $\eta_p^2 = 0.563$. When the analysis was restricted to each anchor duration (Fig. 3, left upper panel), there was a significant decrease in the proportion of correct responses for the devalued duration trial as compared to the nondevalued duration trial when the food associated with the long anchor was devalued, Paired $t_{(14)} = -3.33$, $P = 0.005$, Cohen's $d = -0.859$, whereas no significant difference was obtained when the food associated with the short anchor was devalued, Paired $t_{(14)} = 1.57$, $P = 0.14$. The analysis of incorrect responses essentially mirrored the one for correct responses (Fig. 3, left middle panel) with no main effect of devaluation, $F < 1$, a main effect of anchor duration, $F_{(1,14)} = 8.53$, $P = 0.011$, $\eta_p^2 = 0.379$, and a devaluation × anchor duration
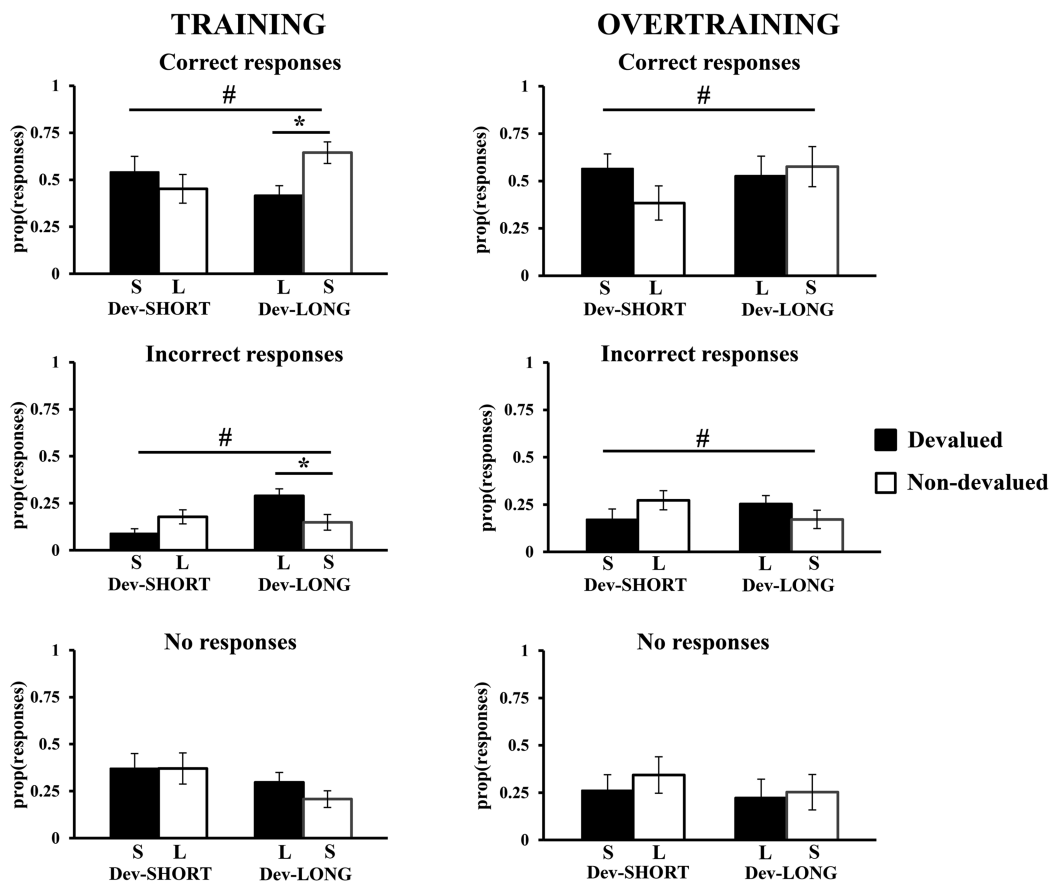

Figure 3. The mean proportion of correct (*upper* histograms), incorrect (*middle* histograms) responses and the proportion of "no response" (*lower* histograms) out of total number of trials for each duration ($n = 9$) during the devaluation extinction session as a function of the devalued anchor duration (*x*-axis) for the associated devalued (black) and nondevalued (white) duration trials (bars) for the training group (*left* panels), and the overtraining group (*right* panels). Devaluation of the outcome (sucrose or grain) associated with a given duration (i.e., short in a Dev-SHORT session, and long in a Dev-LONG session) was expected to affect lever pressing associated with that devalued stimulus (black bars, i.e., short (S) after Dev-SHORT, and long (L) after Dev-LONG), in comparison to the lever pressing associated with the nondevalued duration within the same session for the same rats (white bars, i.e., long (L) after Dev-SHORT, and short (S) after Dev-LONG). Only devaluation of the outcome associated to the long stimulus had an impact in the training group, an effect that was reduced in the overtraining group. Error bars represent ±SEM. #, significant devaluation × anchor duration interaction, $P < 0.05$; *, significant difference between % response to devalued and nondevalued stimuli, $P < 0.05$.

interaction, $F_{(1,14)} = 11.08$, $P = 0.005$, $\eta_p^2 = 0.442$. Again, there was a significant difference in the proportion of incorrect responses between the devalued and the nondevalued duration trials when the long anchor duration was devalued, $t_{(14)} = 2.62$, $P = 0.02$, Cohen's $d = 0.678$; that is, devaluation produced an increase in errors. When the food associated with the short anchor duration was devalued, there was a trend toward a difference, $t_{(14)} = -2.04$, $P = 0.06$, albeit in the opposite direction (i.e., a decrease in errors after devaluation).

For the overtraining group, a 2 (devaluation) × 2 (anchor duration) ANOVA on the proportion of correct responses as a function of the devalued anchor duration for the devalued and the nondevalued durations (Fig. 3, right upper panel) yielded no main effect of devaluation, $F < 1$, no main effect of anchor duration, $F < 1$, and a devaluation × anchor duration interaction, $F_{(1,10)} = 6.25$, $P = 0.031$, $\eta_p^2 = 0.385$. When restricted to each anchor duration, the analyses did not yield a significant difference, $t_{(10)} = 1.80$, $P = 0.101$ and $t_{(10)} < 1$ for short and long anchor devaluation, respectively. Analysis of incorrect responses yielded no main effect of devaluation, $F < 1$, no main effect of anchor duration, $F < 1$, and a devaluation × anchor duration interaction, $F_{(1,10)} = 8.50$, $P = 0.015$, $\eta_p^2 = 0.460$ (Fig. 3, right middle panel). Again, when restricted to each anchor duration, the analyses did not yield a significant difference, $t_{(10)} = -1.71$, $P = 0.118$ and $t_{(10)} = 1.70$, $P = 0.12$ for short and long anchor devaluation, respectively. Finally, analysis of no responses yielded no main effect of devaluation, $F < 1$, or anchor duration, $F < 1$, and no devaluation × anchor duration interaction, $F < 1$.

In all, overtraining induced two main effects depending on the test used. First, overtraining the 2 sec versus 8 sec discrimination task produced a leftward shift in the bisection point. It reflected a decreased point of subjective equality (PSE) compared to the training group, with no modification of the temporal sensitivity (gamma). Brown et al. (2011) already noted a smaller PSE with bisection repetition but also observed an increased temporal sensitivity with repeated testing. It is possible that the repeated exposure to intermediate durations in bisection tests in Brown et al. (four to six bisection sessions were run each month, from 4 to 8 mo) was responsible for the change in gamma. Alternatively, the training procedure may have already led to modified temporal precision, precluding any further change with overtraining. The present study demonstrates that repeated exposure to anchors alone is sufficient to produce a change in bias toward long duration judgment while potentially leaving temporal precision stable. Second, devaluation of the instrumental reward by a specific satiety treatment produced a differential loss of accurate responding depending on stimulus duration (short versus long) and the level of training (training versus overtraining).

The devaluation procedure that followed the bisection tests could provide information about the processes involved during overtraining of the discrimination task. After training as well as overtraining, performance to the short anchor duration was not altered by the selective reward devaluation, indicating insensitivity to the current value of the reward, and thus the possible formation of a habit during the training phase. In contrast, after training, the long anchor duration showed sensitivity to reward devaluation evidenced by decreased instrumental performance thus suggesting goal-directed behavior. This sensitivity to devaluation decreased after overtraining, indicating that habitual responding to the long anchor duration stimulus was acquired with session repetition. Interestingly, these effects were accompanied by an increase in the proportion of incorrect responding, when the long anchor duration was devalued in the training group. Whether the increase in incorrect classification, instead of "no responses," was due to the two-choice discrimination procedure or was related to the timing aspect of the task remains to be determined. Nevertheless, this effect also disappeared after overtraining. This result demonstrates

that time-based behavior can become habitual in a temporal discrimination task, in agreement with the loss of DA sensitivity observed by Cheng et al. (2007) in a peak interval task. The present results augment those findings by demonstrating habit formation with overtraining in a devaluation test. That the rate of lever-pressing to the short anchor duration was unchanged after devaluation of its associated reward constitutes asymmetry, and indicates that durations of stimuli can be processed differently in a goal-directed or habitual mode. It represents another example of asymmetry in independent manipulation of short and long durations in temporal discrimination tasks (Akdoğan and Balcı 2016). The insensitivity to devaluation of the reward associated with the short duration could be interpreted as a more rapid development of habit-based behavior than for the long response. One possible basis for a more rapid habit formation for the short duration is that it is more often presented to the rats, as short durations are always embedded in long durations. Another possibility is that once rats had learned the discrimination rule (duration/lever position/ reward), they developed a behavioral pattern commencing with approach to the short lever, which transforms more rapidly into an S-R habit behavior, i.e., triggered by stimulus onset and independent of long stimulus duration. Machado and Keen's (2003) inference of dominant control by the short anchor duration was based on the observation that manipulation of the long anchor duration had little or no effect on the subjects' times of departure from the vicinity of the short cue on tests with long-cue durations. While the present study used rats in a standard chamber instead of pigeons in a long box, our procedure used one lever and food dispenser pair located on one wall, and the other pair located on the opposite wall of the operant box. As with pigeons, rats might have positioned themselves in front of the short lever first and moved to the other lever as stimulus duration elapsed. The number of training trials (~450 trials per duration in ~11 sessions) could thus have been enough to produce habit-based instrumental performance for short duration stimulus. This post-hoc account would require support showing sensitivity to reward-devaluation at some earlier point in training than that used for the present training group. It is conceivable that with overtraining, the motor patterns associated with approach to both short and long cue locations become habitual, and relatively independent of rewarding consequences.

In conclusion, these results may validate the hypothesis of development of a strategy based on a comparison of the short anchor duration with any longer duration (short/no-short rule) with extended training, a basis for habitual behavior.

Acknowledgments

References

Adams CD, Dickinson A. 1981. Actions and habits: variations in associative representations during instrumental learning. In *Information processing in animals: memory mechanisms* (ed. Spear NE, Miller RR), pp. 143–166. Erlbaum, Hillsdale, NJ.

Akdoğan B, Balcı F. 2016. Stimulus probability effects on temporal bisection performance of mice (*Mus musculus*). *Anim Cogn* **19:** 15–30.

Araiba S, Brown BL. 2017. The effect of the long anchor duration on performance in the temporal bisection procedure. *Behav Processes* **135:** 76–86.

Brown BL, Höhn S, Faure A, von Hörsten S, Le Blanc P, Desvignes N, El Massioui N, Doyère V. 2011. Temporal sensitivity changes with extended training in a bisection task in a transgenic rat model. *Front Integr Neurosci* **5:** 44.

Callu D, El Massioui N, Dutrieux G, Brown BL, Doyere V. 2009. Cognitive processing impairments in a supra-second temporal discrimination task in rats with cerebellar lesion. *Neurobiol Learn Mem* **91:** 250–259.

Cheng RK, Hakak OL, Meck WH. 2007. Habit formation and the loss of control of an internal clock: inverse relationship between the level of baseline training and the clock-speed enhancing effects of methamphetamine. *Psychopharmacology (Berl)* **193:** 351–362.

Church RM, Deluty MZ. 1977. Bisection of temporal intervals. *J Exp Psychol Anim Behav Process* **3:** 216–228.

Dickinson A. 1985. Actions and habits: the development of behavioural autonomy. *Philos Trans R Soc Lond B Biol Sci* **308:** 67–78.

Doyère V, El Massioui N. 2016. A subcortical circuit for time and action: insights from animal research. *Curr Opin Behav Sci* **8:** 147–152.

Es-seddiqi M, El Massioui N, Samson N, Brown BL, Doyère V. 2016. The amygdalo-nigrostriatal network is critical for an optimal temporal performance. *Learn Mem* **23:** 104–107.

Faure A, Haberland U, Condé F, El Massioui N. 2005. Lesion to the nigrostriatal dopamine system disrupts stimulus–response habit formation. *J Neurosci* **25:** 2771–2780.

Galtress T, Kirkpatrick K. 2009. Reward value effects on timing in the peak procedure. *Learn Motiv* **40:** 109–131.

Galtress T, Kirkpatrick K. 2010. Reward magnitude effects on temporal discrimination. *Learn Motiv* **41:** 108–124.

Gibbon J. 1981. On the form and location of the psychometric bisection function for time. *J Math Psychol* **24:** 58–87.

Machado A, Keen R. 2003. Temporal discrimination in a long operant chamber. *Behav Processes* **62:** 157–182.

McClure EA, Saulsgiver KA, Wynne CDL. 2009. Manipulating pre-feed, density of reinforcement, and extinction produces disruption in the Location variation of a temporal discrimination task in pigeons. *Behav Processes* **82:** 85–89.

Platt JR, Davis ER. 1983. Bisection of temporal intervals by pigeons. *J Exp Psychol Anim Behav Process* **9:** 160–170.

Ward RD, Odum AL. 2006. Effects of prefeeding, intercomponent-interval food, and extinction on temporal discrimination and pacemaker rate. *Behav Processes* **71:** 297–306.

Ward RD, Odum AL. 2007. Disruption of temporal discrimination and the choose-short effect. *Learn Behav* **35:** 60–70.