Cognition 139 (2015) 154–167

Contents lists available at ScienceDirect

Cognition

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

Reward and punishment act as distinct factors in guiding behavior

Jan Kubanek [a,*], Lawrence H. Snyder [a], Richard A. Abrams [b]

[a] Department of Anatomy and Neurobiology, Washington University School of Medicine, St. Louis, MO 63110, USA
[b] Department of Psychology, Washington University in St. Louis, St. Louis, MO 63130, USA

ARTICLE INFO

Article history:
Received 10 June 2014
Revised 4 March 2015
Accepted 6 March 2015
Available online 28 March 2015

Keywords:
Instrumental (operant) behavior
Law of effect
Reinforcement
Penalty
Gain
Loss

ABSTRACT

Behavior rests on the experience of reinforcement and punishment. It has been unclear whether reinforcement and punishment act as oppositely valenced components of a single behavioral factor, or whether these two kinds of outcomes play fundamentally distinct behavioral roles. To this end, we varied the magnitude of a reward or a penalty experienced following a choice using monetary tokens. The outcome of each trial was independent of the outcome of the previous trial, which enabled us to isolate and study the effect on behavior of each outcome magnitude in single trials. We found that a reward led to a repetition of the previous choice, whereas a penalty led to an avoidance of the previous choice. Surprisingly, the effects of the reward magnitude and the penalty magnitude revealed a pronounced asymmetry. The choice repetition effect of a reward scaled with the magnitude of the reward. In a marked contrast, the avoidance effect of a penalty was flat, not influenced by the magnitude of the penalty. These effects were mechanistically described using a reinforcement learning model after the model was updated to account for the penalty-based asymmetry. The asymmetry in the effects of the reward magnitude and the punishment magnitude was so striking that it is difficult to conceive that one factor is just a weighted or transformed form of the other factor. Instead, the data suggest that rewards and penalties are fundamentally distinct factors in governing behavior.

1. Introduction

Reinforcement and punishment constitute Nature's arsenal in guiding behavior (Thorndike, 1898, 1911; Skinner, 1963; Tversky & Kahneman, 1986; Davison, 1991; Gray, Stafford, & Tallman, 1991; Ehrlich, 1996; Hackenberg, 2009). It is well established that reinforcers and punishers both critically influence behavior, but it has been unclear whether these factors exert symmetric or qualitatively distinct behavioral effects (Skinner, 1953; Farley & Fantino, 1978; Gray et al., 1991; Dinsmoor, 1998; Lerman & Vorndran, 2002; Critchfield, Paletz, MacAleese, & Newland, 2003; Lie & Alsop, 2007). One-factor theories have proposed a symmetric law of effect (Thorndike, 1927). In this view, reinforcement increases behavior frequency, punishment decreases behavioral frequency, and the magnitudes of these effects are equal, just of opposite signs (Thorndike, 1911; Sidman, 1962; Herrnstein & Hineline, 1966; Schuster & Rachlin, 1968; Rachlin & Herrnstein, 1969; Villiers, 1980). In contrast, two-factor theories view reinforcement and punishment as qualitatively distinct influences on operant behavior (Mowrer, 1947; Dinsmoor, 1954; Epstein, 1985; Yechiam & Hochman, 2013).

* Corresponding author at: Department of Anatomy & Neurobiology, Washington University School of Medicine, 660 S Euclid Ave, St. Louis, MO 63110, USA. Tel.: +1 314 552 1169.

E-mail address: kubanek@wustl.edu (J. Kubanek).

http://dx.doi.org/10.1016/j.cognition.2015.03.005
0010-0277/

This debate remains, for the most part, unresolved (Hineline, 1984; Gray et al., 1991; Dinsmoor, 1998, 2001; Critchfield et al., 2003; Lie & Alsop, 2007). This is mainly due to two reasons. First, it is difficult to compare qualitatively different factors (e.g., food versus electric shock) on a common scale (Schuster & Rachlin, 1968; Farley & Fantino, 1978; Villiers, 1980; Fiorillo, 2013). A solution to this problem is to work with reinforcers and punishers that are of the same kind—using tokens that represent gains and losses (Hackenberg, 2009). Second, previous studies targeting this question have employed relatively complex paradigms (Bradshaw, Szabadi, & Bevan, 1979; Gray et al., 1991; Critchfield et al., 2003; Rasmussen & Newland, 2008). The complex paradigms make it difficult to readily investigate the effect of a reward or a punishment on a behavioral response.

We addressed this question in a simple choice paradigm in which we varied the *magnitude* of a reward or a penalty experienced following each choice. This allowed us to measure subjects' tendency to repeat their previous choice as a function of the magnitude of the experienced reward or penalty. In this simple paradigm, one-factor theories predict that the reward and penalty magnitudes will lead to qualitatively similar, just oppositely signed tendencies to repeat the previous choice. In contrast, two-factor theories predict that the choice repetition tendencies will be qualitatively distinct for the two factors. The data indeed revealed a striking asymmetry in the effects of the reward and penalty magnitudes on the choice behavior. The asymmetry was so profound that it suggests that the two behavioral factors are of distinct natures.

2. Materials and methods

2.1. Subjects

Eighty-eight Washington University undergraduate students participated in this study. The subjects performed an Auditory task or a Visual task. The Auditory task was performed by 54 students (37 females, 17 males), aged 18–21 (mean 19.2). The Visual task was performed by a distinct set of 34 students (24 females, 10 males), aged 18–23 (mean 19.4). All subjects were healthy, had normal hearing capacity, and gave an informed consent. Subjects participated for class credit.

2.2. Auditory task

Subjects sat in a comfortable chair 70 cm in front of a flat-screen monitor. Subjects wore headphones (MDR-V600, Sony), which presented a stereo auditory stimulus (see Auditory stimulus). The subjects' hands were comfortably positioned at a computer keyboard, with the left index finger placed over the left Command key and with their right index finger placed over the right Command key. The control of the experimental design was accomplished using a custom program written in Matlab (The Mathworks, Inc., Natick, MA, RRID:nlx_153890).

Each trial started with the presentation of a red fixation cross, 2° in size. Subjects were instructed to fixate at the center of the cross. At the same time, subjects were presented with a stereo auditory stimulus (click sounds, see Auditory stimulus), 1.0 s in duration (Fig. 1A). After the stimulus has been presented, the fixation cross shrank to 1° and changed its color to green. This event cued the subjects to make a movement (choice). Subjects performed 2 blocks of 300 trials each, with a brief period in between. In the first block of 300 trials, subjects were instructed to press the left Command key with their left index finger if they heard more clicks in the left ear and to press the right Command key with their right index finger if they heard more clicks in the right ear. In the second block of 300 trials, this instructed contingency was reversed. We found similar results in both blocks and therefore pooled the data over the two blocks. In 20% of trials, we randomly interleaved cases in which no auditory stimulus was present. When no sound was heard, subjects were instructed to choose either key (i.e., to either press the left key with the left index finger or the right key with the right index finger). The purpose of these trials was to investigate the effect of outcome on choice when no perceptual stimulus is present (Fig. 3B).

If subjects responded prior to the green cue or if they failed to indicate a response within 1200 ms after the cue, the trial was considered invalid, and was aborted and excluded from the analyses. The type of error was indicated to the subjects in red, large-font text ('TOO EARLY', 'TOO LATE'). The proportion of valid choices over the subjects was 96.0% ± 1.0 (mean ± s.d.). A response was immediately followed by a display of the outcome. Specifically, a correct response was followed by the display of a green string that was randomly drawn from the set $\{+5c, +10c, +15c, +20c, +25c\}$. An incorrect response was followed by the display of a red string randomly drawn from the set $\{-5c, -10c, -15c, -20c, -25c\}$. These strings were chosen to represent "cents"; the subjects received no instruction in this regard. The outcome was displayed for 0.5 s. The next trial started immediately following the offset of the outcome.

2.3. Auditory stimulus

The auditory stimulus was equivalent to that used previously (Kubanek, Snyder, Brunton, Brody, & Schalk, 2013). Briefly, each ear was presented with a train of brief (0.2 ms) clicks sounds drawn from a homogeneous Poisson process. Each train lasted 1.0 s. The stereo stimulus was composed such that the sum of clicks presented to the left ear (C_l) plus the sum of clicks presented to the right ear (C_r) summed to a fixed number $C_l + C_r = \Omega, \Omega \in \{25, 32, 39, 46\}$. Since C_r and C_l were drawn randomly in each trial (and randomly in each subject), the polarity (leftward, rightward evidence) of the stimulus was random in each trial. The value of Ω was drawn randomly on each trial. The Ω randomization was imposed to ensure that subject had to pay attention to the click sounds in both ears.

2.4. Visual task

The Visual task was analogous to the Auditory task. We therefore only specify the differences. In the Visual task,