Behavioural Processes 130 (2016) 71–74

Contents lists available at ScienceDirect

Behavioural Processes

journal homepage: www.elsevier.com/locate/behavproc

Temporal discrimination and delayed reinforcement

Jonathan Buriticá*, Zirahuén Vilchez, Cristiano Valerio dos Santos

CEIC, Universidad de Guadalajara, Mexico

ARTICLE INFO

Article history:
Received 21 March 2016
Received in revised form 11 July 2016
Accepted 13 July 2016
Available online 16 July 2016

Keywords:
Timing
Temporal bisection
Delay of reinforcement
Reinforcement value
Rats

ABSTRACT

We attempted to determine the effect of reinforcement delay on time discrimination in an interval bisection task. Three groups of rats were exposed to immediate, delayed reinforcement and longer signals with immediate reinforcement in acquisition and test. Results show differences in the amount of training necessary to reach the acquisition criteria, the Weber fraction and the range or overall stimulus control. The results suggest an increased difficulty to discriminate the difference among durations rather than an increase in estimated time as main effect of delayed reinforcement.

1. Introduction

The interval bisection is a matching-to-sample procedure with stimulus durations as the sample. A stimulus of shorter or longer duration is presented, after which the subject chooses between two alternatives; if the correct response associated to the stimulus duration is emitted, reinforcement is delivered. After a discrimination has been established, a test phase with nonreinforced intermediate stimulus durations is conducted and responses to the options recorded. The main result is an increase in the probability of long response after longer durations (Church and Deluty, 1977). The task has two main components: acquisition and test. In acquisition, the subject learns the discrimination. In the test, evidence is collected about what the subject learned, how the subjects evaluate whether durations are short or long, and overall stimulus control.

Discrimination training with various stimulus modalities often produces a negatively accelerated acquisition curve and an asymptotic level of performance (Kehoe, 2008). However, this pattern can be disrupted by manipulations of reinforcement delay, pre-feeding or magnitude (Lotfizadeh et al., 2012). In a visual discrimination task with delayed reinforcement, for instance, a slower acquisition was found when reinforcement delays were 0.5–5 s, and no acquisition at all was observed with a 10-s delay (Grice, 1948). Sargisson and White (2003) reported lower discriminability in a color matching-to-sample discrimination task when there was a delay between response and reinforcer, and argued that this effect is similar to that obtained with lower reinforcement magnitude.

In interval bisection tasks, Galtress and Kirkpatrick (2010) trained rats with a high reinforcement magnitude, but the results of acquisition are not reported because subjects did not meet the acquisition criterion. However, they found a shorter Bisection Point (BP) in the test when the long duration was reinforced with a higher magnitude, and a flattening function when short or long durations were reinforced with a small magnitude. McClure et al., 2009 found a longer BP, but no differences in overall stimulus control when they pre-fed pigeons before the test with a high amount of food.

Timing theories such as BeT (Killeen and Fetterman, 1988) and LeT (Machado, 1997) predict that changes in reinforcement value affect timing through variations in the pacemaker rate. The rate of pulses in BeT or the transition rate between behavioral states in LeT both are dependent of arousal and devaluation should decrease arousal, and hence pacemaker rate. Thus, devaluation should move the BP to the right due to the lower number of accumulated pulses or to the elapsed behavioral states. Nonetheless, the BP change should be transient, and the repeated experience of reinforcement in the new behavioral state or pulse should update the moment of response to the actual time of reinforcement. Therefore, in a long-term exposure to devaluation (low arousal), the expected time to reinforcement should be the same to the time without devaluation (high arousal).

Scalar Expectancy Theory cannot explain the effects of reinforcement value manipulations on timing, because in its original version it does not include characteristics of the reward to explain the output (Gibbon et al., 1984). Versions of SET that include the effects of the reward have been proposed to include the effects

* Corresponding author at: 180 Fco. de Quevedo, Col. Arcos de Vallarta, Guadalajara, Jal. 44130, Mexico. Tel.: +52 33 37771150x33310.
E-mail address: jjburitica@gmail.com (J. Buriticá).

http://dx.doi.org/10.1016/j.beproc.2016.07.009
0376-6357/

of manipulations of motivation. Galtress and Kirkpatrick (2009, 2010) propose that motivation may affect attention, instantiated in the switch accumulation process of SET; this proposal, however, cannot fully account for all the reported results (see Galtress et al., 2012). According to Galtress et al. (2012), reward value manipulations can interact with many components of the theory at the same time, such as the switch accumulation processes and the mean or standard deviation of response thresholds. Balcı (2014) proposed that motivation manipulations affect the time of response in anticipatory timing procedures due to changes in thresholds to initiate response associated to dopamine levels. Low levels of dopamine should increase threshold meanwhile high levels should do the opposite. Levels of dopamine should be dependent on motivation manipulations or reinforcement value.

To apply Balcı (2014) account to a temporal bisection procedure, it is necessary to think the task as a dual peak procedure. In a dual peak procedure, subjects learn to respond to one shorter interval and then to a second longer interval. In the test, the response has two peaks, one at the first interval and another at the second interval. The aforementioned is translated in the bisection procedure as a high probability of responding short with shorter signal durations and long with longer signals: the subject will approach to the operandum associated to the short signal and then, as time passes, it approaches the long signal operandum. A similar idea has been proposed by de Carvalho et al. (2016); they suggest that temporal bisection performance could be the result of the combination of two generalization gradients around the trained intervals. Following the logic of Balci, we could expect that reinforcement devaluation should increase the threshold to initiate responding at both short and long intervals thus predicting an increase in the BP.

Temporal discrimination follows Weberís law, and thus the acquisition might be slower in the delayed condition if the subject's behavior is controlled by the sum of stimulus duration (short = 2; long = 8 s) and the reinforcement delay (3 s), because the ratio of the perceived stimulus would be higher (ratio 5/11 = 0.45, if the delay is added) in comparison to the ratio of stimulus with immediate reinforcement (ratio 2/8 = 0.25). The same processes may also affect the bisection function, and produce a longer BP in the delayed group, not because of the devaluation, but because of the increased subjective time. Hence, we added a group with longer signals (5 and 11 s), immediate reinforcement and the same test trials to compare its results with delayed reinforcement. We tried to determine whether the discriminated time is the duration of the signal, the interval between stimulus onset to reinforcement delivery or end of trial and whether the effect is explained by Weberís law or by reinforcement devaluation. We expected to find slower learning in training and less sensitivity at test with delayed reinforcement. If reinforcement delay work through a mechanism of slowing pacemaker rate, or changing the threshold to initiate responding, we should observe an increased BP in that condition as well.

2. Method

2.1. Apparatus

Ten rat conditioning chambers from MED-PC Associates, each with one water dispenser in the central panel with a 0.1 ml cup and housed in sound attenuating boxes, were used. Two retractable levers were positioned adjacent to the water dispenser 2.5 cm above the grid floor; 0.12 N were required to record a response. A house light and a speaker were located in the rear panel. Event programming and data recording were conducted using MED-PC IV.

2.2. Subjects

Twenty-four naïve female Long Evans rats, bred in our own laboratory, aged three months at the beginning of experiment. They had free access to food but received only 15 min of access to water per day, 30 min after the session. One subject did not press the lever steadily, so it was dropped from the experiment. They were divided into three groups: immediate (n = 7), delayed (n = 8) and long-signal immediate reinforcement (n = 8).

2.3. Procedure

In pretraining, subjects were exposed to a conjunctive Fixed-Time (FT) Continuous Reinforcement (CRF) schedule, so that every 30 s the cup was raised and the subject had access to water. Both levers were randomly presented, one at a time and no more than three reinforcers in a row were delivered with the same lever present. If one lever press occurred, the response was reinforced and the lever retracted. The FT was increased by 30 s every day until 120 s if the subject failed to obtain 100 reinforcers per session. After the criterion was met, the CRF was in effect in the subsequent sessions until subjects obtained at least 100 reinforcers in two consecutive sessions. Every pre training session lasted 30 min or 150 reinforcers.

Training and testing were similar to the procedure employed by Orduña et al., 2007. Each session consisted of 81 trials. A trial started with two levers present and the house light on. A tone was presented for 2 or 8 s for the groups with immediate and delayed reinforcement and for 5 or 11 s for the group with long-signal immediate reinforcement. Depending on the duration of the tone, a response to the right or left lever was reinforced. Training and testing trials in the group with delayed reinforcement were similar to those of groups with immediate reinforcement, except that a 3-s delay was programmed between the correct response and reinforcement. If no response occurred within 20 s, the house light was turned off, levers were retracted, and a 30-s Inter Trial Interval (ITI) began. If the response was incorrect, the levers were retracted, the house light went off and the ITI began; the same duration was presented in the next trial until a correct choice was made. If the response was correct, the levers were retracted and the water dispenser was activated for five seconds, then the house light and water dispenser turned off and the ITI began. At least ten sessions were run in acquisition. After the subjects obtained 80% correct responses in three consecutive sessions, reinforcement rate was diminished, and only 75% of correct responses were reinforced.

The test began after 80% correct responses were obtained with 75% reinforcement rate. In test sessions, trials with the durations used in acquisition were interspersed with trials of intermediate durations (2.52, 3.17, 4, 5.04, 6.35 s). Test stimulus durations were presented four times per session, and responses in these test trials were never reinforced. The test was carried out for fifteen sessions.

2.4. Data analysis

The results of the test were fitted to a two parameter model: $P_{long} = 1/(1 + (t/T_{50})^{\epsilon})$. P_{long} is the probability of pressing the lever associated with the long duration, T_{50} and ϵ are free parameters, representing BP and the slope of the function (Orduña et al., 2007) respectively; ϵ had negative values. We also computed the Difference Limen as half the difference between the signal duration producing 0.25 and 0.75 probability of long responses. We calculated the Weber fraction as the ratio Difference Limmen/BP. To assess overall stimulus control, we calculated the range as the difference between the probability of responding long when the