



Published in final edited form as:

Neurobiol Learn Mem. 2016 September ; 133: 61–68. doi:10.1016/j.nlm.2016.06.005.

Extinction and the associative structure of heterogeneous instrumental chains

Eric A. Thrailkill and Mark E. Bouton

University of Vermont

Abstract

Drug abuse, overeating, and smoking are all examples of instrumental behaviors that often involve chains or sequences of behavior. A behavior chain is minimally composed of a procurement response that is required in order for a subsequent consumption response to be reinforced. Despite the translational importance of behavior chains, few studies have attempted to understand what binds them together and takes them apart. This article surveys the development of the heterogeneous instrumental chain method and introduces recent findings that have used extinction to analyze the associative content of (what is learned in) the chain. Chained responses that are occasion-set by their own discriminative stimuli may be directly associated; extinction of the procurement response weakens its associated consumption response, and extinction of the consumption response weakens its associated procurement response. Extinction itself involves learning to inhibit the response. Extinguished chained responses are subject to renewal when they are tested either back in the acquisition context or in a new context. In addition, a consumption response that is extinguished outside its chain is renewed when returned to the context of the preceding response in the chain. Research on heterogeneous behavior chains can provide important insights into an important but often overlooked aspect of instrumental learning.

Keywords

Heterogeneous instrumental chains; behavior sequence; extinction; context; renewal; response inhibition

Individuals who overeat, abuse drugs, or smoke generally perform a series or sequence of behaviors in a chain. For example, the overeater purchases and then eats junk food, the iv drug user must find drugs before injecting them, and the smoker must purchase cigarettes before he or she can smoke them. Each behavior in the chain has a different topography (the responses are *heterogeneous*) and each takes place in its own distinctive discriminative stimulus (SD; the responses are *discriminated*). For example, the smoker might buy his cigarettes in a mini-mart and then light up later outside. Behavior chains minimally involve a response that results directly in the reinforcer (a *consumption* response) and at least one more response that provides access to the consumption response (a *procurement* response; Collier, 1981; see Thrailkill & Bouton, 2015a). Because instrumental behaviors arguably

always occur in a chain (e.g., Pryor, 1984; Skinner, 1934), understanding the variables that influence each component may have important theoretical and translational value (e.g., Ostlund & Balleine, 2009).

A translational perspective on instrumental behavior chains suggests that we need to understand what glues them together and takes them apart. However, work with heterogeneous instrumental chains has been limited to date. Some of this research has revealed and dissociated the motivational processes controlling the first and second responses in the chain. Other work has focused on the pharmacology of behavior chains when drugs are used as reinforcers. More recently, our own laboratory has investigated the associative structure underlying such a discriminated heterogeneous chain, as well as the contextual control of behaviors occurring in such chains. Although the associative analysis of chained instrumental behaviors is relatively new, the field is already beginning to understand some important dimensions of them.

Researchers have often used different terms to refer to the first and second responses in two-component chains (e.g., “R1–R2,” Balleine, Garner, Gonzalez, & Dickinson, 1995; “Seeking-Taking,” Olmstead, Lafond, Everitt, & Dickinson, 2001; “Distal-Proximal,” Corbit & Balleine, 2003). We have chosen to adopt language suggested by Collier (1981), who described the first and second responses as “procurement” and “consumption.” These terms invoke the function of the behaviors in relation to the reinforcer (Thrailkill & Bouton, 2015a), and map most directly onto examples of overeating, drug taking, and smoking like the ones described above. Because of our interest in potentially separable procurement and consumption behaviors, results of procedures deliberately designed to study unitized or “chunked” strings of responses (e.g., Greybiel, 1998; Ostlund, Winterbauer, & Balleine, 2009) will not be discussed in detail in the present review.

Motivational control of responding in heterogeneous instrumental chains

The work of Balleine and Dickinson and their colleagues on instrumental chains has mainly been concerned with the motivational processes that influence them (Balleine et al., 1995; Balleine, Paredes-Olay, & Dickinson, 2005). Their experiments on chains began with the study of *incentive learning* (Balleine, 1992), the phenomenon which suggests that for organisms to adjust their instrumental behavior following a change in their internal motivational state (e.g., from satiety to hunger), they must experience the instrumental outcome in the changed motivational state. Incentive learning is well-documented in studies of simple instrumental behavior (Balleine, 1992; Dickinson & Balleine, 1994), but the phenomenon appears to be restricted to those behaviors in a chain that are remote from, or distal to, the primary reinforcer. Balleine (1992) found that rats required experience with the instrumental outcome in order for the motivational shift to influence lever pressing, but entries into the food magazine that followed lever pressing (which required rats to move a cover flap) were immediately changed in response to the motivational shift. Balleine et al. (1995) hypothesized that motivational shifts have different effects based on the position of the responses in a behavior chain. Balleine et al. (1995) developed a method in which two responses (levers) were concurrently available, but a reinforcer was delivered contingent on a consumption response only if a response on the other (procurement) lever had preceded it.

Although the reinforcer depended on emission of both a procurement and a consumption response, note that there were no SDs to indicate which response was required at particular moment. Rats that learned the chain while hungry immediately suppressed the second (consumption) response following a shift to a nondeprived state. However, no such effect occurred with the first (procurement) response unless the animal had received incentive learning (Balleine, 1992). That is, rats needed experience with the food in the changed (nondeprived) state to learn that it was not desired as a goal and adjust their responding on the first (procurement) lever. Noting the parallel between these results and those from studies of reinforcer devaluation in Pavlovian second-order conditioning (Rescorla, 1977), Balleine et al. (1995) suggested that the processes motivating the first and second responses in the chain are different and dissociable.

Corbit and Balleine (2003) further dissociated the motivational control of procurement and consumption behavior. In their method, a procurement lever was always present; an average of four responses on it (a Random Ratio 4 schedule, or RR 4) was required before the consumption lever was inserted into the chamber. Completing an RR 4 on the consumption lever then delivered a food-pellet reinforcer and retracted the consumption lever. Notice that, in contrast to Balleine et al.'s earlier experiments, the chain procedure was now partially discriminated: Insertion of the consumption lever was an SD for consumption responding and theoretically constituted a conditioned reinforcer that reinforced procurement responding (e.g., Gollub, 1977). With this procedure, Corbit and Balleine (2003) showed that when a food-associated conditioned stimulus (CS) was presented while the rats were responding on the chain, the CS excited consumption responding, but not procurement responding. In contrast, an incentive learning treatment [similar to that studied by Balleine et al. (1995)] influenced procurement responding, but not consumption responding. The motivational control of consumption and procurement was thus doubly dissociated: Consumption responding, but not procurement, was strongly influenced by Pavlovian incentive stimuli, and procurement responding, but not consumption, was influenced by instrumental incentive learning.

In related studies using the same procedure, Wassum, Ostlund, Balleine, and Maidment (2011a) found that systemic blockade of dopamine D₁/D₂ receptors reduced the influence of a CS on consumption responding, but had no effect on the sensitivity of procurement to incentive learning. In a complementary way, manipulation of μ -opioid receptors in the basolateral amygdala influenced the ability of incentive learning to affect procurement responding (Wassum, Cely, Balleine, & Maidment, 2011b; Wassum, Ostlund, Maidment, & Balleine, 2009). Finally, Johnson, Bannerman, Rawlins, Sprengel, and Good (2007) used a similar method (although with different RR schedules) and found that mice with genetic deletion of the GluR-1 AMPA receptor showed less excitement of consumption responding when the Pavlovian CS was presented. In sum, different motivational processes have been shown to affect the different responses in the chain, and research has begun to identify some of their neural substrates.

Heterogeneous behavior chains have also been used in a parallel line of research investigating the processes involved in drug-taking (Olmstead et al., 2001; Olmstead, Parkinson, Miles, Everitt, & Dickinson, 2000). In the method usually used, insertion of a

procurement lever signals the opportunity to make a procurement response. Satisfying a procurement response requirement [often a variable-interval (VI) schedule] then leads to insertion of a second (consumption) lever. A single response on the consumption lever then results in the reinforcer (e.g., intravenous delivery of cocaine), the presentation of drug-paired stimuli (e.g., lights signaling drug delivery), and the retraction of both levers. There is then an intertrial interval (ITI) in which neither response manipulandum is available. This procedure constitutes a fully discriminated heterogeneous chain, because insertion of each lever is an SD that signals the opportunity to make the corresponding procurement or consumption response. Notice, though, that although both responses are occasioned by separate SDs, there is no opportunity to observe the strength or rate of a response in the absence of its SD.

Studies using this method have uncovered several important variables that influence responding in the chain. For example, increasing the dose of cocaine increases the rate of procurement responding (Olmstead et al., 2000). And prolonged training of the chain leads to the procurement response becoming resistant to punishment produced by a response-contingent foot shock or a shock-associated CS (Chen et al. 2013; Vanderschuren & Everitt, 2004). According to these investigators, and others, the procurement response's resistance to punishment might model human addiction behaviors that have become "compulsive" (Economidou, Pelloux, Robbins, Dalley, & Everitt, 2009; Jonkman, Pelloux, & Everitt, 2012; Pelloux, Everitt, & Dickinson, 2007).

Related studies have examined whether procurement responding is *goal-directed*. Goal-directed actions are behaviors that are sensitive to a separate change in the value of the outcome, and under the control of a response-outcome association (see Dickinson, 1994, for review). For example, a single lever pressing response is said to be goal-directed if separately devaluing the reinforcer (e.g., by pairing it with a toxin) reduces the strength of the response when the response is tested in extinction. Olmstead et al. (2001) therefore investigated the sensitivity of procurement to a change in the value of a consumption response that was reinforced with intravenous cocaine. Following acquisition of the chain (using the method just described), one group of rats was allowed to perform the consumption response alone (the procurement lever was removed). At this point, making the consumption response did not produce cocaine or retract the consumption lever. A control group received a treatment in which consumption responding was likewise available, but it was reinforced with cocaine instead of extinguished. When tested with the procurement lever alone, the group that had received extinction of consumption made fewer procurement responses than the group for which consumption had been reinforced. The authors suggested that the rats' cocaine-maintained behavior was thus goal-directed in the sense that procurement was sensitive to the current value of consumption. Zapata, Minney, and Shippenberg (2010) went on to show that extinction of consumption did not affect procurement responding after extended training of the chain, consistent with the literature on the development of control by a stimulus-response association (habit) through overtraining (e.g., Dickinson, Balleine, Watt, Gonzalez, & Boakes, 1995; Thrailkill & Bouton, 2015b).

Leblanc, Ostlund, and Maidment (2012) found that Pavlovian and incentive motivational variables influence chained responses that lead to drug reinforcement. They used a chain

procedure that required rats to complete an RR schedule for procurement and consumption responses that were otherwise signaled by insertion of the corresponding levers. They found that presentation of a cocaine-associated CS invigorated consumption responding, but not procurement responding, when procurement and consumption responses were available simultaneously. The results were thus consistent with Corbit and Balleine's (2003) previous findings with food-reinforced chain responding. Also consistent with the earlier results, Hellemans, Dickinson, and Everitt (2006) demonstrated that incentive learning influenced the ability of a CS associated with heroin withdrawal to invigorate procurement responding. That is, a withdrawal-associated CS enhanced procurement only if rats had received the opportunity to make the consumption response for heroin while they were in the withdrawal state (cf. Balleine, 1992, 2001). Thus, in drug self-administration, behaviors in a chain seem to follow the motivational rules that also influence responses in food-reinforced chains.

Associative control of responding in discriminated heterogeneous chains

In an effort to acquire more information about the associative structure underlying chains, we have developed a heterogeneous chain procedure in which procurement and consumption responses can occur at any time, but are occasioned by separate SDs. The procedure is illustrated in Figure 1. When a procurement SD (a panel light adjacent to the response manipulandum) is turned on, making a procurement response (e.g., pulling a chain suspended from the ceiling) leads to the presentation of a new consumption SD (a separate panel light) for a separate consumption response (e.g., pressing a lever). The consumption response is then reinforced with a food pellet. The rats are required to complete an RR 4 schedule in each link of the chain in the presence of each SD (cf. Corbit & Balleine, 2003); the chain and lever manipulanda are otherwise available continuously. Rats readily learn the chain and increase their responding on procurement and consumption across sessions of acquisition. As shown in Figure 2a, consumption responding usually occurs at a higher rate than procurement due to its proximity to the reinforcer (Gollub, 1977). Note that responding is presented as the elevation of response rate in the SD above responding in a 30-s pre SD period. By the end of acquisition, each response is under strong stimulus control; as can be seen in Figure 2b, each SD selectively increases its correlated response. Responses otherwise occur at low rates without SDs and during the SD for the other response (Thrailkill & Bouton, 2015a, 2016a). By using SDs that are separate from the response manipulanda, we are able to study how each SD influences the animal's choice of which response to perform. And at least as important, as we illustrate below, the procedure allows us to manipulate the responses and their SDs independently.

Effects of extinction of consumption on procurement and procurement on consumption

To further understand the associative structure underlying the chain, Thrailkill and Bouton (2016a) used the procedure described above to examine the effects of consumption extinction on procurement responding. Recall that Olmstead et al. (2001; see also Zapata et al., 2010) found that extinction of the consumption response, in comparison to further reinforced training with it, led to weakened procurement responding. Because their experiment did not compare consumption extinction to an untreated control group, it was not clear whether the results were due to the effects of extinction in the experimental group or

reinforcement in the control group (or both). Second, it was not clear whether emitting the consumption response during the extinction treatment was required to produce the possible extinction effect. Thus, the experiment did not determine whether the current value of the consumption response, or the value of the consumption SD (which theoretically constitutes a conditioned reinforcer for procurement), is the crucial factor influencing the strength of procurement responding. In one of our experiments (Thraikill & Bouton, 2016a, Experiment 1), after acquiring a behavior chain using the procedure just described, three groups of rats received different consumption extinction treatments. All involved repeated presentations of the consumption SD. Two groups were allowed to make the consumption response (without reinforcement) whenever the consumption SD was presented. For Group C+P, both the consumption and the procurement manipulanda were available in extinction, and for Group C only, the consumption manipulandum was available alone. A third group received the consumption SD with both manipulanda removed (Group SC only); this group could not make either the procurement or consumption response. A control group received equivalent handling, but was not placed in the experimental chamber (Group Handle). After consumption extinction, all rats received a test in which they could make the procurement response during presentations of the procurement SD. Both manipulanda were present during testing.

Results from the test are shown in Figure 3. Perhaps consistent with previous results (Olmstead et al., 2001), extinction of consumption weakened procurement responding, and this was true whether or not procurement responding had been available during extinction. However, the effect depended importantly on the rat being able to perform the consumption response during extinction: Groups C+P and C only, which were both allowed to make the consumption response during the SD during extinction, each demonstrated weaker procurement responding than the other groups. But Group SC only, which received Pavlovian extinction exposures to the consumption SD without being able to make the response, did not differ from Group Handle. Thus, extinction presentations of the consumption SD were not alone sufficient to weaken the procurement response. Evidently, it is the value of the consumption response, and not the possible Pavlovian value of the consumption SD, that provides the “goal” directing goal-directed procurement responding.

A subsequent experiment asked whether the effect of consumption extinction was specific to the procurement response with which it had been chained (Thraikill & Bouton, 2016a, Experiment 2). Rats first learned two heterogeneous chains that led to the same food reinforcer. One procurement SD (e.g., panel light) set the occasion for one procurement response (e.g., nose-poking), which then led to a consumption SD that set the occasion for a specific consumption response (e.g., pressing the lever to the right of the food cup) and a reinforcer. Trials with this chain were intermixed with trials with a separate chain composed of two different responses (e.g., a chain-pull leading to pressing the lever to the left of the food cup) and their corresponding SDs. Following acquisition of both chains, rats received extinction of one consumption response during presentations of its SD. Then there was a final test of the two procurement responses during presentations of their respective SDs. In the test, rats suppressed the procurement response that was specifically associated with the extinguished consumption response. The rate of the other procurement response did not differ from the rate observed in a control group that had not received any extinction. The

result clearly suggests that the effect of consumption extinction on procurement is specific to the procurement response with which it was associated. It is not due to generalization between the responses, general frustration that might arise in extinction, or a possible suppression of the animal's representation of the common reinforcer; any of these processes would have affected both procurement responses equally. The suppression of procurement responding is due specifically to the weakened value of its associated consumption response.

The effect of extinguishing a response from one component of the chain on responding in the other component is reminiscent of earlier results with Pavlovian serial compound conditioning (e.g., Holland & Ross, 1981). For example, after training with a serial compound conditioned stimulus (CS1-CS2-US), extinction of responding to one of the CSs results in an attenuation of the response to the other CS. Interestingly, this effect does not depend on the position of the CS in the series: Extinction of CS2 weakens responding to CS1, but extinction of CS1 also weakens responding to CS2 (see Holland, 1990 for review). We therefore reasoned that if the components of the instrumental chain are similarly associated, then the effect of consumption extinction on procurement might be reversed. That is, extinction of procurement responding might likewise weaken consumption. Thrailkill and Bouton (2015a) tested this prediction directly. Following acquisition of a discriminated behavior chain, one group of rats received procurement extinction in which nonreinforced exposures to the procurement SD occurred with both response manipulanda in the chamber (Group P+C), and a second group could only make the procurement response (Group P only). A third group received presentations of the procurement SD with both manipulanda removed (Group SP only), so that they uniquely had no opportunity to learn about the specific value of the response. A fourth group received handling, but was not placed in the experimental chamber (Group Handle). Following the extinction phase, all groups received tests of consumption responding in the presence of the consumption SD. The test results, shown in Figure 4, clearly indicated that procurement extinction weakened consumption responding. And in a further parallel with the other results, this effect depended on whether the rats made the procurement response in extinction: Consumption responding was suppressed in Groups P+C and P only, but not in Group SP only. Extinction exposure to the procurement SD alone was not sufficient to affect the consumption response.

A subsequent two-chain experiment further demonstrated that the effect was specific to the consumption response associated with the extinguished procurement response (Thrailkill & Bouton, 2015a, Experiment 3). After training with two chains, extinction of one procurement response caused rats to make fewer responses on the consumption response that had been specifically associated with it during training. As in the previously-described two-chain experiment, this result suggests that extinction of one response does not weaken the other through simple generalization, frustration, or suppression of the representation of the reinforcer. Alternatively, but consistent with analyses of Pavlovian serial compounds (Holland, 1990), the effect might involve a form of representation-mediated extinction that requires inhibition of the procurement response.

We would note that the importance of making the response during extinction that was evident in both series of experiments is consistent with other results suggesting that making the response during extinction might be essential for the success of instrumental extinction

(Bouton, Trask, & Carranza-Jasso, 2016). For example, after simple discriminated operant training, making the response during extinction may be necessary to weaken the instrumental response (Bouton et al., 2016); Pavlovian exposure to the SD alone is not sufficient. Moreover, extinction of a response that is occasion-set by an SD specifically weakens that response, and not another response that is occasion-set by the same SD; it also weakens the same response, but not a different response, occasioned by other SDs (Bouton et al., 2016). Animals appear to learn to inhibit the instrumental response during instrumental extinction (see also Rescorla, 1993, 1997).

To summarize, extinction procedures have revealed several features of the associative structure learned in performing a chain of discriminated instrumental responses. Extinction of consumption weakens procurement, and extinction of procurement weakens consumption. These effects are specific to the associated response in the chain, and critically depend on animals having the opportunity to learn to inhibit the response in extinction. Recent evidence further suggests that procurement and consumption responses are insensitive to outcome reevaluation after training (Thrailkill & Bouton, 2016b). That is, after a modest amount of training with the discriminated chain, multiple pairings of the food-pellet reinforcer with lithium chloride caused the rat to completely reject the food pellet—but had no impact on either procurement or consumption responding (lever pressing or chain pulling) during subsequent testing. These results further point to the consumption response, and not the reinforcer, as the “goal” of the procurement response (see above; Olmstead et al., 2001; Trailkill & Bouton, 2016a; Zapata et al., 2010). It is worth noting that in a discriminated chain, different SDs partition procurement, consumption, and the reinforcer, perhaps explaining why outcome value may have less role in the discriminated chain than in chains that are not segmented by separate SDs (cf. Balleine et al., 1995; Ostlund et al., 2009). Overall, our evidence has consistently highlighted the importance of an association between the procurement and consumption responses in the discriminated heterogeneous chain.

Contextual control of chained behavior and of chain extinction

The context in which learning occurs is often an important factor in controlling learned performance (Bouton, 2004; Bouton & Todd, 2014). Does the context in which the chain takes place therefore enter into the associative structure of the chain? Studies of extinction in both Pavlovian and instrumental learning suggest that extinction is a form of new learning that is especially dependent on the context for expression (see Vurbic & Bouton, 2014 for review). The best example of the contextual control of extinction is the so-called “renewal effect.” Responses conditioned in one context (Context A) then extinguished in a second context (Context B) return when tested back in the original conditioning context (A) or a third context (Context C) (Bouton, Todd, Vurbic, & Winterbauer, 2011). Renewal also occurs with a response trained and extinguished in Context A and tested in a new context (Context B) (Bouton et al., 2011). The AAB and ABC forms of renewal are important because they demonstrate that context change is sufficient for renewal, and suggest that extinction results in inhibitory learning that is context dependent. Although there has been a substantial amount of interest in the renewal of extinguished instrumental behaviors in drug self-administration (see Bouton, Winterbauer, & Vurbic, 2012 for review; Crombag &

Shaham, 2002; Hamlin, Clemens, & McNally, 2008), renewal has not been studied with extinguished chained responses.

A recent series of experiments therefore addressed this question (Thrailkill, Trott, Zerr, & Bouton, 2016). We first examined the contextual control of procurement and consumption extinction. Rats learned our usual discriminated behavior chain, and then received extinction of either the procurement or the consumption response (separate from the chain) in either the acquisition context (Context A) or in a second context (Context B). (The two contexts were different operant chambers that differed along a number of dimensions.) In either case, the appropriate SD was presented, and responding could occur to turn off the SD without leading to the next part of the chain (in the case of the procurement response), or the reinforcer (in the case of the consumption response). The extinguished SD-response combination was then tested in both contexts (cf. Bouton et al., 2011). The results clearly showed that the extinction of both procurement and consumption was context-specific. Renewal occurred in Context A after extinction in Context B (ABA renewal); it also occurred in Context B following extinction in Context A (AAB renewal). A second experiment examined renewal of the individual responses after extinction of the entire chain. That is, after training the chain in Context A, rats were allowed to perform the entire chain without the final reinforcer in either Context A or B. In the extinction procedure, completion of the procurement requirement (RR 4) led to the consumption stimulus, and completion of the consumption requirement (RR 4) in that stimulus turned it off but did not produce the primary (food-pellet) reinforcer. (Trials in which the rat did not complete the procurement requirement eventually transitioned into the consumption stimulus.) After extinction of the chain, half the rats were tested with the procurement response and half with the consumption response in each of the two contexts. In the test, both procurement and consumption responses recovered (were renewed) when they were tested outside the extinction context. Therefore, in a manner similar to simple operant responses, chained responses extinguished either separately or within the chain readily renew when the context is changed. The context is thus a part of the content of extinction learning.

We have previously noted a difference in the effects of a context switch after simple Pavlovian versus instrumental conditioning. Whereas Pavlovian responding to a CS often transfers very well across contexts (e.g., Bouton & King, 1983; Bouton & Peck, 1989), instrumental responding, even in the presence of an SD, does not (Bouton et al., 2011; Bouton, Todd, & León, 2014; Trailkill & Bouton, 2015b). Thus, the context controls the strength of an instrumental, though not a Pavlovian, response. It is interesting to note that the chain renewal experiments just described included groups that received extinction in either the same or a different context. Remarkably, the evidence suggested that the context switch affected the procurement response, but not the consumption response. That is, groups that received extinction of procurement showed an immediate decrement when switched and extinguished in Context B; however, groups that received extinction of consumption showed no such decrement when switched to Context B. Thus, the context appeared to play a more important role in controlling procurement than consumption. Another observation was that consumption responding was weakened by the context switch if it was tested in the whole chain, i.e., when it followed a procurement response that was itself weakened by the context change. The results thus suggested that the procurement response, but not the physical

background, was the “context” for consumption responding. In chain training, procurement responding always occurs before the consumption response; perhaps this allows the procurement component to compete with the operant chamber for effective “contextual” control of the consumption response.

If procurement is indeed the context for consumption, then a consumption response that receives extinction outside the context of the chain should be renewed when it is returned to the chain. Two further experiments (Thrailkill et al., 2016) tested exactly this. In the first experiment, rats learned the discriminated chain before receiving separate extinction of the consumption response. In the extinction trials, the consumption SD was presented, and completing the consumption response requirement turned the SD off but did not produce a food pellet. Half the rats then received a “renewal” test of consumption with the procurement SD preceding presentation of the consumption SD. One group received the procurement SD (SP) alone (the procurement manipulandum was absent), while the other was allowed to make the procurement response in the SD. The remaining rats received further consumption extinction (No SP); half had the procurement manipulandum available, and half did not. The results are shown in Figure 5. Returning the extinguished consumption response to the chain did indeed cause it to be renewed. However, renewal was only observed when the procurement response was available; renewal of consumption was not produced by the reintroduction of the procurement SD alone. Thus, the procurement response, but not the procurement SD, functioned as a “context” for the consumption response.

It was possible, however, that the opportunity to make the procurement response caused “renewal” in some other way. For example, it guaranteed that the animal was in the front of the chamber, closer to the consumption manipulandum, perhaps enabling more consumption responding in the test. To rule out such effects, a further experiment investigated whether renewal of a consumption response only occurred upon return to the specific chain in which it was trained. All rats first learned two separate chains (cf., Trailkill & Bouton, 2015a, 2016a). They then received separate extinction of both consumption responses; half with procurement manipulanda available (Group With Procurement) and half with procurement manipulanda removed (Group Without Procurement). Next, rats were tested with one consumption response. There were three types of test trials: Those in which the consumption SD occurred alone, those in which the consumption SD was preceded by the procurement SD with which it had been trained (congruent trials), and those in which the consumption SD was preceded by the procurement SD that had been trained with the *other* consumption response (incongruent trials). Importantly, renewal of consumption responding occurred in the congruent trials, but not in the incongruent trials. That result suggests that renewal of consumption occurred exclusively following a return to the context of the chain in which it had been trained. Moreover, this occurred only in a group that was allowed to make the procurement response, and not a group tested with the procurement SDs alone. Thus, the context controlling renewal of the extinguished consumption response was indeed the specific procurement response with which it had been associated.

Our results suggesting renewal of consumption upon return to the chain further underscore one of the main conclusions supported by the earlier experiments: An important part of what

is learned in our discriminated heterogeneous chain procedure is an association between the procurement and consumption responses. The recent experiments strongly suggest that the procurement response (and not merely the procurement SD) is the context for consumption responding. Interestingly, our preliminary results suggest that the background apparatus context is not. Consistent with this pivotal role of the preceding response, the experiments on the effects of consumption and procurement extinction on procurement and consumption responding likewise suggested that the two responses (and not simply their SDs) are crucially associated.

Concluding comments

Until recently, the extant literature on heterogeneous instrumental behavior chains mainly focused on the motivational control of chained behaviors and drug self-administration. These studies had uncovered several important effects that suggest instrumental incentive learning and Pavlovian incentive motivational processes selectively influence procurement and consumption responses, respectively (Corbit & Balleine, 2003; Wassum et al., 2011b). Chains have also been useful for modeling processes involved in the development of habitual and compulsive behaviors (Chen et al., 2013; Zapata et al., 2010).

To further understand the associative “content” of heterogeneous chains, we have introduced a discriminated chain procedure that employs separate SDs for procurement and consumption, but separates the responses from the SDs that occasion them. The procedure has allowed for a systematic analysis of the underlying associative structure supporting performance of the chain, and has uncovered a number of noteworthy phenomena. As just described, whether the responses are extinguished separately or together in the chain, extinction of responses learned in a chain is context-dependent (Thraillkill et al., 2016), as single-trained instrumental responses are (Bouton et al., 2016; Todd, Vurbic, & Bouton, 2014). And when a consumption response is extinguished outside the chain, it is renewed when placed back into the context of the preceding response in the chain. This result expands the definition of “context” by adding responses to list of physical, temporal, and reinforcer variables that are known to control the retrieval and expression of learned performance (Bouton, 2004; Bouton & Trask, 2016).

An integrative understanding of behavior chains will ultimately have translational value. Our results suggesting the importance of making the response during extinction (Bouton et al., 2016; Thraillkill & Bouton, 2015a, 2016a) may be consistent with findings suggesting that simple Pavlovian exposure to drug-associated cues are not sufficient to weaken smoking or drug-taking (e.g., Conklin & Tiffany, 2002). The new idea is that in instrumental situations, the organism must directly learn to inhibit the instrumental response. Studies in smokers also suggest the importance of distal stimuli, which may not be directly associated with actual smoking; for example, they may evoke craving responses in a manner similar to stimuli more proximal to consumption (Conklin, Robin, Perkins, Salkeld, & McClernon, 2008). It is worth noting that distal behaviors may be easier to extinguish directly than more proximal (consumption) behaviors: Drug users do not inject saline, smokers rarely smoke denicotinized cigarettes, and junk food eaters do not chew and swallow without food in their mouths. One of the main discoveries of our work to date is that extinction of one response

can weaken the other. Ordinarily, the focus of treatments is necessarily to inhibit consumption (e.g., smoking); our work suggests that inhibiting procurement can also suppress consumption (Thraikill & Bouton, 2015a). The importance and value of extinguishing procurement is further indicated by the fact that, left uninhibited, procurement responding can cause renewal of consumption when extinguished consumption is returned to the chain (Thraikill et al., 2016). For chained instrumental behaviors, like other behaviors, extinction can be effective, but it must be remembered that the result is often specific to the context in which it is learned.

Acknowledgments

Preparation of the manuscript was supported by Grant R01 DA033123 from the National Institute on Drug Abuse to MEB.

References

- Balleine BW. Instrumental performance following a shift in primary motivation depends on incentive learning. *Journal of Experimental Psychology: Animal Behavior Processes*. 1992; 18:236–250. [PubMed: 1619392]
- Balleine, BW. Incentive processes in instrumental conditioning. In: Mowrer, RR.; Klein, SB., editors. *Handbook of contemporary learning theories*. Mahwah, NJ: Erlbaum; 2001. p. 307-366.
- Balleine BW, Garner C, Gonzalez F, Dickinson A. Motivational control of heterogeneous instrumental chains. *Journal of Experimental Psychology: Animal Behavior Processes*. 1995; 21:203–217.
- Balleine BW, Paredes-Olay C, Dickinson A. Effects of outcome devaluation on the performance of a heterogeneous instrumental chain. *International Journal of Comparative Psychology*. 2005; 18:257–272.
- Bouton ME. Context and behavioral processes in extinction. *Learning and Memory*. 2004; 11:485–494. [PubMed: 15466298]
- Bouton ME, King DA. Contextual control of the extinction of conditioned fear: Tests for the associative value of the context. *Journal of Experimental Psychology: Animal Behavior Processes*. 1983; 9:248–265. [PubMed: 6886630]
- Bouton ME, Peck CA. Context effects on conditioning, extinction, and reinstatement in an appetitive conditioning preparation. *Animal Learning & Behavior*. 1989; 17:188–198.
- Bouton ME, Todd TP. A fundamental role for context in instrumental learning and extinction. *Behavioural Processes*. 2014; 104:13–19. [PubMed: 24576702]
- Bouton ME, Todd TP, León SP. Contextual control of discriminated operant behavior. *Journal of Experimental Psychology: Animal Learning and Cognition*. 2014; 40:92–105. [PubMed: 24000907]
- Bouton ME, Todd TP, Vurbic D, Winterbauer NE. Renewal after the extinction of free operant behavior. *Learning & Behavior*. 2011; 39:57–67. [PubMed: 21279496]
- Bouton ME, Trask S. Role of the discriminative properties of the reinforcer in resurgence. *Learning & Behavior*. 2016 In press.
- Bouton ME, Trask S, Carranza-Jasso R. Learning to inhibit the response during instrumental (operant) extinction. *Journal of Experimental Psychology: Animal Learning and Cognition*. 2016 In press.
- Bouton, ME.; Winterbauer, NE.; Vurbic, D. Context and extinction: Mechanisms of relapse in drug self-administration. In: Haselgrove, M.; Hogarth, L., editors. *Clinical applications of learning theory*. East Sussex, UK: Psychology Press; 2012. p. 103-134.
- Chen BT, Yau H, Hatch C, Kusumoto-Yoshida I, Cho SL, Hopf FW, Bonci A. Rescuing cocaine-induced prefrontal cortex hypoactivity prevents compulsive cocaine seeking. *Nature*. 2013; 496:359–362. [PubMed: 23552889]
- Conklin CA, Robin N, Perkins KA, Salkeld RP, McClernon FJ. Proximal versus distal cues to smoke: The effects of environments on smokers' cue reactivity. *Experimental and Clinical Psychopharmacology*. 2008; 16:207–214. [PubMed: 18540780]

- Conklin CA, Tiffany ST. Applying extinction research and theory to cue-exposure addiction treatments. *Addiction*. 2002; 97:155–167. [PubMed: 11860387]
- Collier, GH. Determinants of choice. In: Bernstein, DJ., editor. *Nebraska Symposium on Motivation*. Lincoln, NE: University of Nebraska Press; 1981. p. 67-127.
- Corbit LH, Balleine BW. Instrumental and Pavlovian incentive processes have dissociable effects on components of a heterogeneous instrumental chain. *Journal of Experimental Psychology: Animal Behavior Processes*. 2003; 29:99–106. [PubMed: 12735274]
- Crombag HS, Shaham Y. Renewal of drug seeking by contextual cues after prolonged extinction in rats. *Behavioral Neuroscience*. 2002; 116:169–173. [PubMed: 11895178]
- Dickinson, A. Instrumental conditioning. In: Mackintosh, NJ., editor. *Animal learning and cognition: Handbook of perception and cognition series*. 2nd. San Diego, CA: Academic Press; 1994. p. 45-79.
- Dickinson A, Balleine BW. Motivational control of goal-directed action. *Animal Learning & Behavior*. 1994; 22:1–18.
- Dickinson A, Balleine B, Watt A, Gonzalez F, Boakes RA. Motivational control after extended instrumental training. *Animal Learning & Behavior*. 1995; 23:197–206.
- Economidou D, Pelloux Y, Robbins TW, Dalley JW, Everitt BJ. High impulsivity predicts relapse to cocaine-seeking after punishment-induced abstinence. *Biological Psychiatry*. 2009; 65:851–856. [PubMed: 19181308]
- Gollub, L. Conditioned reinforcement: schedule effects. In: Honig, WK.; Staddon, JER., editors. *Handbook of operant behavior*. Englewood Cliffs, NJ: Prentice-Hall; 1977. p. 288-312.
- Greybiel AM. The basal ganglia and chunking of action repertoires. *Neurobiology of Learning and Memory*. 1998; 70:119–136. [PubMed: 9753592]
- Hamlin AS, Clemens KJ, McNally GP. Renewal of extinguished cocaine-seeking. *Journal of Neuroscience*. 2008; 151:656–670.
- Hellemans KGC, Dickinson A, Everitt BJ. Motivational control of heroin seeking by conditioned stimuli associated with withdrawal and heroin taking by rats. *Behavioral Neuroscience*. 2006; 120:103–114. [PubMed: 16492121]
- Holland PC. Event representation in Pavlovian conditioning: Image and action. *Cognition*. 1990; 37:105–131. [PubMed: 2269004]
- Holland PC, Ross RT. Within-compound associations in serial compound conditioning. *Journal of Experimental Psychology: Animal Behavior Processes*. 1981; 7:228–241.
- Johnson AW, Bannerman D, Rawlins N, Sprengel R, Good MA. Targeted deletion of the GluR-1 AMPA receptor in mice dissociates general and outcome-specific influences of appetitive rewards on learning. *Behavioral Neuroscience*. 2007; 121:1192–1202. [PubMed: 18085873]
- Jonkman S, Pelloux Y, Everitt BJ. Drug intake is sufficient, but conditioning is not necessary for the emergence of compulsive cocaine seeking after extended self-administration. *Neuropsychopharmacology*. 2012; 37:1612–1619. [PubMed: 22334124]
- LeBlanc KH, Ostlund SB, Maidment NT. Pavlovian-to-instrumental transfer in cocaine seeking rats. *Behavioral Neuroscience*. 2012; 126:681–689. [PubMed: 22866668]
- Olmstead MC, Lafond MV, Everitt BJ, Dickinson A. Cocaine seeking by rats is a goal-directed action. *Behavioral Neuroscience*. 2001; 115:394–402. [PubMed: 11345964]
- Olmstead MC, Parkinson JA, Miles FJ, Everitt BJ, Dickinson A. Cocaine-seeking by rats: Regulation, reinforcement and activation. *Psychopharmacology*. 2000; 152:123–131. [PubMed: 11057515]
- Ostlund SB, Balleine BW. On habits and addiction: an associative analysis of compulsive drug seeking. *Drug Discovery Today: Disease Models*. 2009; 5:235–245. [PubMed: 20582332]
- Ostlund SB, Winterbauer NE, Balleine BW. Evidence of action sequence chunking in goal-directed instrumental conditioning and its dependence on the dorsomedial prefrontal cortex. *Journal of Neuroscience*. 2009; 29:8280–8287. [PubMed: 19553467]
- Pelloux Y, Everitt BJ, Dickinson A. Compulsive drug seeking by rats under punishment: Effects of drug taking history. *Psychopharmacology*. 2007; 194:127–137. [PubMed: 17514480]
- Pryor, K. *Don't shoot the dog: The new art of teaching and training*. New York: Bantam Books; 1984.

- Rescorla, RA. Pavlovian second-order conditioning: Some implications for instrumental behaviour. In: Davis, H.; Hurwitz, HMB., editors. *Operant-Pavlovian Interactions*. Hillsdale, NJ: Erlbaum; 1977. p. 133-164.
- Rescorla RA. Inhibitory associations between S and R in extinction. *Animal Learning & Behavior*. 1993; 21:327–336.
- Rescorla RA. Response inhibition in extinction. *Quarterly Journal of Experimental Psychology*. 1997; 50B:238–252.
- Skinner BF. The extinction of chained reflexes. *Proceedings of the National Academy of Sciences*. 1934; 20:234–237.
- Thraikill EA, Bouton ME. Extinction of chained instrumental behaviors: Effects of procurement extinction on consumption responding. *Journal of Experimental Psychology: Animal Learning and Cognition*. 2015a; 41:232–246. [PubMed: 25915751]
- Thraikill EA, Bouton ME. Contextual control of instrumental actions and habits. *Journal of Experimental Psychology: Animal Learning and Cognition*. 2015b; 41:69–80. [PubMed: 25706547]
- Thraikill EA, Bouton ME. Extinction of chained instrumental behaviors: Effects of consumption extinction on procurement responding. *Learning & Behavior*. 2016a; 44:85–96. [PubMed: 26276367]
- Thraikill EA, Bouton ME. Effects of outcome devaluation on instrumental behaviors in a discriminated heterogeneous chain. Manuscript submitted for publication. 2016b
- Thraikill EA, Trott JM, Zerr C, Bouton ME. Contextual control of chained instrumental behaviors. Manuscript submitted for publication. 2016
- Todd TP, Vurbic D, Bouton ME. Mechanisms of renewal after the extinction of discriminated operant behavior. *Journal of Experimental Psychology: Animal Learning and Cognition*. 2014; 40:355–368. [PubMed: 25545982]
- Vanderschuren LJMJ, Everitt BJ. Drug seeking becomes compulsive after prolonged self-administration. *Science*. 2004; 305:1017–1019. [PubMed: 15310907]
- Vurbic, D.; Bouton, ME. A contemporary behavioral perspective on extinction. In: McSweeney, FK.; Murphy, ES., editors. *The Wiley–Blackwell handbook of operant and classical conditioning*. Chichester, UK: Wiley-Blackwell; 2014. p. 53-76.
- Wassum KM, Cely IC, Balleine BW, Maidment NT. μ -Opioid receptor activation in the basolateral amygdala mediates the learning of increases but not decreases in the incentive value of a food reward. *Journal of Neuroscience*. 2011b; 31:1591–1599. [PubMed: 21289167]
- Wassum KM, Ostlund ST, Maidment NT, Balleine BW. Distinct opioid circuits determine the palatability and the desirability of rewarding events. *Proceedings of the National Academy of Science*. 2009; 106:12512–12517.
- Wassum KM, Ostlund ST, Balleine BW, Maidment NT. Differential dependence of Pavlovian incentive motivation and instrumental incentive learning processes on dopamine signaling. *Learning & Memory*. 2011a; 18:475–483. [PubMed: 21693635]
- Zapata A, Minney VL, Shippenberg TS. Shift from goal-directed to habitual cocaine seeking after prolonged experience in rats. *Journal of Neuroscience*. 2010; 30:15457–15463. [PubMed: 21084602]

Highlights

- This article reviews the basic learning and motivational processes that underlie heterogeneous instrumental behavior chains
- It emphasizes research with *discriminated* chains in which two behaviors required for reinforcement are occasioned by their own stimuli
- Extinction of either the first or second response weakens the other
- Making the response in extinction is necessary to produce these effects
- When the second response is extinguished outside the chain, it is renewed when returned to the context of the chain
- These and other results have theoretical and translational implications for understanding behavior chains

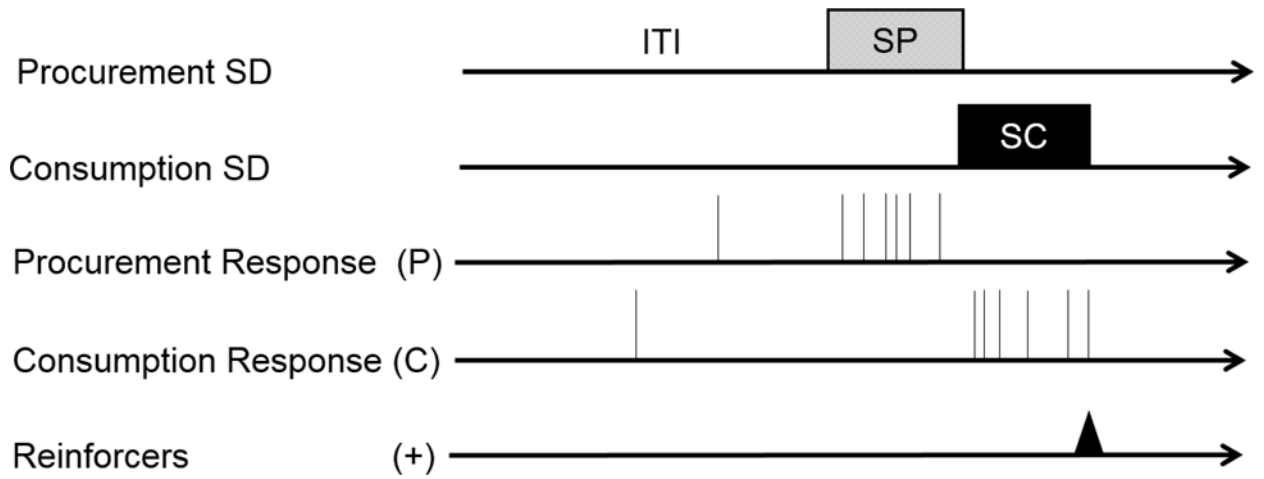


Figure 1. Timeline of events in our discriminated heterogeneous instrumental chain procedure. SD = discriminative stimulus; ITI = intertrial interval; SP = procurement stimulus; SC = consumption stimulus; P = procurement response; C = consumption response.

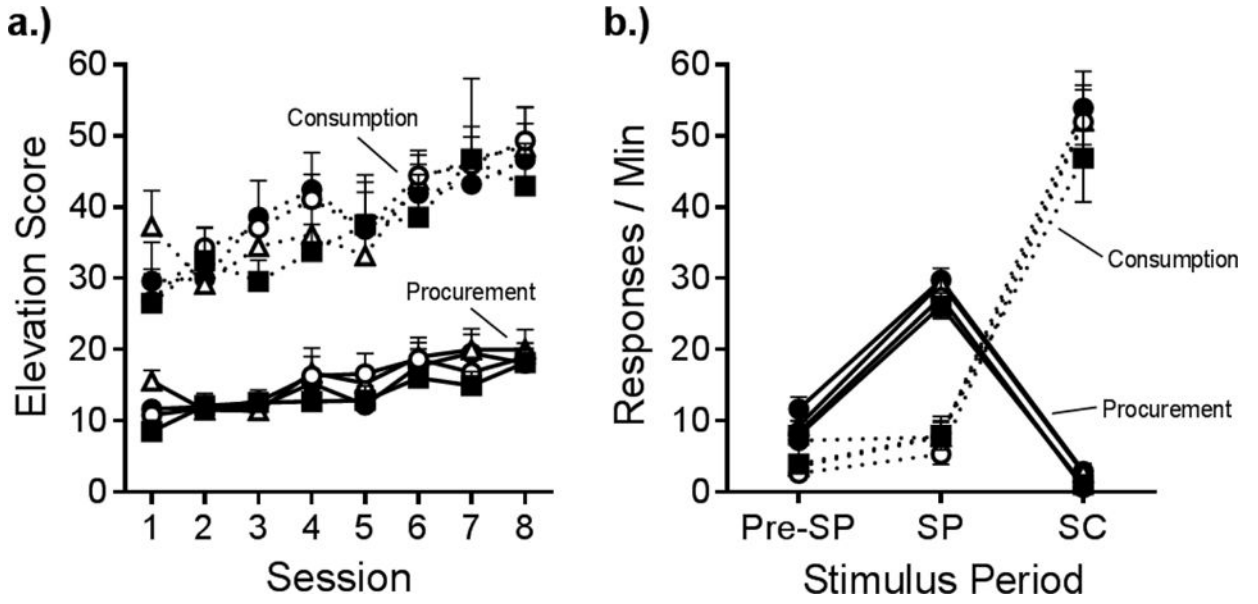


Figure 2. Acquisition and final performance of a discriminated heterogeneous instrumental chain. a.) Mean elevation scores of procurement and consumption responding over the final acquisition sessions in which the response requirement for each was Random Ratio 4. b.) Mean response rates in the final acquisition session on procurement and consumption manipulanda during the 30-s pre-procurement SD period (Pre-SP), the procurement SD (SP), and the consumption SD (SC). Error bars are the standard error of the mean. From Thraikill and Bouton (2015a).

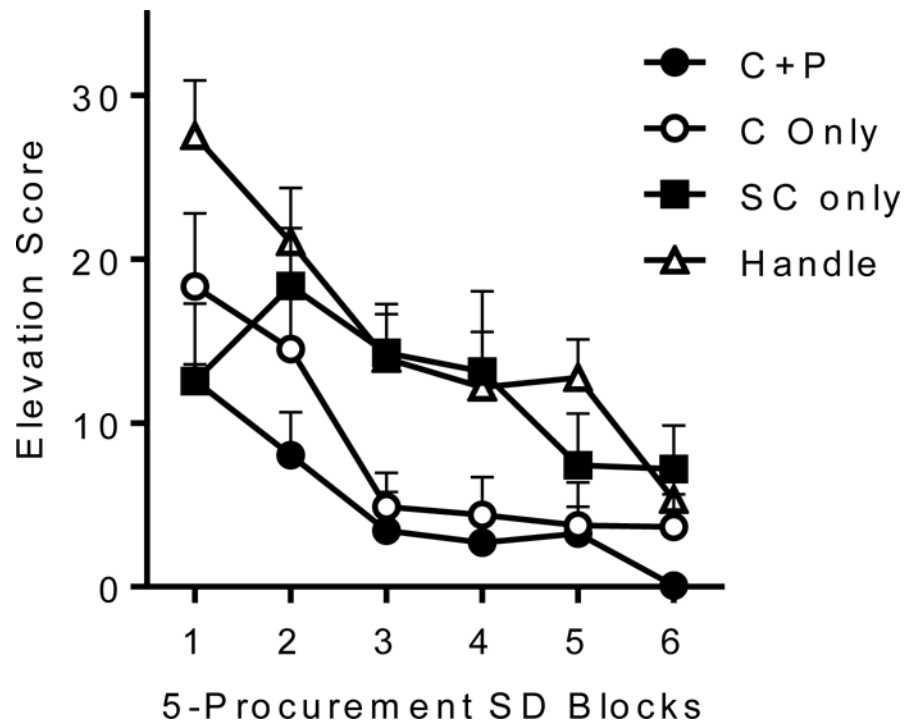


Figure 3. Effects of consumption extinction on procurement responding. Mean elevation scores for the procurement response is shown over blocks of 5 procurement SD presentations during testing. Groups are described in the text. Error bars are the standard error of the mean. From Thraikill and Bouton (2016a).

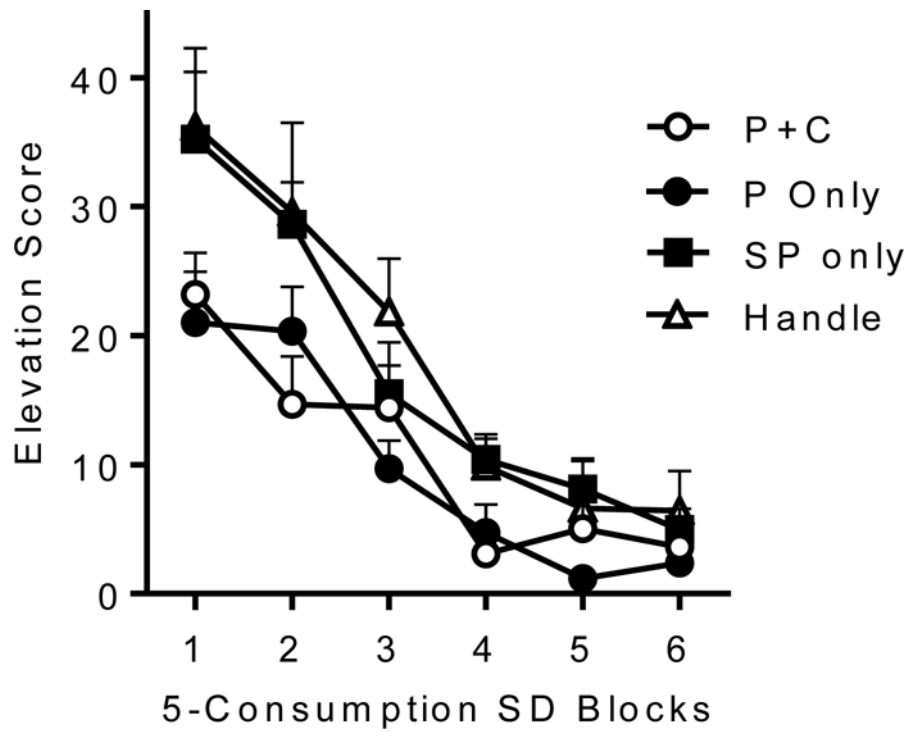


Figure 4. Effect of procurement extinction on consumption responding. Mean elevation scores for the consumption response over blocks of 5 consumption SD presentations during testing. Groups are described in the text. Error bars are the standard error of the mean. From Thraikill and Bouton (2015a).

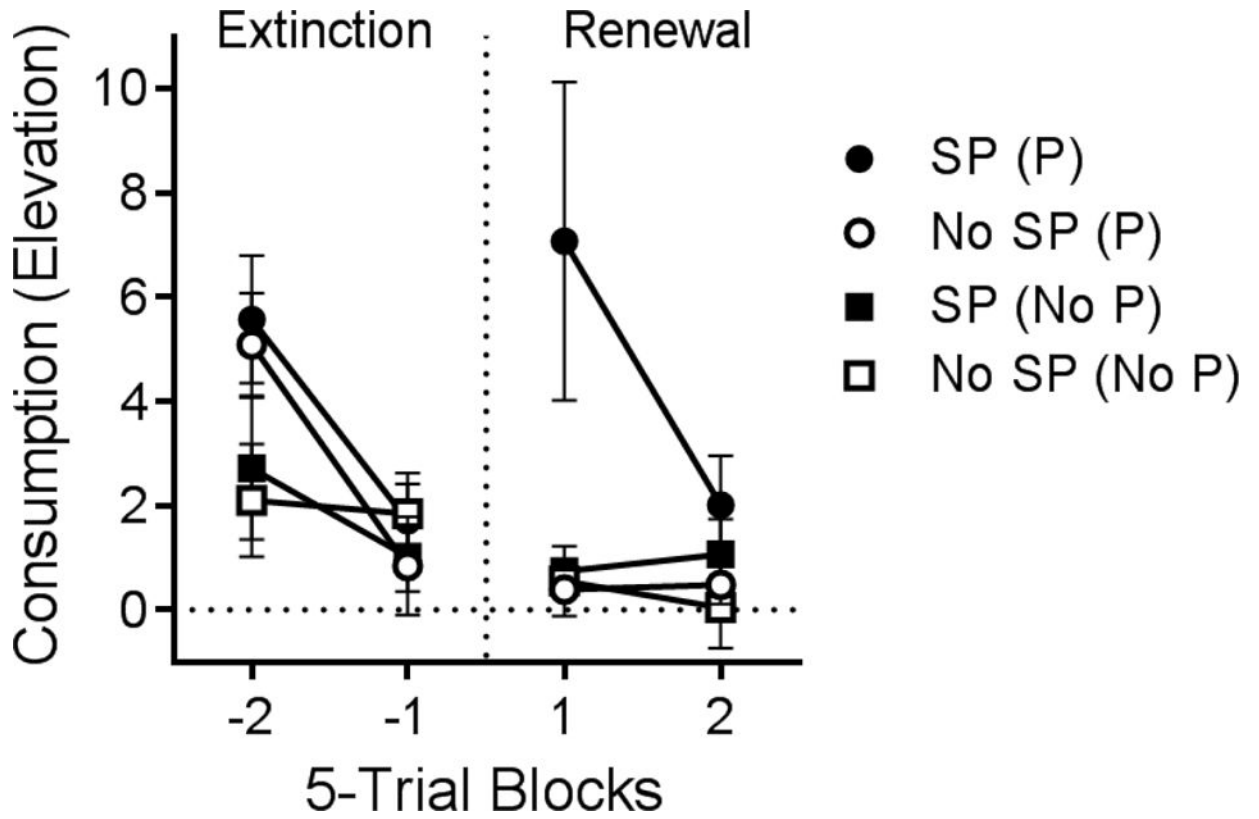


Figure 5. Renewal of extinguished consumption responding when it is returned to the original chain. Mean elevation scores for the consumption response in 5-trial blocks of extinction, followed by 5-trial blocks in which the groups received their respective “renewal” treatments. Only the group that was allowed to make the procurement response to the procurement SD prior to test presentations of the consumption SD showed a renewal of consumption responding [Group SP (P)]. See text for all the other group designations. Error bars are the standard error of the mean and only appropriate for between-group comparisons. Adapted from Thraillkill et al. (2016).