Behavioural Processes 129 (2016) 44–53

Contents lists available at ScienceDirect

Behavioural Processes

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

Increasing the persistence of a heterogeneous behavior chain: Studies of extinction in a rat model of search behavior of working dogs

Eric A. Thrailkill[a,*], Alex Kacelnik[b], Fay Porritt[c], Mark E. Bouton[a]

[a] University of Vermont, United States
[b] University of Oxford and Oxford Risk, UK
[c] Dstl, Fort Halstead, S18, Sevenoaks, Kent TN14 7BP, UK

ARTICLE INFO

Article history:
Received 6 January 2016
Received in revised form 25 May 2016
Accepted 31 May 2016
Available online 13 June 2016

Keywords:
Persistence
Extinction
Heterogeneous chains
Detection
Canine
Rat

ABSTRACT

Dogs trained to search for contraband perform a chain of behavior in which they first search for a target and then make a separate response that indicates to the trainer that they have found one. The dogs often conduct multiple searches without encountering a target and receiving the reinforcer (i.e., no contraband is present). Understanding extinction (i.e., the decline in work rate when reinforcers are no longer encountered) may assist in training dogs to work in conditions where targets are rare. We therefore trained rats on a search-target behavior chain modeled on the search behavior of working dogs. A discriminative stimulus signaled that a search response (e.g., chain pull) led to a second stimulus that set the occasion for a target response (e.g., lever press) that was reinforced by a food pellet. In Experiment 1 training with longer search durations and intermittent (partial) reinforcement of searching (i.e. some trials had no target present) both led to more persistent search responding in extinction. The loss of search behavior in extinction was primarily dependent on the number of non-reinforced searches rather than time searching without reinforcement. In Experiments 2 and 3, delivery of non-contingent reinforcers during extinction increased search persistence provided they had also been presented during training. Thus, results with rats suggest that the persistence of working dog performance (or chained behavior generally) may be improved by training with partial reinforcement of searching and non-contingent reinforcement during both training and work (extinction).

1. Introduction

Learned behavior often takes place in the form of separate but linked sequences or *chains* of behavior (e.g., Skinner, 1934, 1938). A behavior chain minimally consists of one response that provides the opportunity to perform another response that leads to a reinforcer. Each response is occasioned by a unique discriminative stimulus (SD). In a simple heterogeneous behavior chain, an SD sets the occasion for the first response that leads to the presentation of a second SD, which sets the occasion for a second response to be reinforced and serves as a conditioned reinforcer for the first response. Since many behaviors occur in a chain, a comprehensive understanding of learned behavior must include an understanding of behavior chains.

Understanding behavior chains in the laboratory may be useful for addressing behavior in applied settings. We have recently proposed that heterogeneous behavior chains studied in the laboratory describe behaviors analogous to those involved in drug or junk-food procurement and consumption (Thrailkill and Bouton, 2015a,b). In that case, our goal was to reduce undesirable behaviors that take place in a chain. However, the analysis of behavior chains may also be useful for characterizing variables that promote *desirable* behaviors that take place in a chain. For instance, working search dogs are trained to perform a behavioral chain with the goal of detecting explosives or other contraband (see Helton, 2009). A period of search responding (e.g., sniffing) in the presence of an SD (e.g., a car) eventually leads to a target SD (e.g., the odor of an explosive) which signals that a target response (e.g., pointing) will be reinforced. Fortunately perhaps, most working dogs will encounter prolonged periods in which searching does not result in discovery of the target odor. Therefore, a great deal of search behavior will never yield a reinforcer and responding could undergo extinction (Porritt et al., 2015).

* Corresponding author at: University of Vermont, Department of Psychological Science, 2 Colchester Ave. Burlington, VT 05405, United States.
E-mail addresses: eathrail@uvm.edu (E.A. Thrailkill), mbouton@uvm.edu (M.E. Bouton).

http://dx.doi.org/10.1016/j.beproc.2016.05.009
0376-6357/

In order to address this paucity of targets in some working environments, dog trainers commonly extend the period of time between target detections in the initial training phases (e.g. Garner et al., 2001; Goldblatt et al., 2009). However, training time is limited, particularly once dogs are operational and there is a trade-off between extending the length of searches and ensuring that dogs get sufficient reinforcement on all targets during training days. Thus, a better understanding of search extinction and variables that influence search persistence would benefit users and enable them to optimize their use of limited training time.

Extinction generally refers to a procedure in which the reinforcer delivered during response acquisition is removed (Pavlov, 1927). Extinction learning is evidenced by a gradual decline in the response with repeated nonreinforced performance. Pavlovian and instrumental extinction have been a topic of intense interest in experimental and applied settings (e.g., Conklin and Tiffany, 2002; see Vurbic and Bouton, 2014 for a review). Although resistance to extinction has been studied in simple free operant and discrete-trial operant procedures (Amsel, 1967; Capaldi, 1966; Vurbic and Bouton, 2014), there has been no systematic investigation of the extinction of behavior chains. Laboratory studies of both Pavlovian and discrete operant behaviors have identified the crucial role of generalization between the conditions of reinforced training and extinction in creating resistance to extinction (Capaldi, 1966, 1994). Methods that increase the similarity between the conditions of training and extinction decrease the speed of the decline in responding during extinction by enhancing generalization between the two contexts; two such major methods are known to be effective.

First, extinction is slower when training involves reinforcement of only some responses (intermittent or partial reinforcement; PRF) as opposed to every response (continuous reinforcement; CRF). This partial-reinforcement extinction effect (PREE) is a well-known and widely-replicated property of extinction learning; it has been observed in a wide range of species, including dogs (e.g., Mackintosh, 1974; Feuerbacher and Wynne, 2011). The PREE has been demonstrated with responses learned in Pavlovian, free operant, discriminated operant, and discrete-trial operant paradigms. The PREE is thought to occur, at least in part, because introducing occasions when the response is not reinforced during training makes the conditions of extinction more similar to those of training (Capaldi, 1994). Responding persists in extinction, a condition in which all responses are nonreinforced, because the organism learned to respond after recent occasions in which the response was not reinforced.

A second method for slowing extinction is to present the reinforcer used in training independently of the response during extinction (i.e. noncontingent reinforcement). Suspending the response-reinforcer contingency, but continuing to deliver the reinforcer, results in a slower decline compared to removing the reinforcer entirely (e.g., Baker, 1990; Rescorla and Skucy, 1969; Winterbauer and Bouton, 2011). One hypothesis is that reinforcers acquire a discriminative function for the response in addition to strengthening the response. That is, during training, the animal is reinforced for making responses soon after receiving a reinforcer. Noncontingent presentations of the reinforcer slow down extinction because the reinforcer presentations continue to set the occasion for the response. Both partial reinforcement and noncontingent reinforcers can slow extinction because they functionally increase the similarity, or generalizability, between training and extinction conditions.

Although the effects of partial reinforcement and noncontingent reinforcers on extinction have been widely studied in simple operant and Pavlovian procedures, to our knowledge there have been no studies of how these variables affect extinction of behavior trained in a heterogeneous behavior chain. We therefore investigated them here. Recent work by two of us has developed a procedure in which rats learn to make a discriminated heterogeneous behavior chain (Thrailkill and Bouton, 2015a,b). In the method, rats learn to make a search response (e.g., pulling a chain) in the presence of a search SD that leads to the presentation of a target SD that sets the occasion for a target response (e.g., pressing a lever) that is then reinforced with a food pellet. A search stimulus first sets the occasion for a search response, and a target stimulus then sets the occasion for the target response (while potentially also serving as a conditioned reinforcer for search). (Our previous descriptions of the method labeled the first response "procurement" and the second response "consumption," but the "search" and "target" labels used here are better suited for the working-dog application).

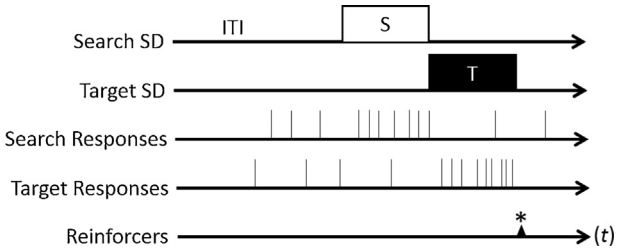

Fig. 1. Diagram of the discriminated heterogeneous behavior chain procedure. Search and target responses can occur freely. After an inter-trial interval (ITI), the Search SD (S) turns on. Search responses during the Search SD can produce the simultaneous offset of the Search SD and onset of the Target SD (T). Target responses during the Target SD can produce the simultaneous offset of the Target SD, presentation of the reinforcer (*), and initiate the next ITI (*t* stands for time.).

The present experiments used the procedure to investigate the contributions of partial reinforcement and noncontingent reinforcement to extinction of search responding. Fig. 1 shows a diagram of the events in the procedure. In our first experiment, we used a factorial design to compare extinction of search responding after training with continuous versus partial reinforcement (sometimes search responding ended the search SD without an opportunity to make the target response), and long versus short search stimulus durations (longer search durations yielded more search responses that were not reinforced). In a second experiment, we then asked how noncontingent pellet deliveries during extinction would influence the persistence of searching in extinction. The results suggest that each of the variables investigated can in fact increase the persistence of search responding in extinction.

2. Experiment 1

In Experiment 1, we studied the contributions of partial reinforcement of search and of search duration as factors influencing extinction in a 2 × 2 factorial design. Rats learned a discriminated heterogeneous behavior chain consisting of presentations of a search SD (panel light located adjacent to the response manipulandum) that set the occasion for a search response (e.g., chain pull) which, according to a variable-interval (VI) schedule, could lead to a target SD that set the occasion for a target response (e.g., lever press) being reinforced. The rats received 30 such trials every training day. One factor in the design was the VI schedule used to reinforce the search response. Different groups could earn access to the target SD (and an opportunity to receive reinforcement of the target response) according to either VI-10 s ("Short" groups) or VI-30s ("Long" groups) schedule. Notice that the "Long" interval potentially provided the opportunity for more search responses that did not produce the target SD. The two groups were further divided into two additional groups that differed in the proportion of search trials that actually led to the target SD. For the "CRF groups" (continuous reinforcement), 30 out of 30 trials led to the target SD; for the "PRF groups" (partial reinforcement), 10 ran-