

Article

Not peer-reviewed version

State-Dependent Valuation Learning Affects Preference in a Delay-Discounting Task

Zirahuén Vílchez , Erick Barrón , Laurent Ávila-Chauvet , Jonathan Buriticá , [Óscar García-Leal](#) *

Posted Date: 15 March 2023

doi: [10.20944/preprints202303.0291.v1](https://doi.org/10.20944/preprints202303.0291.v1)

Keywords: state-dependent valuation learning; self-control; delay-discounting; impulsivity



Preprints.org is a free multidiscipline platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This is an open access article distributed under the Creative Commons Attribution License which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Disclaimer/Publisher's Note: The statements, opinions, and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions, or products referred to in the content.

Article

State-Dependent Valuation Learning Affects Preference in a delay-Discounting Task

Zirahuén Vílchez ¹, Erick Barrón ², Laurent Ávila-Chauvet ³, Jonathan Buriticá ⁴, Óscar García-Leal ^{5,*}

¹ Centro de Estudios e Investigaciones en Comportamiento, Universidad de Guadalajara, México; zira_vilchez@hotmail.com

² Facultad de Psicología, Universidad de Guadalajara, México; erick@academicos.udg.mx

³ Instituto Tecnológico de Sonora, México; laurent.avila231516@potros.itson.edu.mx

⁴ Centro de Estudios e Investigaciones en Comportamiento, Universidad de Guadalajara, México; jburitica@academicos.udg.mx

⁵ School of Doctoral Studies and Research, Universidad Europea de Madrid, España; Centro de Estudios e Investigaciones en Comportamiento, Universidad de Guadalajara, México; oscar.garcia.leal@universidadeuropea.es

* Correspondence: oscar.garcia.leal@universidadeuropea.es

Simple Summary: When an animal makes a choice, the value attributed to the alternatives that constitute the context will determine its preference and final choice. The value of an alternative will depend on its consequence magnitude (how much the animal gets), the time estimated to reach the consequence (how much time it requires) and, at least, the internal state of the animal at the moment of learning about the properties of each alternative. In this experiment, we explore how the state of food deprivation at the moment an animal faces for the first time one alternative, and therefore learn about its properties, will affect its future preference for this alternative. We manipulated the state of food deprivation in two groups of rats: group one faced the alternative under a state of high food deprivation while group two faced the alternative under a state of relatively sated. Later, we presented a choice situation between the previously trained alternative and a new one. Consistently, animals preferred the previously trained alternative. Later, we diminish the value of the previously trained alternative, by increasing progressively the time required to get the consequence. Preference for the previously trained alternative extended on time, independently of diminishing its value.

Abstract: If a subject learns about the properties of an alternative under a high state of food deprivation, it will prefer this alternative over a similar one in a later choice test, although the subject is currently under a relative state of satiety. In this experiment, we probe this effect in a delay-discounting task, where the reward of the experienced alternative is progressively delayed. Eight male Wistar rats were trained to respond to an alternative under a high state of deprivation and eight rats under a relative state of satiety. After 20 sessions, their weight was recalibrated to a state of relative satiety. Then, they were exposed to a delay-discounting task where the long-larger alternative was the one used during training. The group that experienced the alternative under a higher level of food deprivation preferred more this alternative, and its preference stayed across time. On the contrary, the group that experienced the alternative under a lower level of food deprivation systematically preferred the small-shorter alternative. Our results support that the reinforcement value may be sensitive to the organism's state at the moment an alternative is experienced for the first time.

Keywords: state-dependent valuation learning; self-control; delay-discounting; impulsivity

1. Introduction

[1] (p. 2) defines self-control as the ability or capacity to obtain an objectively more valuable outcome rather than an objectively less valuable outcome through tolerating a longer delay or a

greater effort requirement in obtaining the more valuable outcome. Thus, a strong preference for larger-longer (LL) over smaller-shorter (SS) rewards is the behavioral definition of self-control in most choice paradigms [2,3]. On the other hand, impulsivity has been usually defined as the other extreme of a behavioral continuum [4]. Self-control has been also defined as a limited capacity resource that can be improved with practice [5].

The particular function that represents how the value of an alternative diminishes as the delay of gratification increases, has been hypothesized by several models, assuming that whatever the alternative is chosen, its value will decrease as the delay to reward increases [6]. On one hand, economic models assume that delayed rewards are exponentially discounted, diminishing its value as time advances in such a way that the reward value is discounted by a fixed proportion for each unit of time [7,8]. On the other hand, behavioral research supports the reward value does not decrease by a fixed proportion, but proportionally to its delay [9]. Inside the behavioral tradition, a hyperbolic discounting function [10] is arguably the most accepted form the reward value decays as time advances.

The rate at which a reward loses its value as the delay to it increases will determine how self-controlled is the subject's behavior. Hence, subjects that discount a reward faster are assumed to show more impulsive behavior than subjects that discount slower, who will show more self-controlled behavior.

Previous research suggests that reward value is not only dependent on properties like amount and delay to reward, but other variables can contribute to its value. One of these variables, that has received attention in the past, is the context in which a reward is presented and particularly, how stimuli that precede the reward presentation can change or modulate preference [11].

One of the first experiment that probes that events preceding the presentation of a stimulus could bias preference, was made by [12]. Pigeons were first exposed in a Condition 1 to a FR1 schedule on a center key that gave access to a simple discrimination task in which two different color side-keys were available. Choices on the red key were reinforced with food, while choices on the yellow key were not. A second condition was equal, but twenty responses were now required on the center key, to give access to a similar discrimination task between a reinforced green key and a non-reinforced blue key. Tests consisted of a choice task in which pigeons chose between the two reinforced side-keys (red vs. green) and the two non-reinforced side-keys (yellow vs. blue). Results indicated a preference for the two alternatives previously followed by twenty responses on the center key. Thus, pigeons preferred the alternatives that were preceded by a higher effort. [13] proposed the Within-Trial Contrast Model (WTCM) to explain this result. WTCM assumes that preference for the side-keys faced after a higher effort is controlled by a higher positive shift in the organism's hedonic state caused by the reward when the pigeons previously must complete twenty responses in the center key than only one response.

Other evidence seems better to supports State-Dependent Valuation Learning (SDVL) as an alternative model [14]. This model hypothesizes that the relative value of a reinforcer is directly related to the energetic state or the organism's fitness at the point of reinforcer delivery [11] (p. 186). SDVL assumes that if an option is trained while a subject is under a high state of food deprivation, such option will be preferred over a similar one during a later choice test [15,16]. This prediction is supported under the assumption of a decelerated function of value related to objective payoff where reinforcement value increases with the amount of deprivation. The effect of SDVL is hence described as the preference for stimuli associated with the presentation of the reward under a high state of food deprivation compared to situations in which these stimuli are presented under a state of relative satiety [17]. Hence, learning about the reward value is important in SDVL.

Supporting SDVL, [18] suggested that the remembered gain might depend on the amount of wellness associated with that alternative in the past. Therefore, food obtained on an alternative while the animal was very hungry has a higher reinforcement value than food obtained on an alternative while the animal was not so hungry. Extending this argument, [19] argued that when a subject becomes familiar with a source of reward it acquires two forms of knowledge about it: about the physical properties of the reward (e.g., amount and delay) and about the benefit accumulated from

the outcome under specific circumstances, like the reduction of the state of deprivation. If this knowledge about the properties of a source of reward is determinant of its reinforcement value, then preference for that option should be changed as well. The implication of this conjecture is that choice behavior might be controlled more by the food's value to the animal than by the amount of food per se.

More recently, [20] trained pigeons (*Columba livia*) to respond to two keys in separate sessions, one key while pigeons were at 80% of their free-feeding body weight and another key when pigeons were fed prior to the training sessions. Afterwards, both keys were simultaneously presented, and the pigeons chose between them. Half of the subjects were at 80% of their free-feeding body weight and the other half being pre-fed. The pigeons preferred the alternative associated with a higher level of food deprivation during training, no matter whether they were currently deprived or satiated. Nevertheless, the effect seemed to disappear in pigeons when the response-initiated fixed interval between the options during training was matched. This suggests that the relative immediacy between the options could be a factor controlling the effect [17].

Another demonstration of SDVL was reported in fishes (*Banded tetras*) by [19]. Fishes were trained to access to two arms of a Y-maze in separate sessions, one while they were under a closed economy schedule and the other when they were fed prior to the training session. Then they were tested on a choice task, half of them being pre-fed and the other half being in closed economy, and then reversely. The result was a preference for the option initially trained during the hunger state. The effect has also been reported in birds [15,21,18].

Previously, [22] reported similar results but using water instead of food deprivation. A group of rats with different histories of water deprivation were presented a choice situation between two alternatives, one delivering a small amount of water after a short delay and the other delivering a larger amount of water but after a longer delay. Independently, whether they had a previous experience on water deprivation, or they were currently deprived, they showed a higher preference for the LL alternative than rats without previous experience on water deprivation. [22] suggested that such strategy is a result of a long-term strategy learned by the subjects to match the scarcity of the experienced environment.

While SDVL has been reported in several species, even belonging to different classes or even phyla, suggesting it is essentially a primitive adaptive mechanism, as far as we know the protocols used to study its effect have always delivered rewards immediately. It has not been tested if SDVL also bias preference in a delay discounting task, where one alternative delivers the reward delayed in time compared to the other alternative.

In delay-discounting procedures, subjects are exposed to consecutive choice situations in which they must choose one of two alternatives. One of them, the smaller-shorter (SS) alternative, delivers a small amount of reinforcement after a short-constant delay, while the other, the larger-longer (LL) alternative, delivers a larger amount of reinforcement, after a delay that changes, increasing or decreasing, across the experiment. When the amount and delay of each alternative make indifferent the choice, the alternatives are considered equivalent in its reinforcement value.

In our experiment, we study SDVL on a delay-discounting task, to know if the reinforcement value of an alternative, that results of learning about its properties in different state of water deprivation, extends in time even while the alternative is later devaluated through increasing the delay to the reinforcer. Water deprivation in rats is highly related with self-imposed food restriction [23]. With this purpose, two groups of rats learned about one alternative under a state of high and low deprivation. Afterward, we tested groups' preference for this alternative in a delay-discounting choice situation where its consequence was progressively delayed. Our hypothesis is that higher deprived group will persist on the previously trained alternative even though increasing its reward delay. As far as we know, this effect has not been reported in rats [11].

2. Materials and Methods

Subjects. Sixteen males naïve Wistar rats (*Rattus norvegicus*), three months old when the experiment started. They were housed individually and exposed to a 12h:12h light/dark cycle with lights on a 7am. Food was always available throughout the experiment, but the access to water was limited by the experimental phase in which the subject was.

The experimental procedure was approved by the local Ethical Committee of the Center for Studies and Investigations in Behavior, by the University of Guadalajara committee for animal experiments, and met governmental guidelines.

Materials. Four operant conditioning chambers (MED Associates, Inc., Model ENV-007) for rats were used. Each chamber was 29.53 cm long, 23.5 cm width and 27.31 cm high. Each box was equipped with a fan that circulated air. The front panel was divided into three columns. An arm type water dispenser (ENV-202M) was inserted into the central column, located 6 cm from the stainless-steel floor. A 2.8 w house light was located above the water dispenser, 17 cm from the stainless-steel floor. A white noise device was located at the same height of the house light but at the right column to mask extraneous noises. It stayed turned on across the experiment and through each session.

At the back panel, three retractable levers (ENV-112CM) were located 7 cm from the stainless-steel floor, one inserted in each column. A pressure of 0.2 N on the levers was required to count as a response. Above the left and right levers, 12 cm from the stainless-steel floor and 18cm apart center-to-center, were located two white hues (ENV-212M). Finally, a 4500 Hz tone generator was in the central column, 5 cm above the central lever.

Procedure. Rats were pseudo-randomly divided into two groups (n=8) in order to make groups similar in weight before starting to deprive one of them. One of the groups had daily access to 25ml of water. It was labeled as 'Relative Satiety Group' (RS), because the amount of water available made possible to maintain rats at approximately 90% of its free-feeding weight across the experiment. The other group was exposed to a more restricted water access. Rats received only 10ml of water daily, to keep them at approximately 70% of its free-feeding weight. It was labeled as 'Deprived Group' (D). Both groups received vitamin B12 Complenay® liquid supplement across the experiment. The supplement was daily diluted in water and water was delivered at the end of each session, after the rats were moved to their housing boxes.

The experiment was divided into two different phases: training and test. The training phase, started two weeks after the groups' weight reached steady. The delay-discounting test phase started two weeks after finishing the training phase.

Training phase. D and RS rats were trained to press a lever under a continuous schedule of reinforcement (CRF) on the right or left lever. The lever position was counterbalanced among subjects. Such lever was called the Trained Alternative (TA), and it was associated with a fixed-white cue located above the lever. After pressing the lever, it was immediately delivered 0.12 ml of water that served as a reward. Each session ended after 100 rewards were earned. Within 30 minutes after the end of each daily session both groups received the water necessary to maintain their weight depending on the group the rat was assigned. Training phase lasted for 20 sessions.

Once training phase was finished, D group received additional water in the home cages to recalibrate their weight to approximately 90%. After two weeks, the delay-discounting test phase started. Therefore, D and RS groups started the delay-discounting test phase at approximately 90% of their free-feeding body weight.

Delay-discounting test phase. Our experiment followed the procedure previously used in [24], which is an adaptation from [25] where delays were gradually increased in consecutive blocks of trials. This procedure makes easy to measure sensitivity to delays at shorter time intervals and minimizes the possibility of a co-existing effect of habituation.

Each session was divided into 5-blocks, 8-trials each. Inside each block, the first two trials were no-choice and the remaining six were choice trials. Two alternatives were presented: Non-Trained Alternative (NA) and Trained Alternative (TA), where TA was the alternative previously trained during the Training phase under different levels of water deprivation. In no-choice trials, only one of these alternatives was inserted into the chamber to be pressed. In choice trials, both alternatives were

inserted into the chamber, and one of them must be chosen to move to the next trial. A session finished after 40 trials were completed. Each block was presented once per session.

Each trial started with the insertion of the central lever and a 65 db tone. Pressing this lever resulted in its retraction, the tone disappearance, and the insertion of one or two side levers, depending on the trial type (i.e., no-choice or choice). The TA was always inserted at the same side it was previously trained during the Training phase, and a fixed-white cue was activated above it, as previously done during training. Therefore, the NA was always inserted at the other side, and it was correlated with an intermittent-white cue.

No-choice trials served the rats to learn the properties of the alternatives block by block. NA always delivered a constant amount of reward after a constant delay. TA delivered a constant amount of reward, but the delay changed between and across blocks as will be described. Because we presented two no-choice trials by block, one of them permitted to press only the NA and in the other the TA was the only alternative available. Trials order was randomized.

Choice trials served to measure changes in preference for TA, as its reward delay changes between and across blocks. In choice trials, a response on the NA resulted in the extinguishing of both the intermittent and fixed cues, the retraction of both levers, the onset of the general house light and an immediate delivery of 0.06ml of water. A response on the TA resulted in 0.12 ml of water delivered with a delay that increased among blocks, considering three different conditions. In Condition 1, the scheduled delays among blocks were 0, 0.5, 1, 2 and 4 s, in Condition 2 were 0, 1, 2, 4 and 8 s, and in Condition 3 were 0, 2, 4, 8 and 12 s. An inter-block-interval of 15s was programmed, while the inter-trial-interval (i.e., ITI) was set to 5s.

RS group experienced the delays of each condition in ascending order, while D group in descending order. The order of conditions was according to its ordinal number and did not differ across groups. Subjects stayed for 10 sessions in each condition after moving to the next condition, for a total number of 30 sessions in the Delay-discounting test phase.

3. Results

Experiencing an alternative under different levels of food deprivation could influence on its reinforcement value. If such is the case, then we should observe different performance in subjects who learned about the alternative under different levels of food deprivation. In this experiment, we induced food deprivation through restricting the daily water available. Particularly, we would expect a faster approach to the alternative that gained higher reinforcement value, the experienced under higher deprivation.

Figure 1a plots response latency in Training phase to the only alternative available, under different state of food deprivation. Response latency was defined as the time in seconds between the extension of the lever in the chamber and the lever press to earn a reward.

Because we were interested in latencies as a measure of the alternative's reinforcement value, we only considered the last 5 sessions. Besides, due to data loss, we could only plot latencies for half of the subjects assigned to group D. Despite this, as Figure 1a shows, data are not so variable, so we assume they represent well the final reinforcement value of the trained alternative for each group.

The average response latency was shorter for the D group (Average \pm SEM; 1.58 ± 0.32) than for the RS group (3.57 ± 0.83 s). A Mann-Whitney non-parametric comparison results in significant differences ($U=124.5$, $p=.04$), supporting that both groups performed differently in training phase because of water-deprivation level. The session length was also significantly shorter ($U=590.5$, $p=.044$) for the D group (993.3 ± 42.13 s) than for the RS group (880.4 ± 16.25 s). Figure 1b plots session length for each group and for all rats. Together, response latencies and session length suggest the groups performed differently depending on the level of water-deprivation they were exposed. Considering these results, we argue TA gained different reinforcement value for each group: it was more valuable for the D group than for the RS group.

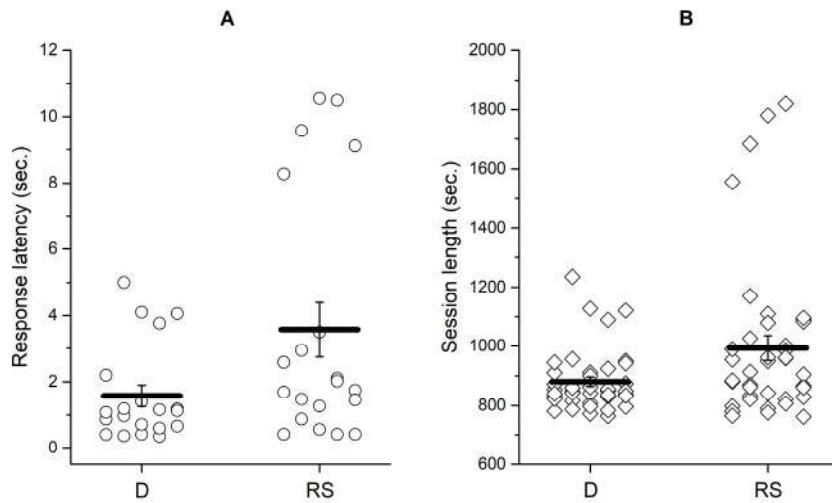


Figure 1. (a) Average response latency for TA in the last 5 sessions of each subject in the training phase. Each dot represents the average latency in a particular session. Only half of the subjects were considered by group. (b) Average session length of the last 5 sessions considering all subjects of both groups. The horizontal line represents the average of the group. D = Deprivation group. RS = Relative Satiety group.

Preference in the Delay-discounting test phase was defined as the relative choice for TA in choice trials by condition and block, considering only the last 5 sessions. Thus, TA preference is the result of dividing the number of times TA was chosen considering choice-trials, between the total number of choice-trials by block. Figure 2 plots TA preference for D and RS groups.

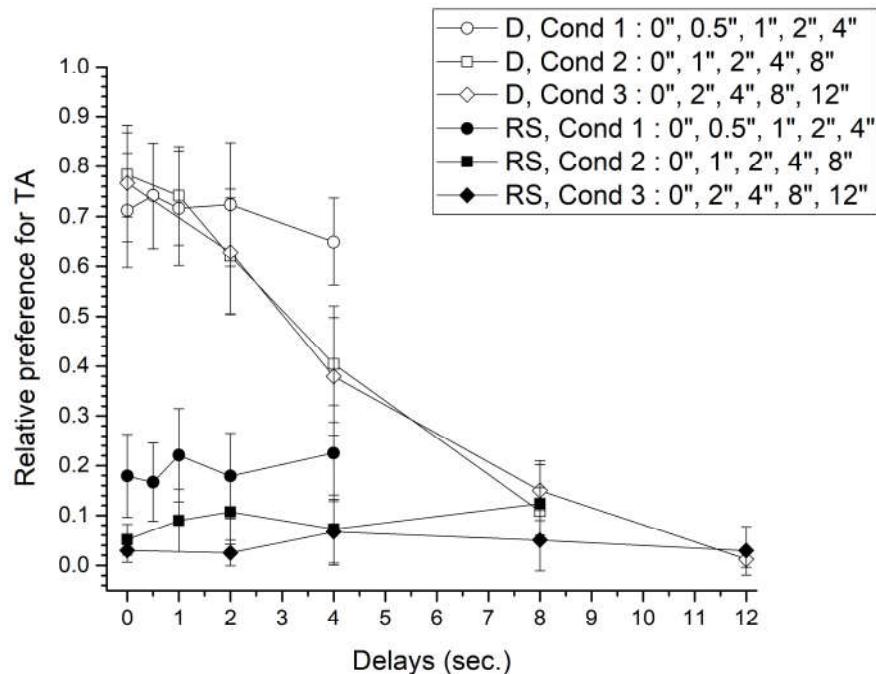


Figure 2. Average preference (\pm SEM) for TA considering the last 5 sessions by block during the test phase. Filled shapes correspond to D group, while white shapes correspond to the RS group. Condition 1: circles. Condition 2: squares. Condition 3: diamonds.

RS group preference for TA did not vary neither between conditions nor across blocks. Importantly, rats preferred NA alternative no matter the programmed delay to reward in TA, despite NA provided half amount of water. Nevertheless, average preference for NA diminishes across conditions. Average delay to reward was 1.5 s, 3 s, and 5.2 s for conditions 1, 2, and 3, respectively. Average preference was 0.194 (SEM \pm 0.012), 0.089 (\pm 0.013), and 0.04 (\pm 0.008) for conditions 1, 2, and 3, respectively. Non-parametric mean comparison revealed significant differences between conditions ($\chi^2(2)=10$; $p<.01$).

On the other hand, D group showed a sharply preference for TA, particularly with the smaller delays to reward. TA preference diminished as the delay to reward increased. Preference for TA reversed above 4 s delay for conditions 2 and 3, such that rats changed to choose the more immediate alternative. Average preference for TA was higher in condition 1 (0.709 \pm 0.016), than in condition 2 (0.532 \pm 0.125) or condition 3 (0.388 \pm 0.141). Nevertheless, differences were non-significant ($\chi^2(2)=5.158$; $p=.076$).

Finally, a mixed GLM with group and delay to reward as factors, revealed significant differences between groups in Condition 1 ($F(1,14)=17.37$, $p=.001$, $\eta^2=.55$), in Condition 2 ($F(1,14)=24.98$, $p<.001$, $\eta^2=.64$), and Condition 3 ($F(1,14)=24.20$, $p<.001$, $\eta^2=.63$). Also, there were an interaction effect between groups and blocks in Condition 1 ($F(4,56)=2.79$, $p=.03$, $\eta^2=.16$), Condition 2 ($F(4,56)=38.29$, $p<.001$, $\eta^2=.73$), and Condition 3 ($F(4,56)=27.94$, $p<.001$, $\eta^2=.66$).

4. Discussion

Our goal was to test the effect of SDVL on a delay-discounting task, particularly in the procedure described in [24]. During the training phase, RS and D groups were exposed to an alternative that delivered 0.12ml of water under different levels of water deprivation. This alternative is referred as TA. A previous study supports that water deprivation is highly related with food self-deprivation in rats [23]. Hence, this experiment assumes that water restriction changes the organism state at the moment of choosing, in such a way that as water deprivation increases, the body-weight decreases. RS group learned about TA at 90% of their free-body weight, but D group rats were at 70% of their free-body weight. Both groups performed differently during the Training phase. Particularly, lower latencies of response were recorded on D group, as well as lower session length.

Latency of response has been proposed as a measure of the reinforcement value. In a series of papers Kacelnik et al. [26,27] explored the relation between response latencies to several alternatives that were faced in sequential encounters, and later preference on simultaneous encounters to these alternatives, in starlings. Usually, the alternative that evoked the lower response latency in sequential encounters becomes later preferred in simultaneous encounters. This effect has been also reported in human participants [28].

After finishing Training phase, group D received additional water to recalibrate their weight to approximately 90%, such that rats from both groups started the delay-discounting test phase with an equivalent weight. Preference for TA was defined as the result of dividing the number of times TA was chosen in choice-trials by the total number of choice-trials in each block.

While RS group always preferred NA, independently of the programmed delay to reward, D group preference for TA resulted evident, especially for shorter delays. Indeed, RS group seemed to be non-sensitive to the reward delay, in such a way that its clear preference for NA was not affected by the reward delay. Experiencing the TA under a state of low food deprivation seemed to contribute to decreasing its reward value in such a way that rats avoid the experienced alternative even though it delivered double amount of water immediately. On the other side, D group preference for TA was modulated by the programmed delay to reward. As delay to reward increased, TA preference decreased. In fact, preference reversed after 4 s delay, but only for Conditions 2 and 3, meaning that rats were not only sensitive to the local delay of reward, but also to the average block delay. Considering our results as an all, our results are consistent with previously mentioned reports about SDVL in pigeons [20], starlings [21], fishes [19], and grasshoppers [18]. This experiment demonstrate that rats also show SDVL.

The fact that D group preferred the previously trained alternative across the first blocks (the lower delays of rewards) between conditions suggests that deprivation increased the alternative's reinforcement value in the long-term. Such idea was previously suggested by [22], also using rats as subjects. Thereafter, other studies have addressed the possible adaptive value of State-Dependent Valuation Learning, SDVL [17,29,16,18]. SDVL assumes that the value of a reinforcement alternative will depend on the state of the animal when it learned about the properties of the alternative. This learned reinforcement value will make the animal prefer an alternative whose properties were initially experienced under a higher state of food deprivation.

This paper also demonstrates that SDVL affect not only preference when both alternative delivers reward immediately, but also when the delay to reward is progressively increases. As far as we know, this is the first evidence supporting that discounting is also affected by SDVL.

A relation between time perception and impulsivity has been suggested in humans. The perception of time as passing more quickly reduces the value of the delayed rewards [30], keeping motivation equal. Other studies in humans have suggested that scarcity (manipulated by an imposed budget) can change how people allocate attention, leading to engage in activities that reduce long-term delayed gains [31]. Thus, procedural and conceptual differences cannot be ignored. For example, scarcity (measured by hypothetical budges) does not necessarily imply food deprivation or high motivation.

Looking for evidence in animals, it has been shown that rats classified as impulsive showed higher preference for the LL option when they were under high levels of deprivation [32]. Therefore, in this case food deprivation increased self-control. Our findings agree with this result.

At least two limitations of our experiment are worth mentioning. First, RS group experienced the delays in ascending order, whereas D group did in descending order. Nevertheless, evidence is not clear about a possible effect of the order of presentation of delays on preference.

On one hand, several studies have reported similar rates of delay discounting with both, ascending and descending delay presentation, using different species. [33] found similar rates of delay discounting in humans exposed to a delay discounting task with ascending and descending delay presentations. Differently, they only found higher rates of delay discounting when delays were randomly presented. Before, [34], also reported absence of ordering effect, using rats exposed to a delay discounting procedure with signaled or unsignaled alternatives. Other studies using the same delay-discounting procedure we used [24,35], and adding a reversal of delay order (increasing order to decreasing order), suggests that the order of delays does not affect the discounting function.

On the other hand, [36] reported that previous experience with delays seems to evoke tolerance to delayed reinforcers in rats, but they did not test the order of presentation of delays. One possibility to account for our results could be that RS group in our study generated tolerance to the delays of the TA, because the ascending order of presentation.

This could have functioned as a training to higher delays, thus facilitating preference for the TA alternative, like [36] reported. However, the D group showed a higher preference for the NA alternative during all the procedure independently of the delay programmed. From our view, both data taken together support that the learned value initially attributed to an alternative depending on the state of the organism, determines how the delays to reward affect preference. Additionally, previous studies using the same delay-discounting procedure than we used in this experiment [24,35] and adding a reversal of the order presentation of the delays (increasing order to decreasing order), have shown that the order of delays does not affect the discounting function.

The second limitation refers to the fact that the values for the delays were taken from an experiment with mice [24], and there could be differences in how rats and mice discount delayed reinforcers. Nevertheless, the form of the discounting function seems similar and stable between experiments, or even between species, regardless the values of delays used [34,37,38].

Our results seem to support SDVL in a delay discounting procedure. Learning about the properties of an alternative under a high state of food deprivation seems to increase later preference for this alternative, even when it is being currently devalued through increasing its delay to reward. SDVL hence can be regarded as an effective procedure to promote self-control. Although the specific

mechanism underlying this effect remains elusive, the value priority hypothesis proposed by [21] and captured in SDVL seems to better account for our findings. The relation between the state of the organism (e. g., motivation, food deprivation, scarcity of resources), previous experience with alternatives and the discounting of delayed reinforcers requires further exploration.

5. Conclusions

- This experiment increases evidence about how the internal state of animals at the moment of learning about the properties of alternatives, will determine its future preference.
- Our results seem to support SDVL in a delay discounting procedure.
- Learning about the properties of an alternative under a high state of food deprivation seems to increase later preference for this alternative, even when it is being currently devalued through increasing its delay to reward.

Author Contributions: Conceptualization, Erick Barrón and Óscar García-Leal; Data curation, Zirahuén Vílchez; Formal analysis, Zirahuén Vílchez; Investigation, Zirahuén Vílchez; Methodology, Zirahuén Vílchez, Erick Barrón, Laurent Ávila-Chauvet, Jonathan Buriticá and Óscar García-Leal; Supervision, Jonathan Buriticá and Óscar García-Leal; Writing – original draft, Zirahuén Vílchez and Jonathan Buriticá; Writing – review & editing, Erick Barrón, Laurent Ávila-Chauvet and Óscar García-Leal.

Funding: This research was funded by the Consejo Nacional de Ciencia y Tecnología (CONACyT, México), grant number 354218.

Institutional Review Board Statement: The experimental procedure was approved by the local Ethical Committee of the Center for Studies and Investigations in Behavior, by the University of Guadalajara committee for animal experiments, and met governmental guidelines.

Data Availability Statement: Data is available from the corresponding author on reasonable request.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Beran, M. J. (2015). The comparative science of “self-control”: What are we talking about? *Frontiers in Psychology*, 6, 1–4. doi: 10.3389/fpsyg.2015.00051
2. Johnson, M. W. & Bickel, W. K. (2002). Within-subject comparison of real and hypothetical money rewards in delay discounting. *Journal of the Experimental Analysis of Behavior*, 77(2), 129–146. doi: 10.1901/jeab.2002.77-129
3. Rachlin, H., & Green, L. (1972). Commitment, choice and self-control. *Journal of the Experimental Analysis of Behavior*, 17(1), 15–22.
4. Evenden, J. L. (1999). Varieties of impulsivity. *Psychopharmacology*, 146, 348–361. doi: 10.1007/pl00005481
5. Baumeister, R. F., Vohs, K. D. & Tice, D. M. (2007). The strength model of self-control. *Current Directions in Psychological Science*, 16(6), 351–355. doi: 10.1111/j.1467-8721.2007.00534.x
6. McKerchar, T. L., Green, L., Myerson, J., Pickford, T. S., Hill, J. C. & Stout, S. C. (2009). A comparison of four models of delay discounting in humans. *Behavioural Processes*, 81(2), 256–259. doi: 10.1016/j.beproc.2008.12.017
7. Lancaster, K. (1963). An axiomatic theory of consumer time preference. *International Economic Review*, 4(2), 221–231. doi: 10.2307/2525488
8. Meyer, R.F. (1976). Preferences over time. In: R. L. Keeney & H. Raiffa (Eds.), *Decisions with multiple objectives: Preferences and value tradeoffs* (pp. 473–514). New York: Wiley.
9. Bickel, W. K., Odum, A. M. & Madden, G. J. (1999). Impulsivity and cigarette smoking: delay discounting in current, never, and ex-smokers. *Psychopharmacology*, 146(4), 447–454. doi: 10.1007/pl00005490
10. Mazur, J.E. (1987). An adjusting procedure for studying delayed reinforcement. In: M.I. Commons, J.E. Mazur, J.A. Nevin & H. Rachlin (Eds.), *Quantitative analyses of behavior*, Vol. 5, Effects of delay and of intervening events on reinforcement value. Hillsdale, NJ: Erlbaum.
11. Meindl, J.N. (2012). Understanding preference shifts: A review of within-trial contrast and state-dependent valuation. *The Behavior Analyst*, 35(2), 179–195. doi: 10.1007/bf03392277
12. Clement, T. S., Feltus, J. , Daren, H. & Zentall, T. R. (2000). “Work ethic” in pigeons: Reward value is directly related to the effort or time required to obtain the reward. *Psychonomic Bulletin & Review*, 7(1), 100–106. doi: 10.3758/bf03210727
13. Zentall, T. R. (2008). Within-trial contrast: When you see it and when you don’t. *Learning & Behavior*, 36, 19–22. doi: 10.3758/lb.36.1.19
14. Kacelnik, A. & Marsh, B. (2002). Cost can increase preference in starlings. *Animal Behaviour*, 63(2), 245–250. doi: 10.1006/anbe.2001.1900

15. Marsh, B., Schuck-Paim, C. & Kacelnik, A. (2004). Energetic state during learning affects foraging choices in starlings. *Behavioral Ecology*, 15(3), 396–399. doi: 10.1093/beheco/arh034
16. McNamara, J. M., Trimmer, P. C. & Houston, A. I. (2012). The ecological rationality of state-dependent valuation. *Psychological Review*, 119(1), 114–119. doi: 10.1037/a0025958
17. Fox, A. E. & Kyonka, E. G. E. (2014). Choice and timing in pigeons under differing levels of food deprivation. *Behavioural Processes*, 106, 82–90. doi: 10.1016/j.beproc.2014.04.018
18. Pompilio, L., Kacelnik, A. & Behmer, S.T. (2006). State-Dependent Learned Valuation Drives Choice in an Invertebrate. *Science*, 311(5767), 1613–1615. doi: 10.1126/science.1123924
19. Aw, J. M., Holbrook, R. I., Burt de Perera, T. & Kacelnik, A. (2009). State-dependent valuation learning in fish: Banded tetras prefer stimuli associated with greater past deprivation. *Behavioural Processes*, 81(2), 333–336. doi:10.1016/j.beproc.2008.09.002
20. Vasconcelos, M., & Urcuioli, P. J. (2008). Deprivation level and choice in pigeons: A test of within-trial contrast. *Learning & Behavior*, 36(1), 12–18. doi: 10.3758/LB.36.1.12
21. Pompilio, L. & Kacelnik, A. (2005). State-dependent learning and suboptimal choice: when starlings prefer long over short delays to food. *Animal Behaviour*, 70, 571-578. doi: 10.1016/j.anbehav.2004.12.009
22. Christensen-Szalanski, J. J. J., Goldberg, A. D., Anderson, M. E., & Mitchell, T. R. (1980). Deprivation, delay of reinforcement, and the selection of behavioural strategies. *Animal Behaviour*, 28(2), 341–346. doi: 10.1016/S0003-3472(80)80042-X
23. García-Leal, O., Saldivar, G. & Díaz, C. A. (2008). Efecto de la disponibilidad de recursos energéticos en la sensibilidad al riesgo en ratas (*Rattus norvegicus*). *Acta Comportamentalia*, 16, 25–40. doi: 10.5514/rmac.v33.i2.16253
24. Isles, A. R., Humby, T., Walters, E. & Wilkinson, L.S. (2004). Common genetic effects on variation in impulsivity and activity in mice. *The Journal of Neuroscience*, 24(30), 6733-6740. doi: 10.1523/JNEUROSCI.1650-04.2004
25. Evenden, J.L. & Ryan, C. N. (1999). The pharmacology of impulsive behaviour in rats VI: the effects of ethanol and selective serotonergic drugs on response choice with varying delays of reinforcement. *Psychopharmacology*, 146(4), 413–421. doi: 10.1007/pl00005486
26. Kacelnik, A., Vasconcelos, M., Monteiro, T., & Aw, J. (2011). Darwin's "tug-of-war" vs. starlings' "horse-racing": How adaptations for sequential encounters drive simultaneous choice. *Behavioral Ecology and Sociobiology*, 65(3), 547–558. doi: 10.1007/s00265-010-1101-2
27. Shapiro, M. S., Siller, S., & Kacelnik, A. (2008). Simultaneous and Sequential Choice as a Function of Reward Delay and Magnitude: Normative, Descriptive and Process-Based Models Tested in the European Starling (*Sturnus vulgaris*). *Journal of Experimental Psychology: Animal Behavior Processes*, 34(1), 75–93. doi: 10.1037/0097-7403.34.1.75
28. García-Leal, O., Rodríguez, E. & Camarena, H. O. (2017). Análisis empírico del Modelo de Elección Secuencial en Humanos. *Avances en Psicología Latinoamericana*, 36(1), 139. doi: 10.12804/revistas.urosario.edu.co/apl/a.4869
29. Halpern, J.Y. & Seeman, L. (2018). Is state-dependent valuation more adaptive than simpler rules? *Behavioral Processes*, 147, 33-37. doi: 10.1016/j.beproc.2017.12.001
30. Bauman, A.A. & Odum, A. (2012). Impulsivity, risk-taking, and timing. *Behavioral Processes*, 90, 408-414. doi: 10.1016/j.beproc.2012.04.005
31. Shah, A.K., Mullainathan, S. & Shafir, E. (2012). Some consequences of having too little. *Science*, 338. doi: 10.1126/science.1222426
32. Zaichenko M.I. & Merzhanova, G.Kh. (2011). Studies of impulsivity in rats in conditions of choice between food reinforcements of different value. *Neuroscience and Behavioral Physiology*, 41(5), 445-451. doi: 10.1007/s11055-011-9435-4
33. Robles, E. & Vargas, P.A. (2007). Functional parameters of delay discounting assessment tasks: Order of presentation. *Behavioral Processes*, 75(2), 237-241. <https://doi.org/10.1016/j.beproc.2007.02.014>
34. Slezak, J.M. & Anderson, K.G. (2009). Effects of variable training, signaled and unsignaled delays, and d-amphetamine on delay-discounting functions. *Behavioral Pharmacology*, 20, 424-436. doi: 10.1097/FBP.0b013e3283305ef9
35. Isles, A. R., Humby, T. & Wilkinson, L. S. (2003). Measuring impulsivity in mice using a novel operant delayed reinforcement task: effects of behavioural manipulations and d-amphetamine. *Psychopharmacology*, 170(4), 376–382. doi: 10.1007/s00213-003-1551-6
36. Stein, J.S. & Madden, G.J. (2013). Delay discounting and drug abuse: Empirical, conceptual, and methodological considerations. In: J. MacKillop & J. de Wit (Eds.), *The Wiley-Blackwell Handbook of Addiction Psychopharmacology*. John Wiley & Sons, Ltd.
37. Bradshaw, C. M. & Szabadi, E. (1992). Choice Between Delayed Reinforcers in a Discrete-trials Schedule: The Effect of Deprivation Level. *The Quarterly Journal of Experimental Psychology Section B*, 44(1), 1–16. doi: 10.1080/02724999208250599

38. Reynolds, B., de Wit, H. & Richards, J.B. (2002). Delay of gratification and delay discounting in rats. *Behavioral Processes*, 59, 157-168. doi: 10.1016/S0376-6357(02)00088-8

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.