Discounting of Future Rewards and Punishments in Rats

Maurice-Philipp Zech*, Sandra Schäble*, and Tobias Kalensch

https://doi.org/10.1523/ENEURO.0452-21.2022

Comparative Psychology, Institute of Experimental Psychology, Heinrich-Heine Universität Düsseldorf, Düsseldorf, 40225, Germany

Abstract
Temporal reward discounting describes the decrease of value of a reward as a function of delay. Decision-making between future aversive outcomes is much less studied, and there is no clear decision pattern across studies; while some authors suggest that human and nonhuman animals prefer sooner over later painful shocks, others found the exact opposite. In a series of three experiments, Long–Evans rats chose between differently timed electric shocks and rewards in a T-maze. In experiment 1, rats chose between early and late painful shocks with identical, long reward delays; in experiment 2, they chose between early reward and early shocks, or late rewards and late shocks; in experiment 3, they chose between early and late rewards, with identical, short delays to the shock. We tested the predictions of two competing hypotheses: the aversive discounting theory assumes that future shocks are discounted, and, hence, less unpleasant than early shocks. The utility from anticipation theory implies that rats derive negative utility from waiting for the shock; late shocks should, hence, be more unpleasant than early shocks. We did not find unanimous evidence for either theory. Instead, our results are more consistent with the post hoc idea that shocks may have negative spill-over effects on reward values, the closer in time a shock is to a subsequent reward, the stronger the reward is devalued. Interestingly and consistent with our theory, we find that, depending on the temporal shock-reward contiguity, rats can be brought to prefer later over sooner rewards of identical magnitudes.

Key words: aversive discounting; delay discounting; intertemporal choice; reward; shock; utility from anticipation

Significance Statement
It is well understood how animals discount future rewards, but much less is known about how they choose between future aversive outcomes. We designed a novel task to examine decision-making between differently timed electric shocks and rewards. Although rats revealed clear preferences for sooner over later shocks, we found no evidence that they derived negative utility from waiting for the shocks (dread), nor that they discounted the disutility of future shocks. Instead, their choices were consistent with the novel hypothesis that shocks have negative, time-dependent spill-over effects on the utility of subsequent rewards. Consistent with this, we find that, depending on the temporal shock-reward contiguity, rats can be brought to prefer later over sooner rewards, thus promoting self-control.

Introduction
Individuals devalue rewards as a function of time until they can be realized, a phenomenon called temporal discounting (Kalenscher and Pennartz, 2008; Woolverton et al., 2012). Not only humans, but also nonhuman animals, including rats, mice, and pigeons, discount future rewards (Kalenscher et al., 2005; Vanderveldt et al., 2016). However, although temporal reward discounting is reasonably well understood, most real-life decisions often also yield negative consequences in the future, e.g., negative health effects of smoking, or diet-related health problems, a phenomenon less well studied. Several competing theories have
The aversive discounting model assumes a decrease of aversiveness as a function of delay. That is, a future event with negative utility would be less unpleasant than the same event now. This theory was empirically supported by several studies in rats who preferred later over earlier electric shocks (Deluty, 1978; Liley et al., 2019). Another study provided further support that the aversiveness of unpleasant events decreases with increasing delay (Woolverton et al., 2012). Interestingly, the decline of the negative value of the aversive event was best described by a hyperbolic function similar to temporal reward discounting.

However, contrary to the predictions of the aversive discounting model, many individuals tend to accelerate, rather than defer, aversive future events. For instance, many people tend to choose earlier over later painful dentist appointments, although the aversive discounting model would predict that the later appointment should be less frightening. The utility from anticipation model (Loewenstein, 1987) can account for this behavior. It states that the final utility of a delayed aversive event is the result of two interacting mental processes: the discounting of the aversiveness of a future aversive event, as hypothesized by the aversive discounting model, plus the disutility derived from anticipating the aversive event (dread). In other words, the final utility of a future outcome is a combination of the utility derived from anticipating the outcome and the discounted utility of future consumption. If the negative utility from anticipation outweighs the discounted negative utility from the event itself, people aspire to reduce the aversive anticipation period and thus accelerate the event (“get it over with”). Applied to the example above, the utility from anticipation model predicts that the dread of waiting for the painful dental procedure would motivate the acceleration of the appointment. Both the utility from anticipation model as well as the aversive discounting model make similar predictions regarding the discounting of future rewards.

In support of this model, human participants have been shown to not only accelerate electric shocks, they were even willing to endure a stronger shock to avoid waiting for it (Berns et al., 2006). In another study, human participants preferred a smaller, sooner over a larger, later monetary loss (Thaler and Shefrin, 1981; Holt et al., 2008). In general, humans discount delayed gains more steeply than delayed losses indicating different processes for the discounting of positive and negative outcomes (Estle et al., 2006; Mies et al., 2016). This effect is called the sign effect and participants exhibiting the sign effect show different neural activity compared with participants not experiencing the sign effect (Tanaka et al., 2014). Sign effects are observed for monetary losses, but also primary punishers like the threat of a shock (Robinson et al., 2015). Nonhuman animals also show behavior consistent with the utility from anticipation model. In one study, rats preferred an immediate electric shock over a delayed shock (Knapp et al., 1959). Additionally, Rodríguez et al. (2018) presented rats the choice between a large reward paired with a shock, and a small reward without shock. Without any delay between the large reward and the shock, the small reward was preferred. By increasing the delay between the large reward and the shock, rats chose the larger reward with shock. However, note that an alternative interpretation of the rats’ preference for early over late shocks is that the uncertainty of the timing of the shock increases with increasing delay. Animals may prefer sooner over later shocks to make more appropriate preparatory responses (Knapp et al., 1959; Seligman et al., 1971).

In summary, there is contradicting evidence in the human and nonhuman literature about decision-making about future aversive events: some evidence suggests that future aversive events are less unpleasant than immediate events, much like future rewards are less appetitive than immediate rewards. However, other studies imply that humans and nonhuman animals prefer earlier over later aversive events. In the human literature, this ambiguity has been resolved by assuming that primary punishment, such as a painful event, is more dreadful, and, hence, generates more disutility from anticipation, than secondary punishment, such as delayed financial payments (Loewenstein, 1987; Benzion et al., 1989). However, this idea cannot explain the contradicting evidence in the animal literature where primary reinforcement and punishment are used predominantly. Hence, it is unclear whether animals prefer an early aversive over a later event, as predicted by the utility from anticipation model, or vice versa, in line with the predictions of the aversive discounting model.

In the current study, we tested whether rats choose early or late electric shocks. Across repeated trials, they could enter one of two arms in a T-maze. Both arms yielded rewards and shocks, both delivered with variable delays. We tested the predictions of both theories in a series of three experiments by altering entry-to-shock delays and shock-to-reward delays. We additionally employed exploratory analyses.

Materials and Methods

Housing and animals

The rats for all experiments were obtained from Charles River Laboratories (Calco) and kept in an inverted 12/12 h light/dark cycle (light off at 7 A.M.). The temperature within the colony room was maintained at 20 ± 2°C and the humidity at 50%. Upon arrival the animals were between eight and nine weeks old, and the food access was set...
Experimental setup

A customized T-maze was used in all experiments (Fig. 1). The T-maze consisted of a start arm (80 × 30 × 45 cm), including a start box (40 × 30 × 45 cm). The start arm was connected to two identical decision arms (50 × 30 × 45 cm) leading left and right, respectively. The floor of the maze consisted of grid floors (9-mm gaps) to apply electric shocks. Each decision arm and the start box were separated by automatic sliding doors that could be lowered. Additionally, pellet dispensers were placed at the end of each decision arm and the start box. The pellet dispensers delivered the food rewards (20 mg, dustless precision pellets, Bio-Serv) into Petri dishes and reward lights were placed above the Petri dishes. The apparatus was controlled by Ethovision 11.5 (Noldus Information Technology).

Shaping and pretraining

The general procedure of shaping and pretraining of all studies was identical. Training and testing were done in the active phase of the animals and only on weekdays. The subjects were habituated to the maze for 1 day (see the Appendix for the details on the habituation and shaping procedures). Subsequently, rats were trained for four sessions in shaping 1. Next was five sessions of shaping 2, followed by 3 d of shaping 3. Independent of the performance of the animals all training phases were conducted. After the last step a performance criterion was applied to determine which animals were promoted to the main experiment consisting of 10 sessions. Food rewards always consisted of three sucrose pellets (20 mg, dustless precision pellets, Bio-Serv). The reward lights signaled the availability of food rewards.

Hypotheses

We tested rats’ preferences between timed rewards and shocks in three experiments. In experiment 1, the underlying logic was that rats choose between alternatives that yield identical rewards, delivered after identical delays, but differ with respect to the timing of the shock. Rats entered each decision arm in the T-maze and received a reward 21 s after arm-entry (Fig. 2). Entering one decision arm yielded a shock after 1 s (early shock + late reward; EL), entering the other arm yielded a shock after 20 s (late shock + late reward; LL). The aversive discounting theory predicts that rats would choose the arm yielding a late shock because the negative value of late shocks should be discounted; late shocks should, thus, be less aversive than early shocks at the time point of decision. By contrast, the utility from anticipation theory would predict choices of the early shock because of increasing dread with longer delays.

Note that this experimental design confounds entry-to-shock delay (the delay between entering the arm and receiving the shock) with shock-to-reward delay (late shocks are closer in time to reward than early shocks). A choice of late rewards could be indicative of utility from anticipation, as outlined above, but, given this confound, it is equally plausible that rats might use the shock as a cue to predict the following reward; hence, if this was true, the delay between cue and reward would be shorter in LL trials, temporal reward discounting would therefore predict choices of the late shock. Therefore, in experiment 2, we kept the shock-to-reward delay constant across choices. Choices of one arm yielded an early shock 1 s after entering, followed by an early reward 1 s after the shock (EE), choices of the other arm yielded a late shock, followed by a late reward (LL; same timings as above). The utility from anticipation theory would predict choices of the early shock/early reward arm to minimize delay-to-shock (dread) and delay-to-reward (sooner rewards are better than later rewards). The predictions of the aversive discounting model are somewhat unclear since the discounted disutility of the late shock (late shocks are better than early shocks) would compete with the discounted utility of the late reward (late rewards are worse than early rewards). Either way, any choice could not be accounted for by shock-to-reward signaling. To tease out the role of the discounted reward in experiment 2, we manipulated the entry-to-reward delay in experiment 3. Rats chose between identically timed, early shocks (1-s latency), and an early (EE, 2-s latency) or a late reward (EL; 21-s latency). Thus, standard temporal reward discounting would predict choices of the earlier over the later reward. Figure 3
summarizes the shock and reward contingencies and theory predictions. The following paragraph provides further details of the experimental designs.

**Experimental sessions**

We used a between-subject design, each experiment was performed by a separate batch of rats. In each experiment, rats performed 10 sessions (one session per day), consisting of 21 trials. Each session began with six forced trials in which only one arm was opened to ensure the rats were sampling both decision arms with their respective, entry-to-reward as well as shock-to-reward delays. Upon completion of the forced trials, the animals performed 16 free trials in which they could choose between both arms. Importantly, the arm-outcome (shock/reward) contingencies were counterbalanced and pseudo-randomized within and across animals and across sessions; they, hence, had to be re-learned in each session. If the arm-outcome contingency was unchanged for more than two consecutive sessions, it was reversed in the subsequent session. Before and after each animal training, the maze was cleaned with a 70% ethyl alcohol solution to remove odor cues.

**Figure 2.** Overview of the shock and reward timings for each experiment. After an animal entered a decision arm (0 s), the doors were closed. Afterwards, the rewards and shocks were delivered. In general, an early reward had a delay of 2 s and a late reward of 21 s. The early shock was administered after 1 s (relative to entering) and the late shock after 20 s.

**Figure 3.** Overview of the shock and reward contingencies in the two arms of the T-maze for each experiment and the predictions for each theory. For the shocks and rewards, each column represents one arm of the T-maze (early reward + late reward [EE]; late shock + late reward [LL]; early shock + early reward [EE]). In the first experiment, the entry-to-reward contingencies are identical in both arms, but the entry-to-shock delays differ between both arms. The second experiment tested constant shock-to-reward delays with different entry-to-shock delays. The third experiment used constant entry-to-shock delays, but the shock-to-reward contingencies differ between arms. In the predictions, “-” represents no predictions. Note that the assignment of shock/reward contingencies to the left or right arm of the T-maze will be pseudo-randomized across sessions for all experiments. The left side is always the condition for which the percentage of decisions was calculated.
entering the start box, the last door was closed and a new trial started. The animal was removed after completing all trials or after a duration of 40 min.

**Analysis**

In all three experiments, animals performed nearly all trials (experiment 1: 99.20%; experiment 2: 99.97%; experiment 3: 99.85%). Additionally, trials in which animals avoided the shock were labeled as omission trials. However, because of the low occurrence rate (experiment 1: 0.80%; experiment 2: 0.02%; experiment 3: 0.13%), we excluded them from analysis. To test whether rats prefer one arm or the other in each experiment, we calculated one-sample t tests (Extended Data 1; two-tailed; Table 1), one for each experiment, against the null hypothesis significance testing IBM SPSS Statistics 27 (IBM) and MATLAB 2019a (The MathWorks) was used. Each estimation of \( \theta \) followed a \( \beta \) function with \( \mu_\theta(\kappa - 2) + 1, (1 - \mu_\theta)(\kappa - 2) + 1 \). For \( \mu_\theta \), an uninformed \( \beta \) prior was used, and \( \kappa \) a \( \gamma \) function (0.01, 0.01) function for prior predictions, see Extended Data Figs. 4-1, 4-2, 4-3).

In addition, we conducted Bayesian inference statistics and calculated the Bayesian posterior distribution for each experiment (Extended Data 1). A Bayesian framework of inference allowed us to calculate the highest density interval (HDI). Thus, assumptions can be made that a specific value is within the 95% most probable data (Wagenmakers et al., 2018).

For the Bayesian parameter estimation RStudio (RStudio Team, 2018) was used. Additionally, the following R packages were used: rstan (Stan Development Team, 2020) and patchwork (Pedersen, 2020).

We calculated the following model (Fig. 4) for each experiment with the general form of \( \text{Decision}_{i,s} \sim \text{Bernoulli}(\theta_s) \) with \( \text{Decision}_{i,s} \in \{0,1\} \) and \( \theta_s \in \mathbb{R}_{0,1} \). \( \text{Decision}_{i,s} \) is the \( i \)th Decision for \( s \)th subject and \( \theta_s \) is the parameter estimate for the \( s \)th subject. We assume for \( \theta_s \sim \text{beta}(\mu_\theta(\kappa - 2) + 1, (1 - \mu_\theta)(\kappa - 2) + 1 \). for EE over LL. In experiment 3, both models predict a preference for EE over LL. In experiment 2, only the utility from anticipation model predicts a preference for EE over LL. In experiment 1, the aversive discounting model predicts a preference for LL over EL, and the utility from anticipation model assumes a preference for EL over LL. For prior predictions, see Extended Data Figs. 4-1, 4-2, 4-3).

| Data structure | Type of test       | Power          |
|---------------|--------------------|----------------|
| Experiment 1  | One-sample two-tailed | 95% confidence |
| Normal distribution   | t tests                  | interval        |
| Bernoulli     | Bayesian parameter estimate | 95% highest density Interval |
| Experiment 2  | One-sample two-tailed | 95% confidence |
| Normal distribution   | t tests                  | interval        |
| Bernoulli     | Bayesian parameter estimate | 95% highest density Interval |
| Experiment 3  | One-sample two-tailed | 95% confidence |
| Normal distribution   | t tests                  | interval        |
| Bernoulli     | Bayesian parameter estimate | 95% highest density Interval |

Listed are all performed tests in the same order as the text.
Results

One-sample t tests confirmed that, in the first experiment, rats preferred the EL (mean±SEM = 54.89±1.07) condition above chance level (two-tailed t test: t_{24} = 4.59, p < 0.001; Fig. 6A), indicating that rats preferred earlier over later shocks. Additionally, the μ was 0.55 with an 95% lower bound of 0.52 and an upper bound of 0.57 (Fig. 6B). On the other side is the prediction for the aversive discounting model (upper bound = 0.62, lower bound = 0.589). In the second row is experiment 2, and the prediction for the utility from anticipation model is displayed (upper bound = 0.42, lower bound = 0.38). Finally, in the bottom row, experiment 3 can be seen, and both models have the same predictions (upper bound = 0.42, lower bound = 0.38).

Figure 5. Mean parameter estimation with simulated data for all predictions according to the Bayesian hierarchy model. The predictions of both models are color coded, and on the top left corner are the specific experiments (EL: early shock + late reward; LL: late shock + late reward; EE: early shock + early reward). The y-axis shows the posterior distribution of μ. The vertical gray lines represent the upper and lower bound for the 95% highest density interval. Experiment 1 is in the top row, and the utility from anticipation model is on the left side (upper bound = 0.42, lower bound = 0.38). On the other side is the prediction for the aversive discounting model (upper bound = 0.62, lower bound = 0.589). In the second row is experiment 2, and the prediction for the utility from anticipation model is displayed (upper bound = 0.42, lower bound = 0.38). Finally, in the bottom row, experiment 3 can be seen, and both models have the same predictions (upper bound = 0.42, lower bound = 0.38).

Figure 6. A, Mean percentage of decisions for each experiment. In each experiment, animals performed 10 sessions and up to 21 trials per session (6 forced trials and 16 free trials). The timings of the reward and shocks are displayed on the x-axis as well as the experiment. Specifically, for experiment 1, the percentage of decisions are calculated for early shock + late reward (EL) versus late shock + late reward (LL). For experiments 2 and 3, the percentage is calculated for early shock + early reward (EE) versus late shock + late reward (LL) and early shock + early reward (EE) versus early shock + late reward (EL), respectively. The vertical lines represent the SEM, and each dot represents a single animal. For all experiments, we calculated one-sample t tests (two-tailed). In the first experiment, animals showed a significant preference above chance level. The second experiment failed to yield any significant results; and in the third experiment, animals revealed a significant preference below chance level. The black horizontal line represents chance level.

Research Article: New Research 6 of 10
November/December 2022, 9(6) ENEURO.0452-21.2022 eNeuro.org
preference for earlier shocks/earlier reward, or later shocks/late rewards. Bayesian analyses confirmed that rats were indeed indifferent between both alternatives. More specifically, \( \mu_b \) was 0.52 with an 95% lower bound of 0.49 and an upper bound of 0.55. Hence, the point of indifference is included in the HDI. Again, the convergence of the chains (nChain = 12, Chain length = 15,000, warmup = 5000) was successful with an effective chain length of 145,906 and a RHAT = 1. Finally, in experiment 3, the EE condition (mean \( \pm SEM = 45.12 \pm 1.13 \)) was chosen significantly below chance level (two-tailed t test: \( t_{23} = 4.32, p < 0.001 \)), suggesting the counterintuitive implication of a preference for the late reward over the early reward. Supporting, this implication, is the fact that \( \mu_b \) was 0.45 with a 95% lower bound of 0.42 and an upper bound of 0.48. The effective chain length was 122,730 (nChain = 12, Chain length = 15,000, warmup = 5000) with a RHAT of 1.

Thus, in summary, we found that rats chose earlier over later shocks if both rewards were equally timed. However, if entry to either shock led to an immediate shock, rats predominantly chose the arm with the late reward. This surprising finding implies that rats can be brought to choose later over sooner rewards by associating both choice alternatives with immediate shocks, thus reversing any time discounting of future reward value.

Finally, we computed additional statistics to further investigate the learning behavior (Fig. 7). We calculated the percentages of choice for the first block of trials (trials 1–8) and the second block of trials (trials 9–16). In experiment 1, the ANOVA revealed no significant block of trials \( \times \) session interaction on choice (\( F_{(9,207)} = 0.740, p = 0.672, \eta^2 = 0.031 \)). The main effect of block of trials (\( F_{(1,23)} = 1.462, p = 0.239, \eta^2 = 0.060 \)) and of session (\( F_{(9,207)} = 0.950, p = 0.483, \eta^2 = 0.040 \)) on choice were not significant as well. In experiment 2, there was a no significant interaction effect of block of trials \( \times \) session (\( F_{(9,171)} = 1.049, p = 0.403, \eta^2 = 0.052 \)). Additionally, no session effect was found (\( F_{(9,171)} = 1.530, p = 0.141, \eta^2 = 0.075 \)). However, there was a significant effect of block

---

**Figure 7.** Mean percentage choice (experiment 1: early shock, late reward vs late shock, late reward, EL vs LL; experiment 2: early shock, early reward vs late shock, late reward, EE vs LL; experiment 3: early shock, early reward vs early shock, late reward, EE vs EL) for the first and second blocks of trials (trials 1–8 vs trial 9–16; left panels) and for the first and second block of sessions (sessions 1–5 vs sessions 6–10; right panels) in all experiments. Repeated-measured ANOVAs revealed significant main effects of block of trials for experiments 2 and 3. Additionally, there was a significant main effect of block of sessions on percentage choice in experiment 3, but not in experiment 1 or 2. Additionally, there were significant main effects of session order within a block on choice in experiments 1 and 2 but not in 3; \( * p < 0.05 \).
of trials ($F_{(1,19) = 18.832, p < 0.001, \eta^2 = 0.498}$). Finally, in experiment 3, there was no significant block of trials x session interaction ($F_{(9,180) = 1.805, p = 0.070, \eta^2 = 0.083}$). As seen before, there was no significant effect for the sessions ($F_{(4,765,85,308) = 1.482, p = 0.205, \eta^2 = 0.069}$).

To check whether the behavior changed over time, we calculated the percentages of choice for the first block of sessions (sessions 1–5) and the second block of sessions (sessions 6–10). In experiment 1, there was no significant interaction of the block of sessions x session order ($F_{(4,96) = 0.176, p = 0.950, \eta^2 = 0.007}$). The block of sessions showed no significant effect ($F_{(1,24) = 0.017, p = 0.989, \eta^2 = 0.001}$). However, the main effect of session order within a block was significant ($F_{(4,96) = 2.541, p = 0.045, \eta^2 = 0.096}$). In experiment 2, the interaction of the block of sessions x session order was not significant ($F_{(4,76) = 0.106, p = 0.980, \eta^2 = 0.006}$) and there was no significant effect of the block of sessions ($F_{(1,19) = 0.227, p = 0.639, \eta^2 = 0.012}$). Furthermore, the main effect of session order within a block was significant ($F_{(4,76) = 3.883, p = 0.006, \eta^2 = 0.170}$). The last ANOVA for experiment 3 revealed no significant interaction of the block of sessions x session order ($F_{(4,60) = 0.509, p = 0.730, \eta^2 = 0.025}$). However, this time the block of sessions was significant ($F_{(1,20) = 7.415, p = 0.013, \eta^2 = 0.270}$). Finally, the session order within a block had no significant influence ($F_{(4,80) = 1.353, p = 0.258, \eta^2 = 0.063}$). Additionally, we compared the rats’ choice pattern in “switch” sessions, i.e., those sessions where the reward and shock contingencies changed from the previous session, with choices in “stay” sessions, i.e., sessions where the contingencies stayed the same between sessions. However, the results yielded no consistent and conclusive results across experiments. Notably, there was a higher number of switch than stay sessions across all experiments, given our randomization algorithm reported above. Hence, the inconclusiveness of this analysis might possibly because of insufficient statistical power.

**Discussion**

It is well established that appetitive events are discounted over time. Thus, a reward loses value as a function of delay. However, the literature is inconsistent about the discounting of aversive events. On the one hand, the utility of anticipation model predicts an acceleration of aversive events. The model assumes that future aversive events evoke dread. On the other hand, several studies showed a discounting of aversive events similar to temporal discounting of appetitive events. Thus, aversive events should be less displeasing in the future; delayed aversive events should, hence, be preferred over earlier ones. For our three experiments, both hypotheses yield specific predictions. In the first experiment, the LL condition would be preferred according to aversive discounting and the EL condition would be preferred according to the utility from anticipation model. Our data, indeed, showed that rats significantly preferred the EL option.

Rats attach positive or negative anticipatory value to cues associated with appetitive or aversive outcomes in many contexts and situations, for instance, during place conditioning (Huston et al., 2013) or secondary reinforcement (Berridge and Robinson, 1998). In other words, associating value with stimuli predicting rewards or punishment implies that rats attribute utility to cues associated with outcomes; they thus derive utility from anticipating the outcome. It is therefore not entirely unreasonable to assume that rats derived negative utility from waiting for the shock associated with the shock-arm, and hence, interpret these data as evidence for the utility from anticipation model. However, other theories could explain this choice pattern, too: animals can make more accurate predictions of the temporal occurrence of earlier compared with later events, including shocks (Church, 2003). Hence, it has been pointed out that the preference for earlier shocks might also stem from the rats’ ability to make more accurate, and, hence, better preparatory responses for sooner than later shocks (Seligman et al., 1971).

In experiment 2, rats showed no consistent preference for EE or LL alternatives. Therefore, experiment 2 also did not provide conclusive evidence for or against either theory.

Both theories make similar predictions about the discounting of future rewards. Hence, in experiment 3, both theories, along with standard models of temporal reward discounting (Kalenscher and Pennartz, 2008) would predict choices of EE over EL options. However, contrary to these predictions, our rats preferred later over earlier rewards. This surprising finding implies that rats can be brought to choose later over sooner rewards if both choice alternatives are associated with immediate shocks, thus reversing any time discounting of future reward value.

How can we explain the somewhat surprising preference for later over sooner rewards in experiment 3? One possibility is that the temporal proximity between shock and reward matters for the evaluation of the reward: the closer in time the reward is after the shock, the less valuable it becomes. Hence, according to this idea, the shocks’ negative spill-over effects on reward values would compete with the discounting of future rewards, potentially resulting in a higher valuation of later over sooner rewards.

This hypothesis provides an alternative account of our rats’ choices in experiments 1 and 2, too. In experiment 1, rewards are farther away in time from the shocks in the EL than the LL condition. Our hypothesis would, hence, predict EL preferences, consistent with our observations. In experiment 2, there is no difference in the reward-to-shock delay between the EE and the LL options. Our hypothesis would therefore predict indifference between both alternatives, again consistent with our observations.

Interestingly, this post hoc hypothesis might explain some contradictory results in the literature. Knapp et al. (1959) showed that rats preferred earlier over later shocks, but Deluty (1978) found the exact opposite choice pattern. The main difference in design between those studies was the timing of reward relative to the shock; in Knapp et al. (1959), rewards were delivered at the end of a trial, after the shocks with variable delays, similar to our experiment 1. Deluty (1978),
However, provided the rewards at the beginning of a trial. Thus, negative spill-over effects of shocks on the values of temporally close rewards, as hypothesized here, would result in a devaluation of the later rewards in the Knapp et al. (1959) design, but would lower the values of the sooner rewards in the Deluty (1978) design, hence explaining the differential choice patterns in both studies.

Another study (Rodriguez et al., 2018) combined a small reward with no shock and a large reward with a shock. At first, the animals preferred the small reward avoiding a shock altogether. With an increased delay between the large reward and the shock, a preference reversal occurred shifting the preference to the large reward with a shock. According to the authors, this effect can be explained with aversive discounting. However, following our hypothesis, it is also possible that the disutility of the shock lowered the value of the reward, but by increasing the delay between shock and large reward, the negative transfer effect of the shock on reward value was gradually reduced. Additionally, it was shown that the reinforcer effectiveness of aversive histamine injections decreased as a function of the delay between histamine and cocaine administrations (Woolverton et al., 2012). Interestingly, the decreased effectiveness of the aversive event was well described by a hyperboloid discounting function. Therefore, the authors argued that the results are in line with aversive discounting. However, those results are also in line with our hypothesis.

Interestingly, the data also indicate learning behavior within sessions, i.e., within each session the animals increase their preference for the chosen option. However, this pattern is not seen in experiment 1 which might be due to the pseudo-randomization. Hence, the learning behavior could reflect the relearning of the side contingencies. Interestingly, we found a significant effect of session order within a block for experiment 1 and experiment 2. It is possible, that this again reflects learning behavior based on experience. However, this result has to be interpreted with caution because there was no general learning effect over all sessions. In experiment 3, the results indicate stronger preferences for the second half of the experiment. Hence, over time the animals strengthened their preference, i.e., in the first five sessions the preferences are lower compared with the last five sessions. In general, it seems like there is a complex pattern of learning within sessions and possibly over sessions as well. However, because of the pseudo-randomization and the complex pattern of significant results, those results have to be interpreted with caution.

Animal experiments, such as ours, use delays in the range of several seconds. However, aversive outcomes in human intertemporal decisions are often in the range of months or years, such as negative health consequences of smoking or unhealthy life styles. It is unknown whether our results translate to longer timescales, but a recent study has shown that intertemporal decisions are comparable across different time scales (Lukinova et al., 2019). It would be intriguing to test in follow-up studies if human participants can be brought to choose delayed over early rewards when associated with differently timed punishments, albeit with much longer delays.

Finally, some authors found diverging results from ours; in analogy to our experiment 2, Renner and Specht (1967) offered rats the decision between an early shock followed by an early reward and a late shock with a late reward, among others. However, unlike in our experiment 2, their rats preferred the early shock and early reward condition. There are several procedural differences between the design by Renner and Specht (1967) and ours that may explain the differences in choice patterns. Most notably, the delay preceding the late shock, late reward in their study was much longer than our delay in experiment 2. It is possible that, with such long delays, standard reward and shock discounting mechanisms might have dominated the decision process, resulting in a devaluation of the future rewards and shocks, thus generating clear preferences. This interpretation is in line with our hypothesis since the shock-to-reward latency was kept constant across choice alternatives, and, therefore, should not matter for the choice process. In other words, it is possible that we would have found a similar choice pattern had we used longer delays, too.

In general, it is worth noting that the reasons for the discrepancy between our results and those in the literature may also be attributable to the fact that animals (Renner and Specht, 1967), much like humans (Thaler and Shefrin, 1981), discount future rewards more steeply than future aversive outcomes. Thus, depending on the magnitude and delay of the aversive and appetitive events, a rat may be biased toward deferring or accelerating the outcomes. Future studies should directly decide between the predictions of those theories.

In conclusion, we did not find unanimous support for either theory outlined in the introduction. By contrast, it seems as if decision-making between timed rewards and punishments involves at least three different mechanisms. The first two mechanisms are temporal reward and aversive discounting. As a third mechanism, we propose that shocks have a negative spill-over effect on the valuation of rewards that are close in time, with decreasing spill-over effects with increasing shock-to-reward latencies. Our theory makes the interesting and counter-intuitive prediction, supported by data from experiment 3, that rats can be brought to prefer later over sooner rewards of identical magnitudes if the later rewards are temporally decoupled from shocks. One intriguing question is if this finding could be used to nudge human participants to make more far-sighted intertemporal decisions, an implication that would have to be tested in future studies.

Appendix

1. Habituation

Each animal was placed into the start box, while all doors were closed and rats received their first reward. Subsequently, all doors were opened and the animal was able to freely explore the maze. Food rewards, indicated by the reward light, were delivered after each zone transition, i.e., changing from either decision arm or the start box. After 10 min, all doors closed and the animal was removed from the maze. All animals were automatically promoted to the shaping sessions.
Shaping
Shaping steps ensured that the rats could learn the functional principles and procedures of the task. The reward was evenly distributed between both decision arms. There was no performance criterion for any of those steps but the last.

Step 1
The first shaping step consisted of four sessions of 16 free trials or the maximum trial duration of 30 min. In free trials, both sliding doors were opened so that rat could choose to enter either decision arm. At the beginning of the session, the animal was placed in the start box followed by food rewards. Afterwards, all doors opened and as soon as the animal entered a decision arm, all doors were closed. Fifteen seconds after entering a decision arm, food rewards were delivered. Another 10 s elapsed before the door of the chosen decision arm and the door of the start box opened. As soon as the animal entered the start arm, the door of the decision arm was closed. After entering the start box, food rewards were delivered, indicating the end of the trial.

Step 2
The second shaping step consisted of five sessions of six forced trials and 16 free trials. The timings and structure were identical to step 1. Each session began with six forced choice trials (3 on each side in a pseudorandom order), in which the rat was directed in one of the two decision arms by just opening one of the two sliding doors. In the following 16 free choice trials, both sliding doors were opened. A session either ended after completing all trials or after 40 min.

Step 3
Finally, the last training step was conducted, consisting of three sessions of six forced and 16 free trials. The general procedure was identical to step 2. However, 13 s after entering a decision arm, a mild electric shock (800 ms, 0.3 mA) was delivered. The shock intensity was chosen such that animals would avoid it, but it would not induce moderate or strong freezing, or any other signs of fear conditioning. The delay between shock onset and reward delivery was 1 s. The animal was removed after completion of all trials or after 40 min. Criterion for promotion to the main experiment was that the animal was removed after completion of all trials or after 40 min. Delay between shock onset and reward delivery was 1 s. The strong freezing, or any other signs of fear conditioning. The animal was removed after completion of all trials or after 40 min. Criterion for promotion to the main experiment was that the animal was removed after completion of all trials or after 40 min.

References
Benzion U, Rapoport A, Yagil J (1989) Discount rates inferred from decisions: an experimental study. Manag Sci 35:270–284.
Berns GS, Chappelow J, Cekic M, Zink CF, Pagnoni G, Martin-Skurski ME (2006) Neurobiological substrates of dread. Science 312:754–758.
Berridge KC, Robinson TE (1998) What is the role of dopamine in reward: hedonic impact, reward learning, or incentive salience? Brain Res Brain Res Rev 28:309–369.
Church RM (2003) A concise introduction to scalar timing theory. In: Functional and neural mechanisms of interval timing, pp 3–22. Boca Raton: CRC/Routledge/Taylor and Francis Group.
Deluty MZ (1978) Self-control and impulsiveness involving aversive stimuli. J Exp Psychol Anim Behav Proc 4:250–266.
Estle SJ, Green L, Myerson J, Holt DD (2006) Differential effects of amount on temporal and probability discounting of gains and losses. Mem Cognit 34:914–928.
Holt DD, Green L, Myerson J, Estle SJ (2008) Preference reversals with losses. Psychon Bull Rev 15:89–95.
Huston JP, de Silva MAS, Topic B, Müller CP (2013) What’s conditioned in conditioned place preference? Trends Pharmacol Sci 34:162–166.
Kalenscher T, Pennartz CM (2008) Is a bird in the hand worth two in the future? The neuroeconomics of intertemporal decision-making. Prog Neurobiol 84:284–315.
Kalenscher T, Windmann S, Diekamp B, Rose J, Güntürkün O, Colombo M (2005) Single units in the pigeon brain integrate reward amount and time-to-reward in an impulsive choice task. Curr Biol 15:594–602.
Knapp RK, Kause RH, Perkins CC Jr (1959) Immediate vs. delayed shock in T-maze performance. J Exp Psychol 58:357–362.
Liley AE, Gabriel DBK, Sable HJ, Simon NW (2019) Sex differences and effects of predictive cues on delayed punishment discounting. eNeuro 6:ENEURO.0225-19.2019.
Loewenstein G (1987) Anticipation and the valuation of delayed consumption. Econ J 97:666–684.
Lukinova E, Wang Y, Lehrer SF, Eriich JC (2019) Time preferences are reliable across time- horizons and verbal versus experiential tasks. Elife 8:e39656.
Mies GW, De Water E, Scheres A (2016) Planning to make economic decisions in the future, but choosing impulsively now: are preference reversals related to symptoms of ADHD and depression? Int J Methods Psychiatr Res 25:178–189.
Pedersen TL (2020) patchwork: the composer of plots. RStudio Team (2018) RStudio: Integrated Development Environment for R. RStudio, PBC, Boston, MA. Available at http://www.rstudio.com/.
Renner KE, Specht L (1967) The relative desirability or aversiveness of immediate and delayed food and shock. J Exp Psychol 75:568–570.
Robinson OJ, Bond RL, Roiser JP (2015) The impact of threat of shock on the framing effect and temporal discounting: executive functions unperturbed by acute stress? Front Psychol 6:1315.
Rodriguez W, Bouzas A, Orduña V (2018) Temporal discounting of aversive consequences in rats. Learn Behav 46:38–48.
Seligman MEP, Maier SF, Solomon RL (1971) Unpredictable and uncontrollable aversive events. In: Aversive conditioning and learning (Brush FR, ed), Chapter 6, pp 347–400. San Diego: Academic Press.
Stan Development Team (2020) RStan: the R interface to Stan. Available at https://mcСтalz.com/.
Thaler RH, Shefrin HM (1981) An economic theory of self-control. J Polit Econ 89:392–406.
Vanderveldt A, Oliveira L, Green L (2016) Delay discounting: pigeon, rat, human—does it matter? J Exp Psychol Anim Learn Cogn 42:141–162.
Wagenmakers E-J, Marsman M, Jamil T, Ly A, Verhagen J, Love J, Selker R, Gronau QF, Smirra M, Epskamp S, Matzke D, Rouder JN, Morey RD (2018) Bayesian inference for psychology. Part I: theoretical advantages and practical ramifications. Psychon Bull Rev 25:35–57.
Woolverton WL, Freeman KB, Myerson J, Green L (2012) Suppression of cocaine self-administration in monkeys: effects of delayed punishment. Psychopharmacology (Berl) 220:509–517.