Learning and Motivation 47 (2014) 1–11

Contents lists available at ScienceDirect

Learning and Motivation

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

Overexpectation in the context of reward timing

Chad M. Ruprecht, Haydee S. Izurieta, Joshua E. Wolf, Kenneth J. Leising*

Department of Psychology, Texas Christian University, United States

ARTICLE INFO

Article history:
Received in revised form 27 January 2014
Accepted 29 January 2014
Available online 1 March 2014

Keywords:
Overexpectation
Timing
Pavlovian conditioning
Rescorla–Wagner model
Temporal difference models

ABSTRACT

One of the many effects predicted by the Rescorla–Wagner model is overexpectation (OX). The OX effect is the finding that following compound training with two asymptotic elements, X and A, animals emit less conditioned responding (CR, e.g., nose poking) during tests of X alone compared to animals that did not receive compound training. We investigated the OX effect in the context of reward timing by training rats to expect sucrose at different times during X and recording the CR throughout the duration of X. Experiment 1 examined the OX effect using a traditional delayed conditioning procedure. In Experiment 2, the period during which sucrose was expected occurred either early or late during X. Tests revealed that less CR occurred in the OX group around the period that sucrose was previously overexpected, and was otherwise similar in response functions to the control group that did not receive the compound manipulation. These are the first studies pitting the effects of OX with an animal's ability to time their expectation of food.

Previous research has shown that the number and quality (e.g., temporal proximity) of pairings of an unconditioned stimulus (US; e.g., sucrose) with an initially neutral, conditioned stimulus (CS; e.g., a tone) influences the magnitude and timing of the conditioned response (CR; e.g., nose poking for food). The Rescorla–Wagner (R-W) model of learning (Rescorla & Wagner, 1972) proposed an equation to describe how a CS comes to control the CR by couching learning as trial by trial alterations to the associative value of a CS. The R-W model notably accounted for many existing conditioning effects (e.g., blocking), and also anticipated, a priori, a variety of conditioning effects that rely on the summation of associative values from more than one CS (e.g., superconditioning and overexpectation) One such effect occurs when the combination of previously trained CSs (hereafter referred to as elements) results in an overexpectation (hereafter referred to as OX) of the US. In Phase 1 of an OX procedure, two elements are trained on separate trials with a common US to asymptotic levels of responding. In Phase 2, the two elements are presented in compound. Given that initial trials of X and A occurred separately, the R-W model posits a summation rule: when X and A are placed in compound their associative values sum together. It is on the initial compound trial that animals should maximally overexpect the single US. Because of the finite amount of learning that can occur to a US, the model predicts that the associative value of each element should drop during subsequent compound trials until each element predicts the appropriate amount of US (i.e., X and A each equal half of the original US value).

The performance of animals trained with an OX procedure is most often compared with a CTL group that receives additional trials with one of the pre-trained elements in place of compound training (e.g., Kehoe & White, 2004; McNally, Pigg, & Weidemann, 2004; Rescorla, 1970, 1999; Sissons & Miller, 2009). Only in group OX should the prediction error lead to an adjustment of the associative value of X and A. The OX effect has been substantiated in the behavior of rats (Kamin & Gaioni, 1974; Kremer, 1978; Lattal & Nakajima, 1998; Rescorla, 1970), pigeons (Khallad & Moore, 1996) and recently, humans (Collins & Shanks, 2006).

* Corresponding author at: Department of Psychology (298920), Texas Christian University, 2800 South University Drive, Fort Worth, TX 76129, United States.
E-mail address: k.j.leising@tcu.edu (K.J. Leising).

0023-9690/$ – see front matter © 2014 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.lmot.2014.01.004

The RW-model makes predictions about the magnitude of the CR following the OX procedure. The model treats a stimulus, regardless of its duration or the timing of reward delivery, as a single element that accrues or loses associative value as a whole unit until the US prediction error is minimized. This assumption predicts the OX effect should result in a diminished CR uniformly throughout the duration of an element and independent of the CS–US temporal relationship. However, there is no question that timing of crucial conditioning events, such as the CS–US interval, influences the magnitude, form, and distribution of the conditioned response during the presentation of an element (e.g., Catania, 1970; Gibbon, Malapani, Dale, & Gallistel, 1997; Kirkpatrick & Church, 1998, 2000; Roberts, 1981).

A temporally specific CR is typically evaluated using post-training trials in which the US is omitted and the temporal distribution of responding is analyzed. For example, Kirkpatrick and Church (1998, Experiment 2) found that rats trained with a 15-s CS–US interval increased the rate of nose poking within 1–2 s after CS onset but peaked at approximately 15 s after CS onset. While the original R-W model uses prediction error to modify the associative value of a stimulus as a unified whole, temporal difference (TD) models assume that each moment (time steps) during a stimulus is distinctly represented (Ludvig, Sutton, & Kehoe, 2008; Sutton & Barto, 1981; Vogel, Brandon, & Wagner, 2003). The pattern of phasic firing by reward processing dopamine neurons is suggested by TD models to encode prediction error at each moment during a stimulus based on the difference between the discounted value of the US predicted at the current time step and the predicted cumulative sum of discounted US value from the remaining time steps. Unlike the R-W model, TD models assume that US prediction strength (i.e., associative strength) varies throughout the duration of the stimulus. This assumption correctly predicts the response peaks observed at the expected time of the US delivery during conditioning studies that vary the CS–US interval. Furthermore, it suggests that moments of prediction error, such as during an OX procedure, may be isolated to distinct time steps within the duration of a stimulus. Temporally specific adjustments in the associative strength of the CS due to prediction error should result in a temporally specific reduction in the magnitude of the CR.

Blaisdell, Denniston, and Miller (2001, Experiment 4) demonstrated the best evidence for temporal relationships between X, A, and the US modulating the strength of the OX effect. Rats were initially trained with a 5-s trace interval between each element (X and A) and the US in a fear conditioning paradigm. The same interval was used in Phase 2 when the elements were presented together in compound training. Rats were then given post-training presentations of A either terminating immediately with the US (Group OX-Diff) or with the US following the same 5-s trace interval from training (Group OX-Same). When tested with X, Group OX-Diff showed a larger CR (i.e., more conditioned suppression) than OX-Same. Blaisdell et al. claimed that A more effectively competed with X when the temporal relations during post-training matched element and compound training. Time was an important factor in OX, but due to their design, Blaisdell et al. were not able to demonstrate whether the decrement due to OX induced a temporally specific drop in the CR. Temporal specificity can be more directly analyzed by measuring fine grained changes in the time course of the CR (e.g., Leising, Sawa, & Blaisdell, 2007; Williams, Johns, & Bindras, 2008).

In the current experiments, we trained rats to nose poke for food and examined whether the OX effect could manifest at a temporally specific time period (Fig. 1). In Experiment 1, we trained two forward-paired elements that differed in duration (40 s vs. 10 s), mirroring the design of recent OX procedures with forward-paired elements (e.g., Rescorla, 2006, 2007; Sissons & Miller, 2009). In Experiment 2, we modified an embedded procedure utilized by Leising et al. (2007) to train rats to expect sucrose either early or late within X and then embedded A directly into these time periods. Following a retraining procedure to further enhance responding to the shorter A (and enhance the decrement during X), we found evidence for timing of the

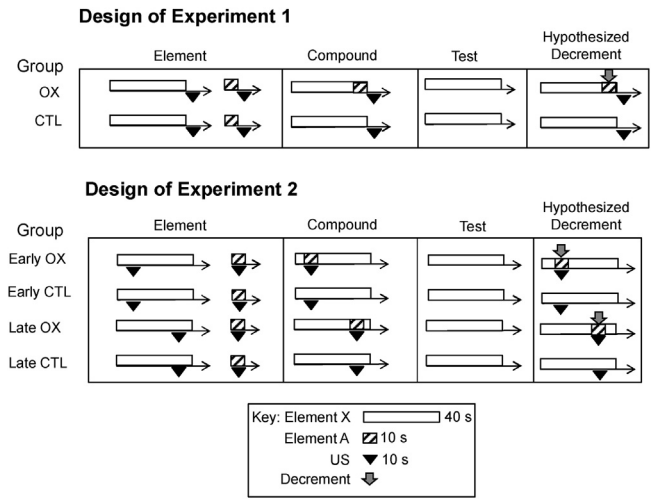

Fig. 1. The design of Experiments 1 and 2. Element training has been collapsed across Phases 1 and 2 and trials of Element B have been excluded for simplicity (see Tables 1 and 2 for specific details).