Personality and Individual Differences 134 (2018) 155–163

Contents lists available at ScienceDirect

Personality and Individual Differences

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

Role of self-criticism in reward and punishment probabilistic learning

O. Duek[a,*], G. Shahar[a], Y. Osher[b], O. Kofman[a]

[a] Department of Psychology, Ben Gurion University of the Negev, Beer Sheva, Israel
[b] Beer Sheva Mental Health Center, Department of Psychiatry, Faculty of Health Sciences, Ben Gurion University of the Negev, Beer Sheva, Israel

ARTICLE INFO

Keywords:
Self-criticism
Depression
Probabilistic-learning
BIS/BAS
Personality

ABSTRACT

Little is known about the effect of personality traits on learning. Thus the aim of this investigation was to better understand the role of depressive personality traits – primarily self-criticism and approach and inhibition tendencies – in reward and punishment learning. In two studies (Study 1: $N = 38$; Study 2: $N = 100$), we used a probabilistic classification task in which participants needed to categorize ambiguous stimuli, and then received probabilistic feedback, according to their choice. In Study 2, we employed a variation of this task with difficult vs. easy contingencies. In both studies we examined the association between performance in the task and approach and avoidance personality traits (BIS/BAS, self-criticism and positive generalization) while controlling for depression and intelligence.

Self-criticism and a tendency to generalize positive events were positively associated with reward, but not punishment, learning. As well, after exposure to difficult contingencies, participants had delayed punishment learning. In light of these results, we suggest that self-criticism might enhance monitoring of errors, which improves reward learning.

1. Introduction

Since its infancy, learning research has been dominated by theoretical behaviourism, characterized by an emphasis on conditioning, shaping, and behaviour modification. However, very early in the development of the field, Breland and Breland (1961) described how species' specific behaviours or traits could dominate conditioning, reopening the enduring debate over the influence of innate behavioural tendencies on learning (Shettleworth, 1978). This approach inspired research on the role of genetics in animal learning (Dalla & Shors, 2009), as well as on the role of human personality in conditioning. Thus, drawing from Cloninger's theory of personality (Cloninger, Przybeck, Svrakic, & Wetzel, 1994), Bodi et al. (2009) showed that *novelty seeking, pertaining to the* tendency to impulsively explore the environment in order to find novel stimuli (Cloninger et al., 1994), was associated with improved *reward* learning, whereas *harm avoidance,* referring to anticipatory worry and fear of uncertainty, was associated with better performance on punishment learning.

Approach and avoidance tendencies are central to research on reward- and punishment-based conditioning. Nevertheless, these traits have not been explored in relation to the ability to self-correct following an unrewarded or punished response. Since learning entails modification of behaviour in the face of loss of reward or of punishment, personality traits that reflect responsivity to success or failure have been the subject of investigation in both healthy participants and those with diagnosed psychopathology. One such trait is self-criticism, pertaining to the tendency to set unrealistically high self-standards and to adopt a punitive stance toward oneself (Shahar, 2015; Shahar, Blatt, Zuroff, & Pilkonis, 2003). Self-criticism appears to be causally implicated primarily in depressive disorders, but also in bipolar, anxiety, and eating disorders, as well as in suicidality and psychosomatics (for review, see Shahar, 2015; Shahar & Henrich, 2013).

Factors related to self-criticism have been found to contribute to learning. Accordingly, Brand and Altstötter-Gleich (2008) found that concern over mistakes and high personal standards, two dimensions of perfectionism that are associated with self-criticism (Dunkley, Blankstein, Masheb, & Grilo, 2006; Dunkley, Zuroff, & Blankstein, 2006), were linked with better performance in the Game of Dice Task (GDT; Brand & Altstötter-Gleich, 2008), but not in the Iowa Gambling Task (IGT; Bechara, Damasio, Damasio, & Anderson, 1994). The GDT is a computerized task in which the participant is expected to maximize gain. To do so, participants are required to gamble on the number that might appear in the next throw of one die, and they can choose to gamble with only one number – maximizing their gain and risk of losing – or choose up to four numbers in one throw – minimizing both their gain and risk (Brand et al., 2005). The net score is calculated by subtracting amounts of risky choices (choosing one or two numbers) from amounts of non-risky choices (choosing three or four numbers), so that

* Corresponding author.
E-mail address: ord@post.bgu.ac.il (O. Duek).

https://doi.org/10.1016/j.paid.2018.06.007
Received 12 March 2018; Received in revised form 31 May 2018; Accepted 4 June 2018

a higher net score suggests that the participant favours a low-risk strategy. Brand and Altstötter-Gleich (2008) found that concern over mistakes and high personal standards were *positively* associated with the net score, thus showing better strategy and maximizing gain.

The IGT is a computerized task in which participants must learn to maximize profit by minimizing high-gain, high-risk choices. Four decks of cards are presented to the participants: two of which produce higher immediate gain, but also higher long-term losses, and two of which have lower immediate gain, but lower losses. The net score in the IGT indicates a preference for advantageous choices over disadvantageous choices. Brand et al. (2005) did not, however, find any association between three facets of perfectionism (concern over mistakes, personal standards, and doubts about action) and performance in IGT. Thus, they concluded that perfectionism is related to risky (as in the GDT), but not ambiguous (as in the IGT), learning.

Both IGT and GDT involve a probability-based feedback to the participant, which affects their performance. However, these tasks do not differentiate between positive – appetitive, or reward-related – and negative – aversive, or punishment-related – feedback. Learning from either reward or punishment involves different neural substrates, which may be differentially affected by personality traits, such as self-criticism. Reward learning is dependent upon mesolimbic dopaminergic (DA) projections, while punishment learning is associated with serotonergic projections (Cools, Nakamura, & Daw, 2010; Cools, Robinson, & Sahakian, 2007).

In order to better differentiate between reward and punishment substrates, we used, in the current study, a probabilistic classification task, in which feedback is based on probability of correct answers (details are in the Methods section). Previous research has found that performance on a probabilistic learning task is grossly impaired in patients with Parkinson's disease (PD), compared to healthy controls and compared to amnesic patients (Knowlton, Mangels, & Squire, 1996). Other studies have found that DA is positively associated with reward. Abler and his colleagues showed that intake of olanzapine (a D2 and 5-HT2 antagonist) reduced activation in the ventral striatum, an area associated with reward learning, and lengthened reaction times in response to reward stimuli (Abler, Erk, & Walter, 2007). Sharot and her colleagues showed that administration of dihydroxy-L-phenylalanine (L-DOPA), a drug which enhances dopaminergic function, increased optimism bias in healthy participants (Sharot, Guitart-Masip, Korn, Chowdhury, & Dolan, 2012). Comparatively, punishment-based learning was mainly associated with serotonergic projections (Cools et al., 2007; Cools et al., 2010). While unmedicated Parkinson's disease patients were impaired on a reward-based probabilistic classification task (PCT), they did not differ from healthy participants when the PCT was based on punishment (Bodi et al., 2009; Shohamy, Myers, Kalanithi, & Gluck, 2008).

In the present context, it is important to consider Gray's Reinforcement Sensitivity Theory (RST; Gray & McNaughton, 2004), wherein Gray describes three behavioural systems that organize human behaviour. The first is the behavioural approach system (BAS), which mediates behavioural approach, and has been correlated with the activation of mesolimbic and mesocortical dopamine pathways in the brain, including the source of the projections in the ventral tegmental area, and the terminal regions of these projections in the nucleus accumbens, the orbitofrontal cortex, anterior cingulate cortex (ACC), and dorsolateral prefrontal cortex (DLPFC; Berns, McClure, Pagnoni, & Montague, 2001; Depue & Collins, 1999; Depue & Iacono, 1989). The second system, the fight-flight-freeze system (FFFS), is responsible for organizing behaviour in response to aversive stimuli. The third system, the behavioural inhibition system (BIS), is responsible for goal-related conflict resolution, which might pertain to the resolution of approach–avoidance conflict, but also to the resolution of approach–approach or avoidance–avoidance conflicts (Gray & McNaughton, 2004). In line with RST, Carver and White (1994) developed a questionnaire that distinguishes between BIS and BAS tendencies. In this questionnaire, BAS includes three different, albeit somewhat related, sub-scales: BAS reward responsiveness (BAS-RR), which focuses on positive responses to reward or the anticipation of reward; BAS drive (BAS-D), which focuses on the persistent pursuit of desired goals; and BAS fun seeking (BAS-F), which focuses on the willingness to approach a potentially rewarding stimulus in the spur of the moment and the desire for the new reward (Carver & White, 1994). The authors of the questionnaire report medium-size correlations between the BAS subscales. Specifically, BAS-D was correlated 0.34 with BAS-RR and 0.41 with BAS-F, and BAS-RR was correlated 0.36 with BAS-F (Carver & White, 1994).

Given the different substrates for reward and punishment, in the current investigation, we examined the role of self-criticism in learning with punitive and rewarding feedback (Barto, 1995) in healthy participants. We hypothesized that, in a healthy population, depression would increase sensitivity to punishment and blunt sensitivity to reward, leading to deterioration of performance. On the other hand, self-criticism in healthy individuals was predicted to counteract the deleterious effects of depression and enhance adaptive response modification following non-reward or punishment. In Study 1, we investigated the effects of self-criticism and other approach-and avoidance-related traits on reward- and punishment-based probabilistic learning. In Study 2, we manipulated the difficulty level of the task by exposing half of the participants to greater task difficulty before exposing them to an easy task. We hypothesized that the hard task would a) impair performance on the easy task, and that b) self-criticism would promote response modification, thereby counteracting the effect of the difficult task.

2. Study 1

2.1. Method

2.1.1. Participants and procedure

Thirty-eight 1st-year psychology students (32 women, $M_{age} = 24.1$; $SD = 1.26$) participated in the experiment in exchange for course credit. Upon arrival at the laboratory, participants provided their informed consent. Participants were then seated in front of a computer screen with 1024 × 768 resolution at eye-level. Next, the experimenter administered the probabilistic classification task (PCT). Following the task, the participants filled out questionnaires (see below) in a different room. Last, the experimenter administered the Raven Standard Progressive Matrices.

2.1.2. Probabilistic classification

We employed a probabilistic classification task based on feedback, like the one used by Bodi et al. (2009). In each trial, participants viewed one of four images, and were asked to guess whether it belonged to category A or B. For each participant, the four images were randomly named S1, S2, S3, and S4. The stimuli were different-coloured geometric shapes with a fractal composition. At any given trial, stimuli S1 and S3 belonged to category A with 80% probability and to category B with 20% probability, while stimuli S2 and S4 belonged to category B with 80% probability and to category A with 20% probability. Stimuli S1 and S2 pertained to *a reward-learning condition*. Thus, if the participant correctly guessed category membership on a trial with either of these stimuli, a reward of +25 points was received. If the participant guessed incorrectly, no points were gained. Stimuli S3 and S4 pertained to a punishment-learning condition. Thus, if the participant guessed incorrectly on a trial with either of these stimuli, 25 points were deducted. However, correct responses in the punishment-learning trials lead to neither loss nor gain of points (Table 1).

The experiment was conducted on a PC and was programmed with PsychToolBox (Brainard & Vision, 1997; Pelli, 1997). The keys for A and B were labelled on the keyboard's S and K keys, respectively. The experiment began with the following instruction screen (translated from Hebrew):

"*You will see different pictures on the screen, one picture in each trial.*