Learning and Motivation 56 (2016) 38–47

Contents lists available at ScienceDirect

Learning and Motivation

journal homepage: www.elsevier.com/locate/l&m

Conditioned reinforcement and backward association

Arthur Prével[a,*], Vinca Rivière[a], Jean-Claude Darcheville[a], Gonzalo P. Urcelay[b]

[a] Laboratoire SCALab, UMR 9193, Université de Lille Nord de France, Campus de Lille 3, Domaine Universitaire du Pont de Bois, B.P. 60149, 59653 Villeneuve d'Ascq Cedex, France

[b] Department of Neuroscience, Psychology and Behaviour, University of Leicester. Centre for Medicine, Lancaster Road, Leicester, LE1 7HA, UK

ARTICLE INFO

Article history:
Available online 30 September 2016

Keywords:
Backward conditioning
Conditioned reinforcement hypothesis
Delay reduction theory
Humans
Signal hypothesis

ABSTRACT

In the present study, excitatory backward conditioning was assessed in a conditioned reinforcement paradigm. The experiment was conducted with human subjects and consisted of five conditions. In all conditions, US reinforcing value (i.e. time reduction of a timer) was assessed in phase 1 using a concurrent FR schedule, with one response key leading to US presentation and the other key leading to no-US. In phase 2, two discrete stimuli, S+ and S−, were paired with US and no-US respectively using an operant contingency. For three groups, backward contingencies were arranged, and two of these were designed to rule out a trace (forward) conditioning interpretation of the results. The two other groups served as control conditions (forward and neutral conditions). Finally, in phase 3 for all groups the CSs were delivered in a concurrent FR schedule similar to phase 1, but with no US. Responding during phase 3 showed conditioned reinforcement effects and hence excitatory backward conditioning. Implications of the results for conditioned reinforcement models are discussed.

1. Introduction

Conditioned reinforcement is an old concept in the study of animal and human behavior. Its core idea is that an initially neutral stimulus (NS), because of its pairing with a primary reinforcer, acquires the capacity to serve as an effective reinforcer. Here, by reinforcement we mean the increase in the frequency of an operant behavior by the contingent presentation of a stimulus. Early evidence of conditioned reinforcement was observed in studies conducted in chimpanzees by Wolfe (1936) and Cowles (1937), with the use of tokens as conditioned reinforcers, or in a study conducted in rats by Bugelski (1938) studying resistance to extinction. But perhaps the most representative demonstration of conditioned reinforcement was the study conducted by Skinner in 1938, with the new response procedure. In Skinner's experiment, the sound of a pellet dispenser was first paired with the delivery of food without the requirement of a response by the rats (i.e. stimulus–stimulus pairing). During a second phase, a lever was introduced in the chamber and pressing the lever (i.e. the new response) produced the sound of the pellet dispenser (without food delivery). Evidence of conditioned reinforcement was shown by an increase in lever press frequency with the contingent delivery of the sound.

Multiple procedures have been designed since the discovery of conditioned reinforcement, and these procedures can be broadly divided in two categories (Williams, 1994). In the first category, the conditioned reinforcer is isolated from the

* Corresponding author.
E-mail address: arthur.prevel@etu.univ-lille3.fr (A. Prével).

http://dx.doi.org/10.1016/j.lmot.2016.09.004

primary reinforcer after the initial pairings and presented contingent on some behavior. The most significant example is the new-response procedure proposed by Skinner (1938) cited above (see also Sosa, dos Santos, & Flores, 2011). Another example is the resistance to extinction procedure (Bugelski, 1938; Urushihara, 2004), where resistance to extinction of an operant response is increased by the contingent presentation of a conditioned reinforcer. This effect is, for example, demonstrated with a comparison group where the operant response is not followed by the primary or the conditioned reinforcer, and where the resistance to extinction is reduced.

In the second category designed for the study of conditioned reinforcement, the conditioned stimulus (CS) is also made contingent on an operant response but the pairings with the primary reinforcer are maintained to avoid Pavlovian extinction. Well known procedures developed in this category are the chain and concurrent-chains schedules (Fantino, 1977; Kelleher & Gollub, 1962). In a concurrent-chains schedule, two concurrent initial-link schedules (e.g. VI 120 conc. VI 120) produce the transition to mutually exclusive terminal-link schedules (e.g. VI 30 and VI 90) producing the primary reinforcer. Transition from initial-link to terminal-link is signaled by different stimuli (e.g. red light and green light) and these stimuli are assumed to develop conditioned reinforcing properties. A second well-known procedure is the observing response procedure (Dinsmoor, 1983; Shahan & Podlesnik, 2005, 2008). In this procedure, an un-signaled reinforcement schedule delivering food alternates with extinction (i.e. a mixed schedule) on one response key. On a second response key (i.e. the observing key), a brief stimulus presentation is associated with the reinforcement schedule (S+) and a second stimulus is associated with the extinction schedule (S−). These stimuli are produced by pressing on the observing key, and responding on the observing key is supposed to be maintained by the conditioned reinforcement properties of S+. Finally, a last procedure known in this second category is the token procedure (Hackenberg, 2009), where tokens are earned and exchanged for accesses to primary reinforcers. Here tokens are supposed to act as conditioned reinforcers.

Most of the experiments on conditioned reinforcement using the procedures described above have been conducted in rats and pigeons. However, there have also been reports of conditioned reinforcement in human subjects. For example, evidence of conditioned reinforcement in a free operant situation was found with psychiatric patients (Levin & Sterner, 1966) and with children (Myers & Myers, 1962), and numerous papers have reported the effect of human attention in the increase of appropriate behaviors (Hall, Lund, & Jackson, 1968; Jones, Drew, & Weber, 2000; Northup, Broussard, Jones, & Herring, 1995). Furthermore, chain schedules have been implemented in children (Long, 1963), and concurrent chain schedules have been studied in adults using a video-game (Leung, 1989, 1993). More recently, the observing response procedure (Fantino & Silberberg, 2010) was studied using a video-game. Finally, studies have also reported conditioned reinforcement effects by tokens in a token economies procedure (Kazdin, 1977).

As discussed above, early evidence of conditioned reinforcement was found in the first half of the 20th century, and a large literature has evolved. But the mechanisms underlying conditioned reinforcement are still debated in the literature (Shahan, 2010), and there is no consensus on what drives conditioned reinforcement. In the following sections, we will review some of the most important hypotheses on conditioned reinforcement.

One of the most influential hypotheses is the conditioned reinforcement hypothesis, or CRH (Dinsmoor, 1983; Skinner, 1938, 1953). CRH was developed in a time dominated by S-R learning theories (Hull, 1943; Spence, 1950) and hence was influenced by this framework. Its core idea is that a neutral stimulus (NS) will develop its own reinforcing value because of its pairing with a primary appetitive stimulus. In other words, a stimulus will develop the capacity of strengthening a stimulus-response association because of previous stimulus–stimulus pairings with the primary reinforcer. Conditioned reinforcement was an influential concept for theorists working in the S-R framework because it permitted them to translate the results from labs to natural situations. Different forms of the CRH were developed (see for example Kelleher & Gollub, 1962; for a review), but the core idea of CRH is that the temporal contiguity between the conditioned stimulus (CS) and the appetitive unconditioned stimulus (US) is an important variable in the development of appetitive value by the CS, even a necessary and sufficient condition (Skinner, 1953). The influence of CS-US delay was demonstrated for example by Bersh (1951) and Jenkins (1950), who showed that the number of lever presses made before CS delivery was reduced with an increased delay between the CS and US. CRH has more recently been supported by Donahoe and Palmer (2004) and Donahoe (2014).

A second well-known hypothesis on conditioned reinforcement is the Delay Reduction Theory, or DRT (Fantino, 2008; Fantino, Preston, & Dunn, 1993; Preston & Fantino, 1991). DRT was originally developed to explain choices in concurrent-chain schedules of reinforcement and the influence of the stimuli signaling the transition from initial to terminal links on response allocation. Although multiple forms of DRT have been developed, its core idea is that the effectiveness of a stimulus as a conditioned reinforcer may be predicted by its reduction in the length of time to primary reinforcement, measured from the onset of the conditioned reinforcer. In its simplest form, DRT may be stated by:

Reinforcement strength of stimulus A

$$f\left(\frac{T-t_A}{T}\right) \quad (1)$$

where T is the averaged time between primary reinforcer presentations and t_A is the time between the conditioned reinforcer and primary reinforcer onset. So, DRT assumes that the more a conditioned reinforcer is correlated with reduction in waiting time to reinforcement, the more it will develop reinforcing properties. This effect was for example demonstrated in an experiment by Fantino (1969), where he found a large preference in a concurrent chain schedule paradigm for a VI 90 VI 30 schedule over a VI 30 VI 90 schedule. Finally, we may note the similarity between DRT and some models of conditioning,