



# HHS Public Access

Author manuscript

*Behav Anal (Wash D C)*. Author manuscript; available in PMC 2020 May 01.

Published in final edited form as:

*Behav Anal (Wash D C)*. 2019 May ; 19(2): 202–212. doi:10.1037/bar0000108.

## Prediction and control of operant behavior: What you see is not all there is

Mark E. Bouton<sup>1</sup>, Bernard W. Balleine<sup>2</sup>

<sup>1</sup>Department of Psychological Science, University of Vermont, USA

<sup>2</sup>Decision Neuroscience Laboratory, School of Psychology, University of NSW, Australia,

### Abstract

Prediction and control of operant behavior are major goals of behavior analysis. We suggest that achieving these goals can benefit from doing more than identifying the three-term contingency between the behavior, its setting stimulus, and its consequences. Basic research now underscores the idea that prediction and control require consideration of the behavior's *history*. As one example, if an operant is a goal-directed action, it is controlled by the current value of the reinforcer, as illustrated by the so-called reinforcer devaluation effect. In contrast, if the behavior is a habit, it occurs automatically, without regard to the reinforcer's value, as illustrated by its insensitivity to the reinforcer devaluation effect. History variables that distinguish actions and habits include the extent of their prior practice and their schedule of reinforcement. Other operants can appear to have very low or zero strength. However, if the behavior has reached that level through extinction or punishment, it may precipitously increase in strength by changing the context, allowing time to pass, presenting the reinforcer contingently or noncontingently, or extinguishing an alternative behavior. Behaviors that are not suppressed by extinction or punishment are not affected the same way. When predicting the strength of an operant behavior, what you see is not all there is. The behavior's history counts.

### Keywords

action; habit; extinction; punishment; behavioral inhibition; behavioral history

---

A major goal of behavior analysis is the prediction and control of operant behavior. Accordingly, since at least Skinner (1938, 1969), and extending through today's use of functional analysis (e.g., Hanley, Iwata, & McCord, 2003), it has been common to emphasize the three-term contingency in the analysis of an operant response: i.e., (1) the occasion within which the response occurs, (2) the response itself, and (3) the reinforcing consequences of the response. Various terms for the elements of the contingency have been used over the years: for example, the occasion has been variously called a discriminative stimulus, state, establishing operation, or context; the response an operant or action; and the reinforcer a reward, consequence, or outcome. But, whatever terms are used, the approach has implied that to predict and control a response it is sufficient to identify its specific setting conditions – its specific occasion - and to develop the means to selectively apply reinforcement. Furthermore, although it has been conceded that identifying and applying these key events might prove challenging in practice, from this perspective it is not a

problem in principle: An important assumption of the behavior analytic tradition is that each term of the contingency is open to direct observation. Therefore, accurately describing the situation, the response, and its consequences will be sufficient to specify the factors that control any operant.

Fundamentally, this means that if a behavior analyst wants to increase or decrease the strength of some specific operant, s/he needs to find its antecedents or consequences and manipulate them accordingly. But basic research from the animal learning laboratory now provides additional information that might help predict what will be effective. Two operants that look very much alike may actually have very different properties; depending on their learning histories, they may be controlled by different kinds of event, or influenced quite differently by the same event. In the present article, we will consider two sets of examples drawn from the current literature. In the first, we will show that two seemingly indistinguishable operants that occur at a stable rate might have a different status as either *goal-directed actions* or *habits*. This influences the variables or events that control them. In the second, we consider what we call *silent operants* that have very low or zero strength. Depending on their actual reinforcement histories, they can respond quite differently to the same event. Both examples violate what Daniel Kahneman (e.g., 2011) has called the What You See is All There Is heuristic used in human decision making. When it comes to operant behavior, What You See is *Not* All There Is. Knowledge of a behavior's history is useful to predict it and control it accurately.

## Actions vs. habits

Historically, the three-term contingency was developed to describe responses that, at that time, were commonly called habits (e.g., Hull, 1943). In accord with the then-dominant neobehaviorist position, the performance of any operant was thought to reflect a stimulus-response association that was strengthened by the selective application of a reinforcer. In the paradigm case, a hungry rat trained to press a lever for food was argued to do so in the presence of various situational, contextual, or discriminative cues (serving as the S) with which the lever press response (the R) became associated due to the delivery of food (the reinforcer) after the response. Such a description accords with the argument that merely observing the rat lever pressing provides the basis for an adequate description of the response and of the contingency controlling that response.

Almost immediately a number of findings began to question the sufficiency of this account. Perhaps the earliest indication that something more than the simple situation-response association was controlling performance came from *incentive contrast* studies. Generally, these studies demonstrated that rats were more concerned with the relative value of the consequences of their actions than they should be if the consequence merely served a reinforcing function. Early studies reported that suddenly changing the value of an earned outcome, e.g., changing the food reward from banana to lettuce in monkeys (Tinklepaugh, 1928), caused an immediate and powerful change in behavior from the pursuit of food to frustration and aggression. In one of the more famous studies involving rats running in a straight runway, a sudden reduction in the amount of food earned by running to the goal box caused an immediate reduction in performance of the response on subsequent trials to a level

that was below that of animals whose responding was maintained by that reduced amount of food from the outset; the so-called negative contrast effect (Crespi, 1942). Although reinforcement theories could address the overall reduction in performance, the rapidity with which the reduction occurred and the actual depression in responding relative to controls – the contrast effect – was difficult to explain in reinforcement terms. What these results suggested was that rather than merely being controlled by the situational cues – or the occasion – within which the operant was performed, it was also influenced by incentive motivation, which required mediation by some sort of representation or expectation of the consequence that the operant produced (e.g., compare Spence, 1956, and Tolman, 1932).

There have been numerous similar findings in mazes and runways supporting this general claim (see Flaherty, 1996 for review), as well as studies showing that even lever pressing in rats can be controlled by a representation of the reward value of the food. Adams and Dickinson (1981), for example, reported that hungry rats trained to lever press for sugar solution were strongly sensitive to changes in the value of the sugar when it was altered offline by conditioned taste aversion. In this study, acquisition of lever pressing for sugar and then taste aversion conditioning to sugar each occurred in separate phases. In a final test, the rats were returned to the lever press apparatus and lever pressing performance was assessed in extinction. (Note that conducting the test in extinction prevented the opportunity to pair the lever-press response directly with the now devalued sugar in experience.) Importantly, the rats with the new taste aversion to sugar nevertheless reduced their lever pressing immediately relative to a group that had received the sugar and the illness unpaired. This is the so-called *reinforcer devaluation effect*. As a consequence of this finding the authors concluded that lever pressing in rats, as implied by the performance observed in mazes and runways, is goal-directed and influenced by the organism's knowledge of the current value of the reinforcer. The organism learns about the behavior and its reinforcer, and when the reinforcer's value changes, the strength of the operant adjusts accordingly.

The benefit of encoding the actual consequence of an action is the flexibility it provides when the value of that consequence changes even without any obvious shift in the context or occasion in which the response is produced. Rather than requiring an animal to perform a response to learn that its consequences are no longer valuable (and potentially noxious), integrating the relationship between the action and its consequences with its experience of the altered value of those consequences is sufficient to alter the performance of the response. This is not always true, however. Although the above studies suggest that operants can be goal-directed, their performance can also appear to be consistent with simpler reinforcement theory. For example, Adams (1982) trained two groups of rats to lever press on a continuous reinforcement schedule for sugar solution. One group was given the opportunity to make 100 reinforced lever presses whereas the other was given the opportunity to make 500 reinforced lever presses. After this training, half of each group was given the taste aversion (reinforcer devaluation) treatment in which the sugar was paired with illness whereas the other half were given the control treatment with the sugar unpaired with illness. Lever pressing was then assessed in all animals in extinction. Despite the fact that the rats were all lever pressing in a seemingly identical manner before devaluation – suggesting that, from a purely observational perspective, similar contingencies should be controlling their performance – the different amounts of training produced very different forms of behavioral

control. The rats given only a moderate amount of training were sensitive to the outcome devaluation treatment and reduced responding compared to the non-devalued control group. In contrast, rats given extended training did not show this effect; although the rats given the devaluation treatment had a strong taste aversion to the sucrose, they lever pressed on test in a similar manner to the unpaired control group. As Adams observed, it appeared that with moderate training the rats' actions were goal-directed whereas when given extended training they had become habits.

These early studies suggested that, depending on the amount of training, it is possible for lever pressing to become goal-directed or habitual. That is, the contingency controlling an operant can either be driven by a situation-response-reinforcement contingency, in the case of habits, or by the association between the action and outcome, in the case of goal-directed actions. Subsequent studies have established a number of other important differences between goal-directed actions and habits. For example, (1) goal-directed actions develop and change very quickly and so can govern *choice* performance: Colwill and Rescorla (1985) were the first to establish that the devaluation of an action alters the rats' willingness to choose its associated action without affecting the performance of other, simultaneously-available actions. Furthermore, (2) rather than being controlled by the context or situation, goal-directed actions are controlled by their association with a specific outcome (Colwill & Rescorla, 1986). As a consequence, goal-directed actions are sensitive to degradation of the action-outcome contingency and selectively so; in a situation where two actions are trained with different reinforcers, adding deliveries of one or other reinforcer that are not contiguous with its associated action reduces the performance of that action while leaving other actions unaffected (Dickinson & Mulatero, 1989; Balleine & Dickinson, 1998). Indeed, (3) when lengths are taken to make the performance of a goal-directed action conditional on a discriminative stimulus (SD), control by the SD is not over the response itself but rather over the response-outcome association (Rescorla, 1990; Bradfield & Balleine, 2013; see also Trask & Bouton, 2014). Other studies have investigated the influence of different forms of devaluation and revaluation of the instrumental reinforcer using shifts in primary motivation (Dickinson & Dawson, 1987; Balleine, 1992), and incentive learning manipulations (Dickinson & Balleine, 1994; Balleine, 2001) including sensory-specific satiety (Balleine & Dickinson, 1998).

In contrast, habits have been found to be difficult to develop in choice situations – despite extensive overtraining, when an explicit choice between two actions is always made available, rats continue to show sensitivity to outcome devaluation (Colwill & Rescorla, 1988; Kosaki & Dickinson, 2010). Furthermore, in contrast to goal-directed actions, habits are strongly and immediately affected by shifts in motivational state (Dickinson et al, 1995) and by shifts in context (Thrailkill & Bouton 2015). And, as their strong relationship to the context suggests, habits have been found to be insensitive to treatments that would otherwise degrade the response-outcome contingency (Dickinson et al., 1996; Dezfouli & Balleine, 2014).

It is important to note that the distinction between goal-directed actions and habits is not unique to rats but is also found in humans. Relatively moderately trained actions in children above the age of three (but not below) show sensitivity to reinforcer devaluation (Klossek,

Russell & Dickinson, 2008), whereas overtraining can render actions insensitive to outcome devaluation and make them habitual (Tricomi, Balleine & O'Doherty, 2010). It should also be noted that actions and habits are not only mediated by distinct contingencies and learning rules but also by distinct neural systems: Balleine and colleagues have established that a cortical basal ganglia circuit involving the medial prefrontal cortex and caudate/dorsomedial striatum in humans and in rodents mediates goal-directed actions whereas a parallel circuit involving the sensorimotor cortices and the putamen/dorsolateral striatum mediates habits (see Balleine & O'Doherty, 2010 for a review). The dissociation of these circuits provides even more evidence for a division between goal-directed actions and habits as fundamentally distinct forms of behavioral control.

Finally, the division between goal-directed actions and habits extends into abnormal behavior (Griffiths et al., 2014). For example, addiction is often characterized both as a loss of behavioral control and the development of a drug-taking habit (e.g., Everitt & Robbins, 2005). Consistent with this characterization, although operant responding for everyday rewards (like chocolate or sugar) in animals and humans can appear relatively normal compared to healthy controls, tests have found that drug exposure attenuates goal-directed action control and causes downregulation of the neural circuitry associated with goal-directed action while causing increases in habitual control and in habit-related circuitry (Hogarth et al., 2013; Furlong et al., 2015). Similarly, deficits in goal-directed action control have been found in adolescents with depression, social anxiety or autism spectrum disorder and in adults with chronic schizophrenia-- based on devaluation tests (Alvarez et al., 2014, 2016; Morris et al., 2015). Conversely, there is some evidence to suggest that disorders such as ADHD and degenerative conditions, such as Parkinson's and Huntington's disease, produce deficits in the habitual control of operant responding causing deficits in operant performance due to the demands of multitasking (Griffiths et al., 2014; Redgrave et al., 2010).

In summary, many studies have pointed to the fact that, depending on the training conditions, the performance of an operant behavior can be controlled by very different contingencies. Goal-directed actions are malleable, flexible, rapidly acquired and relatively readily suppressed or even eliminated from the response repertoire. In contrast, habits are inflexible, automatic and difficult to change or to suppress. Goal-directed actions are less dependent on the situation or occasion for support than habits, which are highly situationally dependent, and whereas goal-directed actions depend on the value of their specific consequences, habits are not concerned with the specific properties or values of their consequences at all. Overall, therefore, in operant or instrumental conditioning, what you see can be ambiguous; an action can appear to be controlled by situational cues but on closer inspection it may be found that shifts in context have little affect performance; it may appear to be controlled by its consequences but subsequent changes in the value of those consequences may demonstrate that it is not. To specify the contingency controlling a response requires, therefore, knowledge of the operant's history and/or additional testing and assessment.

## Operants with low or zero strength

A related case may be made concerning operant behaviors that have very low or even zero rates. How they react to treatment in the future depends crucially on how they achieved those rates. As we describe next, if the response was once reinforced, but was then reduced or eliminated through extinction or punishment, it may return or relapse after some new, precipitating event occurs. Importantly, the return of the response cannot be predicted from simple observations taken before the new event. Again, what you see in behavior is not all there is.

Contemporary research has identified a number of triggering or precipitating events. One of the best studied is a change in the background context, typically the Skinner box or operant chamber in which the animal performs, which can be individuated by different floor composition, spatial locations, and scents. If the context is changed after extinction, an extinguished operant can return or “renew.” The *renewal effect* can take different forms. In the most-studied version, a behavior is reinforced in one context (Context A) and then extinguished in another (Context B). If the behavior is then tested in Context A, the response can return—a phenomenon known as *ABA renewal*. Similarly, if the response is tested in a third context (Context C) after reinforcement and extinction in A then B, responding also recovers (*ABC renewal*). And if responding is reinforced in A and then extinguished in the same context (Context A), the response also renews when tested in a second context (*AAB renewal*). All three forms of renewal have been tested and confirmed with operant behavior (e.g., Bouton, Todd, Vurbic, & Winterbauer, 2011; Todd, 2013). ABA renewal also occurs after extinction in children with developmental disabilities (Kelley, Liddon, Ribeiro, Greif, & Podlesnik, 2015; see Podlesnik, Kelley, Jimenez-Gomez, & Bouton, 2017, for a recent review). And it is also the case that, in animals at least, renewal also occurs if the operant has been suppressed by punishment instead of extinction. ABA and ABC renewal have both been observed after positive punishment, where the response is suppressed by contingent footshock (Bouton & Schepers, 2015); and ABA renewal occurs if the response has been suppressed by an omission (DRO) contingency, i.e., negative punishment (Nakajima, Urushihara, & Masaki, 2002). The renewal effect suggests that neither extinction nor punishment destroys the original operant learning. Instead, the animal learns to refrain from making the response (e.g., Bouton, Trask, & Carranza-Jasso, 2016), and this inhibition is relatively specific to the context in which it is learned (Todd, Vurbic, & Bouton, 2014). Thus, a behavior that has very low or apparently zero strength can come alive again if the context is changed. A suppressed or inhibited operant can be distinguished from a behavior with true zero strength by testing the effect of changing the context.

There are other events that can cause the return or relapse of extinguished or punished operants. In spontaneous recovery, the mere passage of time after extinction (e.g., Rescorla, 2004) or punishment (Estes, 1944) can cause the response to return again. In our view (e.g., Bouton, 1988), spontaneous recovery is another renewal effect in which the context change is provided by the passage of time. Just as extinguished responding renews when it is tested in a new physical context, it recovers when it is tested in a new temporal context. In fact, many different types of stimuli can play the role of context (e.g., Bouton, 2002). In state-dependent learning, drug states provide the context (e.g., Overton, 1985); for example, when

extinction occurs when the organism is under the influence of a drug like a benzodiazepine or alcohol, the response renews when the animal is tested without the drug, i.e., sober, again (e.g., Bouton, Kenny, & Rosengard, 1990; see also Cunningham, 1979; Lattal, 2007). And recent experiments have suggested that hunger state can be a context: When rats lever pressed for sucrose or sweet-fatty pellets (rodent junk food) while they were satiated and then received extinction while they were hungry, the response recovered when it was tested while the rats were satiated again (Schepers & Bouton, 2017). Such renewal may explain the difficulties faced by dieters who might eat when they do not need food and then starve themselves on a diet—the inhibition of eating when dieting may not transfer to well to the satiated state. Other recent research has shown that when the organism must make a sequence of two separate responses to earn a reinforcer (specifically, in a discriminated heterogeneous chain), the first response provides a kind of context for the second response (Thraikill & Bouton, 2016). For example, if the second response is extinguished alone, apart from the chain, it is renewed when it is returned to the chain and tested after the rat has made the first response (Thraikill, Trott, Zerr, & Bouton, 2016). The first response is more affected by changing the physical (Skinner box) context than is the second response. The point is that many kinds of stimuli or events can play the role of context.

Another precipitating event that causes the return of extinguished or punished behavior is free presentations of the reinforcer. In the rat laboratory, when a few reinforcers are presented freely after an operant response has been extinguished, the behavior will return (e.g., Reid, 1958; Rescorla & Skucy, 1969; Ostlund & Balleine, 2007). We recently found that even reinforcers presented contingent on a second response (instead of freely) were also effective at reinstating an extinguished target response (Winterbauer & Bouton, 2011). One explanation is that reinforcer presentations are merely part of the background (or context) that sets the occasion for more responding (e.g., Ostlund & Balleine, 2007). Another is that the reinforcer presentations may condition the background context (e.g., the Skinner box), which might itself re-invigorate the extinguished behavior (Baker, Steinwald, & Bouton, 1991). Presentations of the reinforcer after extinction are also known to reinstate operant behaviors that are reinforced by drugs (e.g., de Wit & Stewart, 1981, 1983). And reinstatement effects have been reported with operants that have been suppressed by punishment (Panlilio, Thorndike, & Schindler, 2003). Once again, low- or zero-rate operants that have achieved that level through extinction or punishment can return through the occurrence of a precipitating event.

Recent research on extinction has focused on another “relapse” effect that may be especially relevant to behavior analysis. In differential reinforcement of alternative behavior (DRA), a new operant (R2) is reinforced at the same time a previously-reinforced target operant (R1) is extinguished. The DRA procedure is widely used by behavior analysts in treating problem behavior in children with developmental disabilities or autism spectrum disorder, and it seems reasonable to think that in nature extinguished responses are also often replaced by reinforced alternative responses. However, if the alternative behavior is now put on extinction, the first behavior can return (e.g., Leitenberg et al., 1970). This *resurgence* effect is now being studied in several laboratories. Although several explanations have been suggested (Leitenberg et al., 1970; Shahan & Craig, 2017; Shahan & Sweeney, 2011), most evidence favors a view that is based on the principles described above: Removal of

reinforcement for R2 during testing changes the reinforcer context, so the inhibited R1 renews (e.g., Winterbauer & Bouton, 2010). A number of findings support this hypothesis. Perhaps the most straightforward is that when R1 is reinforced with one reinforcer and R2 is reinforced with a different one while R1 is extinguished, free (noncontingent) presentations of the *second* reinforcer during testing eliminates resurgence, whereas presentation of the *first* reinforcer does not (Bouton & Trask, 2016). Thus, the second reinforcer is demonstrably a cue or context that in this case controls R1's extinction performance (see also Trask & Bouton, 2016). When the reinforcer is removed, the response returns; when it is maintained, the response remains inhibited and suppressed. The key to preventing resurgence is thus to encourage generalization between treatment and testing. For a more complete review of resurgence and the evidence supporting the context explanation, see Trask, Schepers, and Bouton (2015).

It is worth mentioning that much of the early research on renewal, reinstatement, and spontaneous recovery was actually done in Pavlovian (respondent) conditioning (e.g., Bouton, 1988, 2004, 2017). (Resurgence has been studied exclusively in the operant domain, although we expect that an analogous effect would occur in respondent conditioning; e.g., see Lindblom & Jenkins, 1981). The characteristics of extinguished respondents and operants are thus highly similar. We know that a conditional stimulus (CS) that elicits a very weak response or even no response at all can elicit the response again with a change of context (renewal), presentation of the Pavlovian unconditional stimulus or reinforcer (reinstatement), or the passage of time (spontaneous recovery); see Bouton (2017) for one recent review. Renewal, reinstatement, and spontaneous recovery have also been shown after counterconditioning, in which (for example) CS-shock pairings are followed by CS-food pairings or CS-food is followed by CS-shock (e.g., Bouton & Peck, 1992; Brooks, Hale, Nelson, & Bouton, 1995; Holmes, Leung, & Westbrook, 2016; Peck & Bouton, 1990). The effects of the precipitating events in turn depend on the CS's conditioning history.

In a clear illustration of this, Bouton (1984, Experiment 5) paired a CS with a strong footshock in an initial phase. He then extinguished the response ("fear") that the CS elicited by presenting the CS several times without shock. Importantly, extinction was stopped before the response completely disappeared. A second set of rats received only a small number of pairings of the CS with a weak shock-- without extinction. They received just enough conditioning trials so that the CS elicited a low level of responding that was indistinguishable from that in the conditioned-then-partially-extinguished group. Then both groups received footshock presentations (a reinstatement treatment). When the CS was then tested, there was a robust increase in responding to the conditioned-then-extinguished CS, but no increase to the conditioned-only CS. Thus, the effects of the reinstating shocks (and the conditioning of the context they demonstrably produced) depended crucially on the conditioning history of the CS. One could not have predicted the effect of the reinstatement shocks from prior responding to the CS alone. In respondent conditioning, like operant conditioning, behavioral silence can be misleading. What you see is not all there is.



## Conclusions

To summarize the argument, it may be easiest to predict what events will control a particular behavior by first considering the behavior's history. Silent operants (those that have been inhibited, for example, by extinction or punishment) may "inexplicably" seem to increase in strength after context change, the passage of time, reinforcer presentation, or the extinction of an alternative behavior that replaced it. Non-extinguished or non-punished behaviors would not change in response to these events in the same way. To predict or control future behavior, as well as predict the effects of context change, the passage of time, reinforcer presentation, or the extinction of an alternative, one would need to know if the behavior has been inhibited or suppressed below some earlier rate. Goal-directed actions, on the other hand, can be strengthened or weakened by degradation of the action-outcome contingency or by offline changes in the value of the reinforcing outcome—as shown in the reinforcer devaluation effect. In contrast, habits are operants that are relatively immune to the immediate effects of reinforcer devaluation and in reducing the action-outcome contingency. Actions and habits (like silent and non-silent operants) can be distinguished by the kinds of events that control them. But knowing *ahead of time* whether a behavior is an action or habit will again depend on understanding the behavior's history: actions have usually had little practice, have developed in choice situations, or have been reinforced on ratio schedules. Habits have contrastingly had larger amounts of practice, and have often been reinforced on interval schedules.

We already know that a behavior's history is important. For example, it is widely understood that a behavior that has been intermittently reinforced can be more resistant to extinction than one that has been consistently reinforced (e.g., Capaldi, 1967). However, the concepts of actions, habits, and inhibited or silent operants may have unique implications for the control and treatment of problem behavior. For one, if a problem behavior is a habit (rather than an action), current thinking suggests that it might be difficult to change. And of course, extinction (and punishment) treatments, including DRA or DRO, can be vulnerable to lapse and relapse effects (e.g., Bouton, 2014; Podlesnik et al., 2017). On the more positive side, extinction of an undesirable behavior might in principle allow a more desirable one (that was previously learned and replaced by an undesirable one) to resurge. And reinforcement of a positive new operant to the point of becoming a habit might make it especially resistant to change. These ideas and applications beg for more research to give them more richness and detail.

More generally, it is proposed that the accuracy and precision of prediction and control, the behavior analyst's two main desiderata, will be enhanced by considering the basic research and principles just reviewed. Moreover, prediction and control are also crucially enhanced by understanding a behavior's history. What you see is not necessarily all there is.

## Acknowledgement.

Preparation of this paper was supported by NIH Grant RO1 DA 033123 to MEB and by both a grant from the Australian Research Council (DP150104878) and a Senior Principle Research Fellowship from the National Health and Medical Research Council of Australia (GNT1079561) to BWB. We thank Eric Thraillkill for comments.

## References

- Adams CD & Dickinson A (1981). Instrumental responding following reinforcer devaluation. *Quarterly Journal of Experimental Psychology*, 33B, 109–122.
- Adams CD (1982) Variations in the sensitivity of instrumental responding to reinforcer devaluation. *Quarterly Journal of Experimental Psychology* 33B,109–122.
- Alvares GA, Balleine BW, Whittle L & Guastella AJ (2016). Reduced goal-directed action control in autism spectrum disorder. *Autism Research*, 9, 1285–1293. [PubMed: 26999793]
- Alvares GA, Balleine BW, Guastella AJ (2014) Impaired goal-directed actions in social anxiety disorder predicts treatment response to cognitive-behavioral therapy. *PLoS One*, 9, e94778. [PubMed: 24728288]
- Baker AG, Steinwald H, & Bouton ME (1991). Contextual conditioning and reinstatement of extinguished instrumental responding. *Quarterly Journal of Experimental Psychology*, 43B, 199–218.
- Balleine BW (2001). Incentive processes in instrumental conditioning In Mowrer R & Klein S (Eds) *Handbook of Contemporary Learning Theories* (pp 307–366). Hillsdale, NJ: LEA.
- Balleine BW & Dickinson A (1998). Goal-directed instrumental action: Contingency and incentive learning and their cortical substrates. *Neuropharmacology*, 37, 407–419. [PubMed: 9704982]
- Balleine BW, O’Doherty JP (2010). Human and rodent homologies in action control: corticostriatal determinants of goal-directed and habitual action. *Neuropsychopharmacology*, 35, 48–69. [PubMed: 19776734]
- Bouton ME (1984). Differential control by context in the inflation and reinstatement paradigms. *Journal of Experimental Psychology: Animal Behavior Processes*, 10, 56–74.
- Bouton ME (1988). Context and ambiguity in the extinction of emotional learning: Implications for exposure therapy. *Behaviour Research and Therapy*, 26, 137–149. [PubMed: 3365204]
- Bouton ME (2002). Context, ambiguity, and unlearning: Sources of relapse after behavioral extinction. *Biological Psychiatry*, 52, 976–986. [PubMed: 12437938]
- Bouton ME (2004). Context and behavioral processes in extinction. *Learning & Memory*, 11, 485–494. [PubMed: 15466298]
- Bouton ME (2014). Why behavior change is difficult to sustain. *Preventive Medicine*, 68, 29–36. [PubMed: 24937649]
- Bouton ME (2017) Extinction: Behavioral mechanisms and their implications In: Menzel R (ed.), *Learning Theory and Behavior*, Vol. 1 of *Learning and Memory: A Comprehensive Reference*, 2nd edition, Byrne, J.H. (ed.). pp. 61–83. Oxford: Academic Press.
- Bouton ME, Kenney FA, and Rosengard C (1990). State-dependent fear extinction with two benzodiazepine tranquilizers. *Behavioral Neuroscience*, 104, 44–55. [PubMed: 2317285]
- Bouton ME, & Peck CA (1992). Spontaneous recovery in cross-motivational transfer (counterconditioning). *Animal Learning & Behavior*, 20, 313–321.
- Bouton ME, & Schepers ST (2015). Renewal after the punishment of free operant behavior. *Journal of Experimental Psychology: Animal Learning and Cognition* 41, 81–90. [PubMed: 25706548]
- Bouton ME, Todd TP, Vurbic D, & Winterbauer NE (2011). Renewal after the extinction of free operant behavior. *Learning & Behavior*, 39, 57–67. [PubMed: 21279496]
- Bouton ME, & Trask S (2016). Role of the discriminative properties of the reinforcer in resurgence. *Learning & Behavior*, 44, 137–150. [PubMed: 26486932]
- Bouton ME, Trask S, & Carranza-Jasso R (2016). Learning to inhibit the response during instrumental (operant) extinction. *Journal of Experimental Psychology: Animal Learning and Cognition*, 42, 246–258.. [PubMed: 27379715]
- Bradfield L & Balleine BW (2013). Hierarchical and binary associations compete for behavioral control during instrumental biconditional discrimination. *Journal of Experimental Psychology: Animal Behavior Processes*, 39, 2–13. [PubMed: 23316974]
- Brooks DC, Hale B, Nelson JB, & Bouton ME (1995). Reinstatement after counterconditioning. *Animal Learning & Behavior*, 23, 383–390.

- Capaldi EJ (1967). A sequential hypothesis of instrumental learning In: Spence KW and Spence JT (eds.) *Psychology of Learning and Motivation*, vol. 1, pp. 67–156. New York: Academic Press.
- Colwill RM & Rescorla RA (1985). Postconditioning devaluation of a reinforcer affects instrumental responding. *Journal of Experimental Psychology: Animal Behavior Processes*, 11, 120–132.
- Colwill RC & Rescorla RA (1986). Associative structures in instrumental learning In Bower GH (Ed.), *The psychology of learning and motivation* (Vol. 20, pp.55–104). New York: Academic Press.
- Colwill RM, & Rescorla RA (1988). The role of response–reinforcer associations increases throughout extended instrumental training. *Animal Learning & Behavior*, 16, 105–111.
- Crespi LP (1942). Quantitative variation in incentive and performance in the white rat. *American Journal of Psychology*, 55, 467–517.
- Cunningham CL (1979). Alcohol as a cue for extinction: State dependency produced by conditioned inhibition. *Animal Learning & Behavior*, 7, 45–52.
- De Wit H, & Stewart J (1981). Reinstatement of cocaine-reinforced responding in the rat. *Psychopharmacology*, 75, 134–143. [PubMed: 6798603]
- De Wit H, & Stewart J (1983). Drug reinstatement of heroin-reinforced responding in the rat. *Psychopharmacology*, 79, 29–31. [PubMed: 6403961]
- Dezfouli A & Balleine BW (2012). Habits, action sequences, and reinforcement learning. *European Journal of Neuroscience*, 35, 1036–1051 [PubMed: 22487034]
- Dickinson A & Mulatero CW (1989). Reinforcer specificity of the suppression of instrumental performance on a non-contingent schedule. *Behavioural Processes*, 19, 167–180. [PubMed: 24895910]
- Dickinson A & Balleine BW (1994). Motivational control of goal-directed action. *Animal Learning & Behavior*, 22, 1–18.
- Dickinson A, Balleine BW, Watt A, Gonzales F & Boakes RA (1995). Motivational control after extended instrumental training. *Animal Learning & Behavior*, 23, 197–206.
- Dickinson A, Squire S, Varga Z, Smith JW (1998). Omission learning after instrumental pretraining. *Quarterly Journal of Experimental Psychology*, 51B, 271–286.
- Estes WK (1944). An experimental study of punishment. *Psychological Monographs: General and Applied*, 57, i–40.
- Everitt BJ, & Robbins TW (2005). Neural systems of reinforcement for drug addiction: from actions to habits to compulsion. *Nature Neuroscience*, 8, 1481–1489. [PubMed: 16251991]
- Flaherty CF (1996). *Incentive relativity*. Cambridge: Cambridge University Press.
- Furlong TM, Supit ASA, Corbit LH, Killcross S & Balleine BW (2015). Pulling habits out of rats: Adenosine 2A receptor antagonism in dorsomedial striatum rescues methamphetamine-induced deficits in goal-directed action. *Addiction Biology*, 22, 172–183. [PubMed: 26515740]
- Griffiths KR, Morris RW, Balleine BW. (2014). Translational studies of goal-directed action as a framework for classifying deficits across psychiatric disorders. *Front Syst Neurosci*, 5 26; 8:101. [PubMed: 24904322]
- Hanley GP, Iwata BA, & McCord BE (2003). Functional analysis of problem behavior: A review. *Journal of Applied Behavior Analysis*, 36, 147–185. [PubMed: 12858983]
- Hogarth L, Balleine BW, Corbit LH & Killcross S (2013). Associative learning mechanisms underpinning the transition from recreational drug use to addiction. *Annals of the New York Academy of Science*, 1282:12–24.
- Holmes NM, Leung HT, & Westbrook RF (2016). Counterconditioned fear responses exhibit greater renewal than extinguished fear responses. *Learning & Memory*, 23, 141–150. [PubMed: 26980781]
- Hull CL (1943). *Principles of behavior: An introduction to behavior theory*. New York: Appleton-Century-Crofts.
- Kahneman D (2011). *Thinking fast and slow*. New York: Farrar, Straus & Groux.
- Kelley ME, Liddon CJ, Ribeiro A, Greif AE, & Podlesnik CA (2015). Basic and translational evaluation of renewal of operant responding. *Journal of Applied Behavior Analysis*, 48, 390–401. [PubMed: 25891414]

- Klossek UM, Russell J & Dickinson A (2008). The control of instrumental action following outcome devaluation in young children aged between 1 and 4 years. *Journal of Experimental Psychology: General*, 137, 39–51. [PubMed: 18248128]
- Kosaki Y & Dickinson A (2010). Choice and contingency in the development of behavioral autonomy during instrumental conditioning. *Journal of Experimental Psychology: Animal Behavior Processes*, 36, 334–342. [PubMed: 20658864]
- Lattal KM (2007). Effects of ethanol on the encoding, consolidation, and expression of extinction following contextual fear conditioning. *Behavioral Neuroscience*, 121, 1280–1292. [PubMed: 18085881]
- Leitenberg H, Rawson RA, & Bath K (1970). Reinforcement of competing behavior during extinction. *Science*, 169, 301–303. [PubMed: 5450360]
- Lindblom LL, & Jenkins HM (1981). Responses eliminated by noncontingent or negatively contingent reinforcement recover in extinction. *Journal of Experimental Psychology: Animal Behavior Processes*, 7, 175–190. [PubMed: 7241053]
- Morris RM, Quail S, Griffiths K, Green MJ & Balleine BW (2015). Corticostriatal control of goal-directed action is impaired in schizophrenia. *Biological Psychiatry*, 77, 187–195. [PubMed: 25062683]
- Nakajima S, Urushihara K, & Masaki T (2002). Renewal of operant performance formerly eliminated by omission or non-contingency training upon return to the acquisition context. *Learning and Motivation*, 33, 510–525.
- Ostlund SB, & Balleine BW (2007). Selective reinstatement of instrumental performance depends on the discriminative stimulus properties of the mediating outcome. *Learning & Behavior*, 35, 43–52. [PubMed: 17557390]
- Overton DA (1985). Contextual stimulus effects of drugs and internal states In: Balsam PD and Tomie A (eds.) *Context and learning*, pp. 357–384. Hillsdale, NJ: Erlbaum.
- Panlilio LV, Thorndike EB, & Schindler CW (2003). Reinstatement of punishment-suppressed opioid self-administration in rats: An alternative model of relapse to drug abuse. *Psychopharmacology*, 168, 229–235. [PubMed: 12845420]
- Peck CA, & Bouton ME (1990). Context and performance in aversive-to-appetitive and appetitive-to-aversive transfer. *Learning and Motivation*, 21, 1–31.
- Podlesnik CA, Kelley ME, Jimenez-Gomez C, & Bouton ME (2017). Renewed behavior caused by context change and its implications for treatment maintenance: A review. *Journal of Applied Behavior Analysis*, 50, 675–697. [PubMed: 28608584]
- Redgrave P, Rodriguez M, Smith Y, Rodriguez-Oroz MC, Lehericy S, Bergman H, Agid Y, DeLong MR, Obeso JA (2010). Goal-directed and habitual control in the basal ganglia: implications for Parkinson's disease. *Nature Reviews Neuroscience*, 11, 760–72. [PubMed: 20944662]
- Reid RL (1958). The role of the reinforcer as a stimulus. *British Journal of Psychology*, 49, 202–209.
- Rescorla RA (1991) Associative relations in Instrumental learning: The Eighteenth Bartlett Memorial Lecture. *Quarterly Journal of Experimental Psychology*, 43B, 1–23.
- Rescorla RA (2004). Spontaneous recovery. *Learning & Memory*, 11, 501–509. [PubMed: 15466300]
- Rescorla RA, & Skucy JC (1969). Effect of response-independent reinforcers during extinction. *Journal of Comparative and Physiological Psychology*, 67, 381–389.
- Schepers ST, & Bouton ME (2017). Hunger as a context: Food-seeking that is inhibited while hungry renews in the context of satiation. *Psychological Science*, 28, 1640–1648. [PubMed: 28957015]
- Shahan TA, & Craig AR (2017). Resurgence as choice. *Behavioural Processes*, 141, 100–127. [PubMed: 27794452]
- Shahan TA, & Sweeney MM (2011). A model of resurgence based on behavioral momentum theory. *Journal of the Experimental Analysis of Behavior*, 95, 91–108. [PubMed: 21541118]
- Skinner BF (1938). *The behavior of organisms*. New York: Appleton.
- Skinner BF (1969). *Contingencies of reinforcement*. New York: Appleton.
- Spence KW (1956). *Behavior theory and conditioning*. New Haven, CT: Yale University Press.
- Thrailkill EA & Bouton ME (2015). Contextual control of instrumental actions and habits. *Journal of Experimental Psychology: Animal Learning & Cognition*, 41, 69–80. [PubMed: 25706547]

- Thraill EA, & Bouton ME (2016a). Extinction and the associative structure of heterogeneous instrumental chains. *Neurobiology of Learning and Memory*, 133, 61–68. [PubMed: 27296700]
- Thraill EA, Trott JM, Zerr CL, & Bouton ME (2016). Contextual control of chained instrumental behaviors. *Journal of Experimental Psychology: Animal Learning and Cognition*, 42, 401–414. [PubMed: 27598060]
- Todd TP (2013). Mechanisms of renewal after the extinction of instrumental behavior. *Journal of Experimental Psychology: Animal Behavior Processes*, 39, 193–207. [PubMed: 23627796]
- Todd TP, Vurbic D, & Bouton ME (2014b). Mechanisms of renewal after the extinction of discriminated operant behavior. *Journal of Experimental Psychology: Animal Learning and Cognition*, 40, 355–368. [PubMed: 25545982]
- Trask S, & Bouton ME (2014). Contextual control of operant behavior: Evidence for hierarchical associations in instrumental learning. *Learning & Behavior*, 42, 281–288. [PubMed: 24961749]
- Trask S, & Bouton ME (2016). Discriminative properties of the reinforcer can be used to attenuate the renewal of an extinguished instrumental behavior. *Learning & Behavior*, 44, 151–161. [PubMed: 26400498]
- Trask S, Schepers ST, & Bouton ME (2015). Context change explains resurgence after the extinction of operant behavior. *Mexican Journal of Behavior Analysis*, 41, 187–210. [PubMed: 27429503]
- Tinklepaugh OL (1928). An experimental study of representative factors in monkeys. *Journal of Comparative Psychology*, 8, 197–236.
- Tolman EC (1932). *Purposive behavior in animals and men*. New York: The Century Co.
- Tricomi E, Balleine BW & O'Doherty JP (2009). A specific role for posterior dorsolateral striatum in human habit learning. *European Journal of Neuroscience*, 29, 2225–2232. [PubMed: 19490086]
- Winterbauer NE, & Bouton ME (2010). Mechanisms of resurgence of an extinguished instrumental behavior. *Journal of Experimental Psychology: Animal Behavior Processes*, 36, 343–353. [PubMed: 20658865]
- Winterbauer NE, & Bouton ME (2011). Mechanisms of resurgence II: Response-contingent reinforcers can reinstate a second extinguished behavior. *Learning and Motivation*, 42, 154–164. [PubMed: 21479120]