Learning and Motivation 51 (2015) 50–61

Contents lists available at ScienceDirect

Learning and Motivation

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

Choice behavior under differential outcomes: Sample stimulus control versus expectancy control

John M. Holden[a,*], J. Bruce Overmier[b]

[a] Winona State University, Phelps Hall 231, Winona, MN 55987, United States
[b] University of Minnesota, 75 East River Road, Minneapolis, MN 55455, United States

ARTICLE INFO

Article history:
Received 23 December 2014
Received in revised form 23 April 2015
Available online 16 May 2015

Keywords:
Differential outcomes
Matching-to-sample
Expectancies
Stimulus control
Pigeons

ABSTRACT

It has been hypothesized that in the differential outcomes (DO) procedure, each discriminative stimulus comes to evoke a specific expectancy or representation of its unique reward and that expectancy exerts "stimulus" control over choice behavior in the same manner as a discriminative stimulus. It has also been suggested that expectancy control may reduce—or even replace—control by the discriminative stimulus under DO, and that under DO, subjects may show considerably more ability to choose a correct response on the basis of expectancy than on the basis of the discriminative stimulus. The present experiment, using DO in the delayed matching-to-sample procedure with pigeons, demonstrates that subjects under DO show more ability to match based on expectancy alone than sample alone when direct comparisons are made, and when presented with two choices, one indicated by sample and one indicated by expectancy, subjects typically made the choice indicated by expectancy. The implications of these findings of greater control of choice behavior by expectancies for traditional accounts of behavioral control in choice behavior are discussed.

Published by Elsevier Inc.

Differential outcomes (DO) refers to the procedure in which separate correct stimulus-response sequences are rewarded with different and distinct outcomes. Under DO, acquisition of that task is typically found to be significantly accelerated relative to tasks when all sequences are rewarded with the same reinforcer (*common outcomes* or CO) or with random reinforcers (*nondifferential outcomes* or NDO) (DeMarse & Urcuioli, 1993; Hochhalter, Sweeney, Bakke, Holub, & Overmier, 2000; Kruse, Overmier, Konz, & Rokke, 1983; Peterson & Trapold, 1980; Savage & Langlais, 1995; Trapold, 1970; Williams, Butler, & Overmier, 1990). Notable DO effects have even been shown to occur in procedures in which one of the "reinforcing" outcomes is only secondarily reinforcing—for example, where one of the outcomes is food and the second outcome is a neutral event such as a tone or light (Peterson, Wheeler, & Trapold, 1980; Zentall & Sherburne, 1994). Performance across delays is also enhanced under DO (e.g., Brodigan & Peterson, 1976), as is resistance to memory disrupting drugs (Savage, Stanchfield, & Overmier, 1994). This DO approach has shown great promise in enhancing instrumental performance of both human adults (Esteban, Plaza, López-Crespo, Vivas, & Estévez, 2014; Hochhalter et al., 2000; Martella, Plaza, Estévez, Castillo, & Fuentes, 2012; Miller, Waugh, & Chambers, 2002; Plaza, López-Crespo, Antúnez, Fuentes, & Estevez, 2012) and children (Esteban et al., 2014; López-Crespo, Daza, & Méndez-López, 2012).

The source of this enhancement is worth exploring further. Some authors have assumed that presentation of unique reinforcers for each stimulus-response sequence is conducive to the development of a specific *expectancy* of a given outcome.

* Corresponding author.
E-mail addresses: jholden@winona.edu (J.M. Holden), psyjbo@umn.edu (J.B. Overmier).

http://dx.doi.org/10.1016/j.lmot.2015.04.002
0023-9690/Published by Elsevier Inc.

Trapold and Overmier (1972) argued that expectancies themselves have interoceptive cue properties that can control overt behavior according to the same rules by which exteroceptive discriminative stimuli acquire such control. Classically, overt discriminative stimuli are thought to acquire control over particular response by virtue of those responses being reinforced in the presence of those stimuli, as is the case in the classic stimulus-response (S-R) theory of Thorndike (1898). Under DO, each discriminative stimulus is capable of evoking an expectancy of the reinforcer with which it is uniquely associated. Each expectancy also acquires control over the same response as the stimulus that evoked that expectancy. Expectancy provides an extra source of stimulus control over choice behavior and thereby enhance correct responding above and beyond what would be expected based on control based on the sample alone.

The view that reinforcer-specific expectancies have discriminative, behavior-controlling properties is supported by studies that show that reversing or eliminating DO following initial training causes performance to deteriorate, even when no other aspects of the task are changed (Honig, Matheson, & Dodd, 1984; Peterson & Trapold, 1980; Peterson, Wheeler, & Armstrong, 1978), and by findings that matching-to-sample (MTS) performance under DO readily transfers to novel sample stimuli that have been differentially associated off-baseline with the same outcomes as the samples used in training or to comparison stimuli trained in a separate MTS task (Kruse et al., 1983; Urcuioli, DeMarse, & Lionello, 1998). Some investigators have questioned whether expectancies merely provide redundant stimulus control that is additive to control provided by the discriminative stimuli. Using DO in delayed matching-to-sample (DMTS) tasks, for example, may not merely provide redundant stimulus control but may also affect discriminative control exerted by the sample stimuli themselves, by reducing sample stimulus control despite the overall enhancement of performance. That is, the expectancy as a source of control reduces the subjects' reliance on or ability to use the sample itself as a source of control so that, if they are required to use the sample to choose, their ability to do so is less than it would be under a NDO procedure. Peterson and Trapold (1980) and Urcuioli (1990) demonstrated that when the testing situation is altered such that subjects trained under DO can no longer utilize expectancy as a cue, their performance drops below that of subjects trained comparably under NDO. Furthermore, Urcuioli (1991) found that training under DO retards acquisition of matching relative to NDO if the expectancies evoked are unreliable cues for choice (that is, when differential outcomes were associated with samples but uncorrelated with comparison alternatives). These findings are consistent with the idea that sample stimulus control is in fact reduced/overshadowed by the specific expectancies evoked by those discriminative stimuli.

It may be that the memory strategies employed under NDO or common outcomes (where the same reinforcer is used for both correct sample-comparison sequences), which is of necessity a retrospective strategy where memory of the sample is used to guide choice behavior, is incompatible with a prospective strategy where a specific expectancy of the reinforcer to be delivered guides choice behavior (as seems to be the case in DO). Thus, control over behavior by expectancy is not merely additive to the control by sample memory, but rather supplants it to some extent. In line with this possibility, Holden and Overmier (2014) demonstrated that subjects in a delayed matching-to-sample task (DMTS) trained under DO were unable to employ retrospective memory of the sample to distinguish between a comparison stimulus that had been extensively trained with that sample and a separate comparison stimulus that, while associated with the same outcome as the sample, had never been trained instrumentally with that sample. Such a possibility has also been suggested by recent fMRI studies of memory (Mok, Thomas, Lungu, & Overmier, 2009).

While work from our laboratory has previously suggested that the presence of an outcome expectancy reduces sample memory under DO, it is unclear to what extent. The present work is designed to make a direct comparison of sample and expectancy control within the same procedure. We attempted to pit sample-control and expectancy-control against each other in a complex experiment in which correct choice was sometimes dependent on sample and sometimes dependent on expectancy. Under some conditions, available choices were arranged with the intent that the sample and its evoked expectancy would guide subject behavior in opposite directions on the same trial. We hypothesize generally that expectancy control is stronger than sample control under DO. Perhaps most important, when given a choice between two responses, one indicated by sample and the other indicated by expectancy, subjects should make the choice indicated by expectancy. We explored this choice across a set of delay intervals commonly employed in studies of short-term memory in pigeons, to determine whether the influence of sample and expectancy varied across different delays.

1. Materials and methods

1.1. General design

The general design of the present experiment is shown in Fig. 1. Subjects were first trained in two DMTS tasks, the first of which was established using DO and the second of which was established using NDO. Subjects were then trained in a new DMTS task which utilized the same sample stimuli used in the NDO task. However, the new task was a DO task, involving training with two new comparison stimuli. The sample stimuli from this nondifferential task were trained exclusively with one or the other outcome in the new task. As a result, these samples should, according to expectancy theory, come to evoke a specific expectancy of reinforcement, which may or may not reduce control by the samples themselves.

Once subjects reached criterion on this new task, they were tested under three different probe conditions, using probe trials intermixed with regular trials identical to those from the new task. At the beginning of each probe trial, one of the sample stimuli was presented. After the appropriate delay period (0, 2, 4, or 8 s), subjects were presented with one of the following pairs of comparison stimuli: (a) the comparison stimuli from the DO task, including the comparison stimuli that were trained