Flexible control of Pavlovian-instrumental transfer based on expected reward value

Andrew T. Marshall[1], Briac Halbout[2], Christy N. Munson[2], Collin Hutson[2], and Sean B. Ostlund[2]

1. Children's Hospital Los Angeles, and the Department of Pediatrics, University of Southern California, Los Angeles, CA, USA
2. Department of Anesthesiology and Perioperative Care, Department of Neurobiology and Behavior, Irvine Center for Addiction Neuroscience (ICAN), Center for the Neurobiology of Learning and Memory (CNLM), Center for Neural Circuit Mapping (CNCM), University of California, Irvine School of Medicine, Irvine, CA, USA

Correspondence Dr Sean Ostlund, Department of Anesthesiology & Perioperative Care, University of California, Irvine School of Medicine, Irvine, CA 92697, USA. E-mail: sostlund@uci.edu

Acknowledgements

The authors are supported by NIH grants DA046667 (SBO), MH12685 (SBO), MH106972 (SBO), DA050116 (BH).

Abstract

The Pavlovian-instrumental transfer (PIT) paradigm is widely used to assay the motivational influence of reward-predictive cues, reflected by their ability to invigorate instrumental behavior. Leading theories of incentive learning assume that cues are assigned motivational properties based on the value of the rewards they predict. We outline an alternative account of PIT which recognizes that reward-predictive cues may suppress rather than motivate instrumental behavior under certain conditions. We posit that cues capable of triggering a strong reward expectancy will tend to inhibit instrumental behavior, which is exploratory by nature, in order to facilitate retrieval of the expected reward before it is lost or stolen. According to this view, the motivation to engage in instrumental behavior during a cue should be *inversely* related to the value of the predicted reward, since there is more to lose by failing to secure a high-value reward than a low-value reward. We tested this hypothesis in rats using a high-conflict PIT protocol which elicits competing instrumental and food-port entry responses. In Experiment 1, cues signaling different reward magnitudes (1, 3, or 9 pellets) differentially influenced these behaviors. Whereas the 1-pellet cue increased instrumental behavior, cues signaling 3 or 9 pellets suppressed instrumental behavior and elicited high levels of food-port activity. Experiment 2 found that reward-predictive cues suppressed instrumental behavior and increased food-port activity in a flexible manner that was disrupted by post-training reward devaluation. We discuss how the PIT task provides a useful tool for studying cognitive control over cue-motivated behavior in rodents.

Reward-paired cues acquire motivational properties which allow them to invigorate instrumental reward-seeking behavior (Dickinson, Smith, & Mirenowicz, 2000; Estes, 1948; Lovibond, 1981), a phenomenon referred to as Pavlovian-instrumental transfer (PIT). This motivational influence seems to serve an adaptive function by promoting risky and effortful foraging activity in environments that signal potential reward availability. The PIT paradigm is widely used to study the mechanisms of cue-motivated behavior (Cartoni, Balleine, & Baldassarre, 2016; Corbit & Balleine, 2016) and how they contribute to pathological reward seeking in addiction and related disorders (Corbit & Janak, 2007, 2016; Garbusow et al., 2016; Genauck et al., 2020; LeBlanc, Maidment, & Ostlund, 2013, 2014; LeBlanc, Ostlund, & Maidment, 2012; Marshall & Ostlund, 2018; Ostlund, LeBlanc, Kosheleff, Wassum, & Maidment, 2014; Saddoris, Stamatakis, & Carelli, 2011; Sebold et al., 2021; Shiflett, 2012; Shiflett, Riccie, & DiMatteo, 2013; Vogel et al., 2018; Wyvell & Berridge, 2001). However, despite decades of research, much remains unclear about how fundamental variables such as expected reward value influence expression of the PIT effect.

Leading computational models of incentive learning (Dayan & Balleine, 2002; McClure, Daw, & Montague, 2003; Zhang, Berridge, Tindell, Smith, & Aldridge, 2009) assume that motivational value is assigned to cues based on the total amount of delay-discounted reward that they predict (i.e., their state value). The motivational influence of cues should therefore directly depend on basic Pavlovian conditioning parameters such as reward probability, cue-reward interval, and reward magnitude. This account makes some intuitive predictions. For instance, a cue that reliably signals the immediate

delivery of a large reward should acquire strong motivational properties, whereas a cue signaling that an upcoming reward will be small, delayed, or unlikely to occur at all, should acquire weak motivational properties.

However, in contrast to these predictions, there is growing evidence that the motivational impact of reward-paired cues is instead *inversely* related to their ability to predict reward. For instance, cues signaling a high probability of imminent reward do not invigorate and may even suppress instrumental performance (Azrin & Hake, 1969; Crombag, Galarce, & Holland, 2008; Lovibond, 1981; Vandyne, 1971), which is replaced by Pavlovian conditioned approach behavior directed at the food port (Marshall, Munson, Maidment, & Ostlund, 2020). Instead, it is cues that signal a low probability of reward that become potent motivators of instrumental behavior, while eliciting more modest levels of food-port activity (Marshall et al., 2020). The motivational influence of cues also varies with their temporal relationship with reward delivery (Delamater & Holland, 2008; Delamater & Oakeshott, 2007; Lovibond, 1981; Matell & Della Valle, 2018). For instance, we have shown that cues signaling a fixed 30-sec interval between cue onset and reward delivery produce a gradual suppression of instrumental behavior and a coincident increase in food-port activity as the expected reward delivery time draws near (Marshall & Ostlund, 2018).

These findings suggest that reward-paired cues can acquire distinct motivational and predictive properties that evoke different kinds of behavior, with the former promoting the pursuit of new rewards through instrumental behavior and the latter eliciting

attempts to retrieve a reward that is expected imminently (Ostlund & Marshall, 2021). Organizing behavior in this manner, based on reward expectancy, is important for efficient foraging and is a central pillar of behavior systems theory (Timberlake, Wahl, & King, 1982). While it is adaptive to seek out new rewards through instrumental behavior (or other general search activities) when rewards are scarce, such behavior is unnecessary and may even interfere with the retrieval of a reward that is expected soon (i.e., focal search), increasing the chance that it will be pilfered or otherwise lost.

However, this simplistic description of foraging sidesteps the complexity involved in ambiguous situations, when cues may elicit conflicting tendencies to both seek out new rewards and collect an expected reward. The factors involved in resolving such conflict are not well understood, though presumably the value of the expected reward plays an important role, as discussed below. Early studies on the interaction between Pavlovian and instrumental learning processes may also provide some insight. For instance, Konorski and colleagues demonstrated through a series of studies that Pavlovian cues predicting imminent reward do not simply elicit conditioned consummatory responses – in their case, salivation and orienting toward the food cup – they also acquire the ability to acutely interrupt ongoing instrumental behavior (Ellison & Konorski, 1965; Konorski, 1967). Moreover, such cues were able to prevent other discriminative cues from motivating instrumental performance (Soltysik, Konorski, Holownia, & Rentoul, 1976) and were themselves extremely resistant to acquiring motivational properties if later used as discriminative stimuli for instrumental performance (Konorski & Wyrwicka,

1950). These findings suggest that cues signaling imminent, response-independent reward actively inhibit the expression of cue-motivated behavior.

Following this logic, we have hypothesized that Pavlovian cues acquire the potential to motivate instrumental behavior, but that this motivational response is subject to cognitive control and will therefore be suppressed in situations where it would be disadvantageous (Marshall et al., 2020; Marshall & Ostlund, 2018; Ostlund & Marshall, 2021). Cognitive control broadly refers to the ability to regulate thoughts, emotions, and actions based on an internally represented behavioral goal (Braver, 2012). A core function of cognitive control is to override prepotent response tendencies – particularly those which are impulsive or habitual in nature – in situations where an alternative course of action is more advantageous (Botvinick & Braver, 2015; Miller & Cohen, 2001). We can apply this framework to PIT by assuming that reward-predictive cues signal the value of searching for reward at the food port, which if sufficiently high will suppress the impulse to engage in instrumental reward-seeking behavior. Thus, the value of the reward predicted by a cue should play an important albeit indirect role in modulating instrumental performance, indicating the degree to which such behavior should be suppressed to facilitate reward retrieval. A cue that signals imminent delivery of a high-value reward should therefore be less effective in motivating instrumental behavior and more effective in eliciting food-port activity than a cue that signals a less valuable reward. These predictions are readily distinguished from those made by theories of incentive learning that assume a positive correlation between motivational and predictive value (Dayan & Balleine, 2002; McClure et al., 2003; Zhang et al., 2009).

Interesting, the limited data that exist on this subject suggests that expected reward value rarely affects the expression of PIT performance. For instance, several studies have shown that cues retain their ability to motivate instrumental behavior despite predicting a reward that has been devalued through conditioned taste aversion learning (Colwill & Rescorla, 1990; Holland, 2004; Rescorla, 1994). Such studies have generally used PIT protocols designed to maximize motivational effects by using cues that are relatively poor predictors of reward delivery. Therefore, these findings suggest that cues normally motivate instrumental behavior through an implicit process, without considering expected reward value, but do not address whether expected reward value modulates the suppressive influence of cues that signal imminent reward.

We investigated this issue using an alternative, “high-conflict” version of PIT in which reward-predictive cues tend to suppress instrumental behavior and increase food-port activity in apparent anticipation of reward delivery (Marshall & Ostlund, 2018). The role of expected reward value on these responses was assessed using a parametric manipulation of reward magnitude in Experiment 1 and post-training reward devaluation in Experiment 2.

Experiment 1

Experiment 1 investigated the influence of expected reward magnitude on the expression of instrumental reward-seeking and food-port entry behavior during a high-conflict PIT task. Briefly, hungry rats were first trained to lever press for food pellets

before undergoing Pavlovian conditioning, in which two distinct 30-sec auditory cues signaled food pellet delivery at cue offset. Reward magnitude was varied across cues and groups. While all groups had at least one cue that signaled three food pellets, the alternate cue signaled either one (Group 1/3), three (Group 3/3), or nine food pellets (Group 3/9). PIT testing was then performed by intermittently presenting these reward-predictive cues in a noncontingent manner while rats were free to press the lever and enter the food port, in the absence of reward delivery.

Methods

Animals and Apparatus

Twenty-six experimentally naïve adult male Long Evans rats (Envigo) were used in this experiment. They arrived at the facility (University of California, Irvine; Irvine, CA, USA) at approximately 10 weeks of age. They were pair-housed in a colony room set to a standard 12:12 hr light:dark schedule. Rats were tested during the light phase. Water was always provided ad libitum in the home cages. Rats were fed between 10-14 g of standard lab chow per day during the experiment to maintain them at ~85% of their estimated free-feeding bodyweight. Husbandry and experimental procedures were approved by the UC Irvine Institutional Animal Care and Use Committee (IACUC) and conducted in accordance with the National Research Council Guide for the Care and Use of Laboratory Animals.

The experiment was conducted in 14 operant chambers (Med-Associates; St. Albans, VT), each housed within sound-attenuating, ventilated boxes. Each chamber was

equipped with a stainless-steel grid floor; two stainless steel walls (front and back); and a transparent polycarbonate side-wall, ceiling, and door. Pellet dispensers, mounted on the outside of the operant chamber, were equipped to deliver 45-mg food pellets (Bio-Serv) into a recessed food port centered on the lower section of the front wall. Head entries into the food port were transduced by an infrared photobeam. A retractable lever was located to the left of the food port, on the front wall. The chamber was also equipped with a house light centered at the top of the back wall. Auditory stimuli were presented to animals via a speaker located on the back wall. Experimental events were controlled and recorded with 10-ms resolution by the software program MED-PC IV (Tatham & Zurn, 1989).

Procedure

Magazine training. All sessions of all phases began with the onset of the houselight. In each of two 30-minute sessions of magazine training, grain-based food pellets were delivered on a random-time (RT) 60-s schedule.

Instrumental training. During initial instrumental (lever-press) training, rats were continuously reinforced with a grain-based food pellet delivery for pressing the left lever (fixed-ratio, FR-1), earning a maximum of 30 pellets per session. These FR-1 sessions lasted no more than 30 min. Seven rats required an extra session of FR-1 training, which lasted until these rats earned 30 pellets. During subsequent training sessions, lever pressing was reinforced according to a random-interval (RI) schedule, such that the lever remained available but was inactive for an average of *t* seconds after each

reward delivery, where individual *t* values were selected randomly from an exponential distribution. The RI schedule was changed over training days with 1 day of RI-5 (*t* = 5 sec), 1 day of RI-15 (*t* = 15 sec), 2 days of RI-30 (*t* = 30 sec), and 10 days of RI-45 (*t* = 45 sec) training. Each RI session lasted 30 minutes.

Pavlovian training. Pavlovian training involved exposure to two 30-s conditioned stimuli (CS; 3-kHz tone and 10-Hz clicker) paired with reward (grain-based food pellets). Rats were assigned to one of three groups with different CS-reward magnitude arrangements. For Group 1/3 (*n* = 9), one CS terminated with 1 pellet and the other with 3 pellets. For Group 3/3 (*n* = 8), both CSs terminated with 3 pellets. For Group 3/9 (*n* = 9), one CS terminated with 3 pellets and the other with 9 pellets.

In each 20-min session, a 60-s interval preceded onset of the first CS. There was a 120-s inter-stimulus interval (ISI) between consecutive CS presentations (i.e., between previous CS offset and subsequent CS onset), and a 60-s interval following the final CS presentation prior to the end of the session. Pavlovian training lasted for 25 sessions, each involving 4 pseudorandomly-alternating presentations of each CS (8 total trials per session).

Pavlovian-to-instrumental transfer (PIT). Following Pavlovian training, rats received two sessions of instrumental retraining (identical to earlier sessions with RI45 schedule) followed by one 30-min session of instrumental extinction, in which the lever was continuously available but was inactive. Rats then received a PIT test session (30.25

minutes in duration), during which the lever was once again continuously available but inactive. During the test, rats received 4 noncontingent presentations of each 30-s CS in pseudorandom order (ABBABAAB). The ISI was 150 s, and a 5-min interval preceded onset of the first CS. No food pellets were delivered at test.

Data Analysis

All summary measures were obtained from the raw data using MATLAB (The MathWorks; Natick, MA, USA), and analyzed with mixed-effects regression models (Pinheiro & Bates, 2000), a powerful analytical framework that is both well established and highly recommended for behavioral research (Boisgontier & Cheval, 2016). Mixed-effects models are comparable to repeated-measures regression analyses, and allow for parameter estimation per manipulation condition (fixed effects) and the individual (random effects) (Bolker et al., 2009; Hoffman & Rovine, 2007; Pinheiro & Bates, 2000; Schielzeth, Nakagawa, & Freckleton, 2013). Mixed-effects regression models (1) effectively handle missing data and (2) permit the inclusion of categorical and continuous predictors in the same analysis, thus allowing detection of group-level changes across ordered data samples (i.e., continuous time points) while also accounting for corresponding individual differences. All relevant fixed-effects factors were included in each model. Categorical predictors were effects-coded (i.e., codes sum to 0), and continuous predictors were mean-centered. For analyses of Pavlovian training and PIT, the fixed-effects structure included main effects of group and reward magnitude, and the random-effects structure included a by-rat random intercept. Instrumental training analyses incorporated generalized linear mixed-effects models

(family: gamma, link: log) with predictors of group and time since the previous reward delivery. The random-effects structure included a random by-rat slope of time since reward delivery and a random intercept of rat, restricted to be uncorrelated. The alpha level for all tests was .05. Sample sizes were not predetermined based on statistical analysis but are similar to those reported in previous publications (Halbout et al., 2019; Marshall et al., 2020; Marshall & Ostlund, 2018). Main effects and interactions are reported as the results of ANOVA *F*-tests (i.e., whether the coefficients for each fixed effect were significantly different from 0).

Our primary dependent measures were lever pressing and food-port entry behavior. We quantified cue-induced changes in behavior by subtracting the mean response rate during pre-CS periods (30 sec each) from the mean response rate during CS periods (30 sec each). Pre-CS (baseline) data were averaged across all CS trials (within subject). Food-port entry behavior can fall into two categories: spontaneous entries and entries that are performed as part of an instrumental *press-entry* action sequence (Halbout et al., 2019; Marshall et al., 2020; Marshall & Ostlund, 2018), the latter indicated by an increased likelihood of food-port entry shortly after execution of the lever-press response. To focus our analysis on cue-elicited (and not instrumental) reward-retrieval behavior, we excluded all entries that occurred within a 2.5 sec post-lever press period from our analysis, as in our previous studies (Halbout et al., 2019; Marshall et al., 2020; Marshall & Ostlund, 2018). The final five sessions of Pavlovian training were used to assess conditioned food-port entry behavior during CS trials relative to pre-CS baseline periods. Analyses of instrumental training included the final

three sessions of training. Data points were considered outliers if their values were at least three scaled median absolute deviations from the median (Leys, Ley, Klein, Bernard, & Licata, 2013); for PIT analyses, outliers were based on the rats' mean difference scores within each condition. For the current experiment, 24 individual data points were removed from the instrumental training analysis (i.e., 24 of 1,170 data points [26 rats × 45 bins]).

Results

Instrumental and Pavlovian Training

Rats were first trained to lever press for food reward. While the rats had not yet been assigned to groups for Pavlovian training, Figure 1A shows the groups' mean lever-press rates on an RI-45 s schedule of reinforcement as a function of time since previous reward delivery. Per a generalized linear mixed-effects model (distribution = gamma, link = log) on response rates during the 15-45 s time window, there was neither a main effect of group, $F(1, 786) = 0.21$, $p = .812$, nor a Group × Time interaction, $F(2, 786) = 0.77$, $p = .461$.

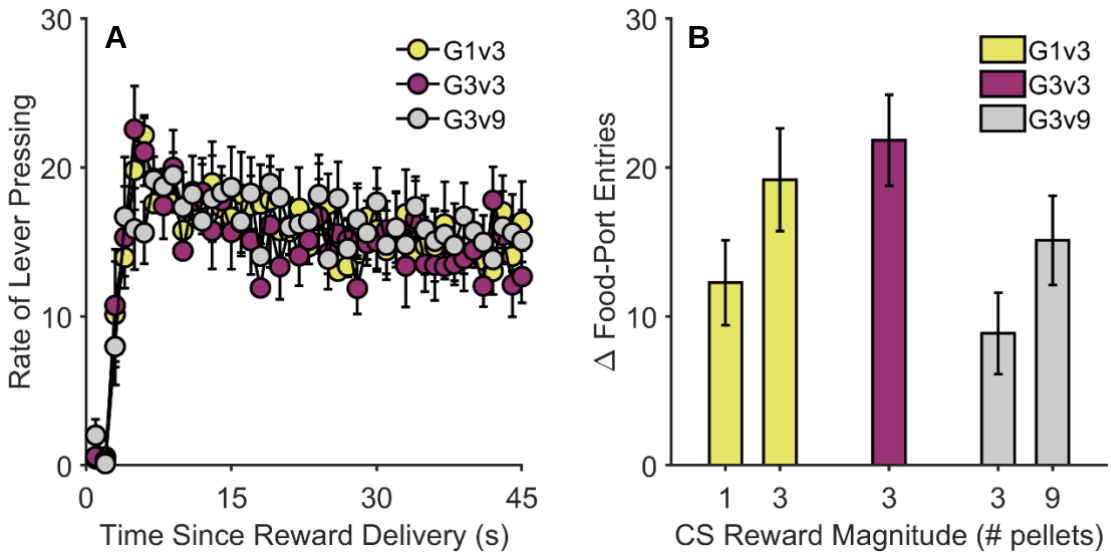

Figure 1. **Instrumental and Pavlovian training in Experiment 1**. (A) Lever-press rates on a random-interval (RI) 45-s schedule of reinforcement as a function of time (in seconds) since previous reward delivery. (B) CS-induced increase in food-port entries relative to the pre-CS baseline. CS = conditioned stimulus. G1v3 = Group 1/3. G3v3 = Group 3/3. G3v9 = Group 3/9. Error bars reflect ± 1 between-subjects standard error of the mean.

Following instrumental training, rats were trained to associate two 30-s CS cues with food reward, delivered upon termination of the cues. Figure 1B shows conditioned food-port entry behavior (relative to the pre-CS baseline) for the final 5 sessions of training. Analysis of these data revealed a main effect of group, $F(2, 40) = 5.69$, $p = .007$, and a main effect of CS reward magnitude, $F(1, 40) = 16.27$, $p < .001$. The latter effect appeared to be driven by Groups 1/3 and 3/9, which showed higher levels of food-port activity during whichever CS signaled the larger of the two possible reward magnitudes.

Pavlovian-instrumental transfer (PIT)

Following two sessions of instrumental retraining and one session of instrumental extinction, rats were given a PIT test, in which the 30-s CSs were presented while the rats were able to freely lever press without reinforcement. Figures 2A and 2B show rats' CS-induced change in lever-press rate and food-port entry rate, respectively, relative to pre-CS baseline periods. Analyses revealed main effects of group, $F(2, 127) = 3.59$, $p = .030$, and CS reward magnitude on lever pressing, $F(1, 127) = 5.74$, $p = .018$, in which CS-induced lever pressing decreased with increases in predicted reward magnitude. For food-port entry rate, there was a main effect of group, $F(2, 127) = 5.26$, $p = .006$, and a significant increase in CS-induced food-port entries with predicted reward magnitude, $F(1, 127) = 8.32$, $p = .005$.

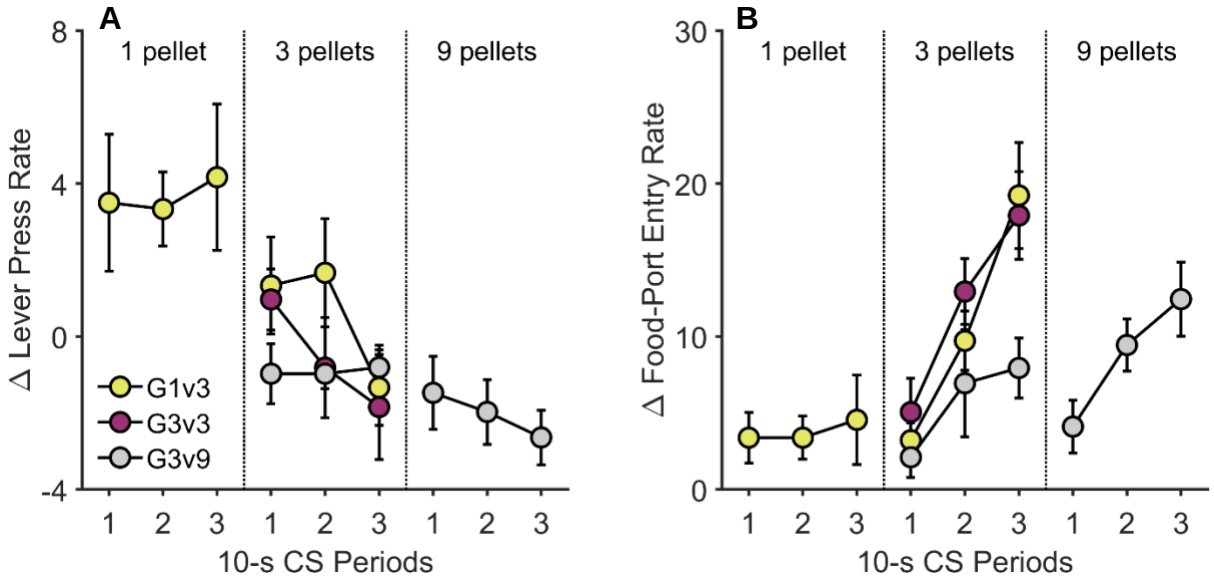

Figure 2. **Pavlovian-instrumental transfer in Experiment 1**. CS-induced changes in (A) lever-press rates and (B) food-port entries relative to the pre-CS baseline on the Pavlovian-instrumental transfer (PIT) test. CS = conditioned stimulus. G1v3 = Group 1/3. G3v3 = Group 3/3. G3v9 = Group 3/9. Error bars reflect ± 1 between-subjects standard error of the mean.

Given the between-groups manipulation of reward magnitude and that one group experienced only one reward magnitude (i.e., Group 3/3), a second set of within-group analyses were conducted to better elucidate how differential reward magnitude influenced CS-induced changes in behavior. For Group 1/3, the 3-pellet CS reduced press rates more than the 1-pellet CS, $F(1, 50) = 8.37$, $p = .006$, and also led to higher levels of food-port entry, $F(1, 50) = 16.66$, $p < .001$. The food-port entry analysis also detected a significant CS Reward Magnitude × CS 10-s Period interaction, $F(1, 50) = 12.67$, $p = .001$, which indicated that the increase in food-port entries over time was steeper for the 3-pellet CS than the 1-pellet CS. Similarly, for Group 3/9, lever pressing was suppressed more by the 9-pellet CS than the 3-pellet CS, $F(1, 50) = 5.36$, $p = .025$. This was paralleled by higher levels of food-port entry during the 9- versus 3-pellet CS, $F(1, 50) = 4.27$, $p = .044$.

Discussion

The results of Experiment 1 indicate that the motivational influence of reward-predictive cues on instrumental behavior varies inversely with expected reward magnitude, with the 1-pellet cue increasing reward seeking and the 3- and 9-pellet cues decreasing this behavior, particularly near the expected time of reward delivery. In contrast, attempts to retrieve reward at the food port were more apparent when larger rewards were expected. It is notable that cue-elicited food-port entries did not monotonically vary with the magnitude of the predicted reward during either Pavlovian training or PIT testing. Importantly, the specific measure we used captures information about how often rats initiate a new bout of food-port activity but is less reflective of the duration of those

bouts. Our informal observations suggest that cues signaling larger rewards tend to elicit more persistent bouts of waiting at the food port than the other cues. Although food-port dwell-times were not recorded in the current study, this measure may more precisely track changes in behavior related to reward expectancy.

Experiment 2

One limitation of Experiment 1 is that reward value (magnitude) was manipulated during Pavlovian conditioning. This complicates data interpretation because it conflates the influence that reward value has on conditioned responding with the potential influence of expected reward value on PIT performance. Consider the finding that rats decrease their rate of lever pressing when presented with a cue that predicts the delivery of 3 or 9 pellets. This may reflect a previously-established tendency for such cues to elicit a competing food-port response. Moreover, this response tendency may be performed in an automatic and inflexible manner. Thus, the results of Experiment 1 do not provide definitive support for the hypothesis that reward-predictive cues prompt a flexible shift from lever pressing to checking the food port based on expected reward value, though they are generally compatible with this account.

The post-training reward devaluation procedure avoids this issue and provides a more direct test of the degree to which a response is performed flexibly, based on an internal representation of the expected reward (Colwill & Rescorla, 1990; Holland, 2004; Holland & Straub, 1979; Rescorla, 1994). We applied this approach to probe the role of expected reward value using the same high-conflict PIT task used in Experiment 1. In

Experiment 2, rats were once again trained to press a lever for a grain-based food pellet, after which they were given differential Pavlovian conditioning with two distinct CS-reward contingencies. One CS signaled the delivery of 3 banana pellets and the other CS signaled the delivery of 3 chocolate pellets. Based on the results of Group 3/3 in Experiment 1, we anticipated that this training would establish cues that suppress lever pressing and elicit high levels of food-port activity during the PIT test. However, in the current experiment, we established a conditioned taste aversion to one of the two rewards used during Pavlovian conditioning prior to testing. Thus, while the two CSs signaled equally valuable rewards during training, only one of these CSs signaled a valuable reward at test, with the other signaling the recently devalued reward. According to the cognitive control account, this treatment should selectively disrupt the behavioral effects of the CS associated with the devalued reward since this cue no longer signals a need to suppress lever pressing and check the food port. However, if the behavioral effects of these CSs are established during training and expressed in a habitual or automatic manner at test, then reward devaluation should have no impact on task performance.

Methods

Animals and Apparatus

Sixteen experimentally naïve adult male Long Evans rats (Charles River) were used in Experiment 2. As in Experiment 1, they arrived at the facility at approximately 10 weeks of age. Husbandry, feeding, and experimental conditions were identical to that of

Experiment 1. The experiment was conducted in 16 operant chambers identical to the ones used in Experiment 1.

Procedure

Magazine training. Rats received two magazine-training sessions as described in Experiment 1.

Instrumental training. Instrumental training was conducted as in Experiment 1 with the exception that all rats received two daily sessions of FR-1 training, which were followed by 1 day of RI-15, 2 days of RI-30, and 10 days of RI-45 training.

Pavlovian training. Pavlovian training was conducted according to the same conditions that were used for Experiment 1's Group 3/3, except that each cue predicted a different type of food pellet. The food pellets used during Pavlovian training also differed from the one used during instrumental training. For half of the rats (n = 8) a 3-kHz tone predicted the delivery of three chocolate-flavored, sucrose-based pellets and a 10-Hz clicker predicted the delivery of three banana-flavored, sucrose-based pellets (45 mg; Dustless Precision Pellets, BioServ). The other rats (n = 8) received the alternative arrangement (i.e., the tone predicted banana pellets and the click predicted chocolate pellets). As for Experiment 1, Pavlovian training sessions involved 4 pseudorandomly-alternating presentations of each CS (8 total trials per session) separated by 120-s ISI. Pavlovian training took place over 25 sessions.

Specific reward devaluation by conditioned taste aversion (CTA). Following Pavlovian training one of the two food pellets used during Pavlovian conditioning was devalued through conditioned taste aversion learning. One type of food pellet was paired with nausea induced by an injection of lithium chloride (LiCl), whereas the other food pellet was paired with a saline injection (counterbalanced across conditioning contingencies). Rats first received 20 g of one type of food pellet in a metal cup in a neutral housing plexiglass cage for 60 min, after which they were given an injection of saline (20 mL/kg) before being returned to the home cage. The following day they received 20 g of the other type of food pellet for 60 min, followed by an injection of 20 mL/kg LiCl (0.15 M) (Balleine & Dickinson, 1992; Bouton, Broomer, Rey, & Thrailkill, 2020) before being placed back in their home cage. Rats received a total of three trials (days) with each food-injection arrangement in alternation over six days.

Pavlovian-to-instrumental transfer (PIT). Following CTA, rats received two sessions of instrumental retraining (identical to earlier sessions with an RI-45 schedule) followed by one 30-min session of instrumental extinction as in Experiment 1. Rats then received a PIT test session as in Experiment 1, during which the lever was once again continuously available but inactive. During the test, rats received 4 noncontingent presentations of each 30-s CS in pseudorandom order (ABBABAAB). The ISI was 150 s, and a 5-min interval preceded onset of the first CS. No food pellets were delivered at test.

Data Analysis

As in Experiment 1, summary measures were obtained from the raw data using MATLAB (The MathWorks; Natick, MA, USA) and analyzed with mixed-effects regression models (Pinheiro & Bates, 2000). Even though devaluation had yet to occur, Pavlovian training analyses included fixed effects of group (i.e., banana- vs. chocolate-flavored pellet devaluation) and CS type (i.e., devalued vs. nondevalued) and a by-rat random intercept. Unlike Experiment 1, due to programming error, instrumental-training analysis collapsed across all time bins, and a Wilcoxon rank sum test was used to compare average lever-pressing rates between the rats who were to have banana pellets devalued and those who were to have chocolate pellets devalued. For PIT analyses, because there were no group differences during Pavlovian and instrumental training (as reported below), the fixed-effects structure included CS type and 10-s CS time bin (CS Time; continuous), along with the corresponding two-way interaction (CS Type × CS Time); the random-effects structure included a by-rat random intercept. Similarly, for the analysis of conditioned taste aversion (CTA), which preceded PIT, the mixed-effects model's fixed-effects structure included the main effects of and interaction between day (1, 2, 3; continuous) and CTA condition (saline, LiCl; categorical), with by-rat random intercepts and slopes as a function of day (restricted to be uncorrelated). For CTA analysis, because the criterion was grams consumed, we employed a Poisson distribution with a log link function (i.e., data were converted to decagrams for analysis). Categorical predictors were effects-coded (i.e., codes sum to 0), and continuous predictors were mean-centered.

Our primary dependent measures were the same as in Experiment 1. Outlier detection and removal were also conducted as in Experiment 1. Data from one rat was removed from instrumental training analyses and data from one rat was removed from Pavlovian training analyses. For PIT analyses, one rat's lever-press data for trials with the nondevalued CS was removed. As in Experiment 1, the alpha level for all tests was .05, and main effects and interactions are reported as the results of ANOVA *F*-tests.

Results

Instrumental and Pavlovian Training

As in Experiment 1, rats were first trained to lever press for food, with the final days of training involving an RI-45-s schedule of reinforcement. While the rats had yet to experience any reward devaluation, we compared the extent to which lever-press rates differed by the eventual groups. Across the final three sessions of instrumental training, mean (*SEM*) lever-press response rates for the rats in which the banana- and chocolate-flavored pellets were later devalued were 15.3 (1.5) and 14.8 (1.4), respectively, which were not significantly different, $p = .397$ (Wilcoxon rank sum test). Likewise, there were no group differences in cue-elicited food-port entry behavior during the final five sessions of Pavlovian training, $F(1, 27) = 0.01$, $p = .942$ [banana devalued: 10.3 (1.3); chocolate devalued: 10.0 (2.1)], and there were no differences in cue-elicited food-port entry behavior between the subsequently devalued versus nondevalued cues, $F(1, 27) = 1.88$, $p = .182$ [devalued: 9.5 (1.9); nondevalued: 10.8 (1.7)].

Conditioned taste aversion (CTA) training

After initial training, one of the two types of food pellet used during Pavlovian conditioning was devalued through CTA training. As expected, rats selectively avoided consuming the food pellet that was paired with LiCl versus the saline-paired food, $F(1, 92) = 652.85$, $p < .001$ (Figure 3). Analysis also revealed a main effect of day, $F(1, 92) = 126.25$, $p < .001$, and a significant Day × CTA Condition interaction, $F(1, 92) = 760.88$, $p < .001$.

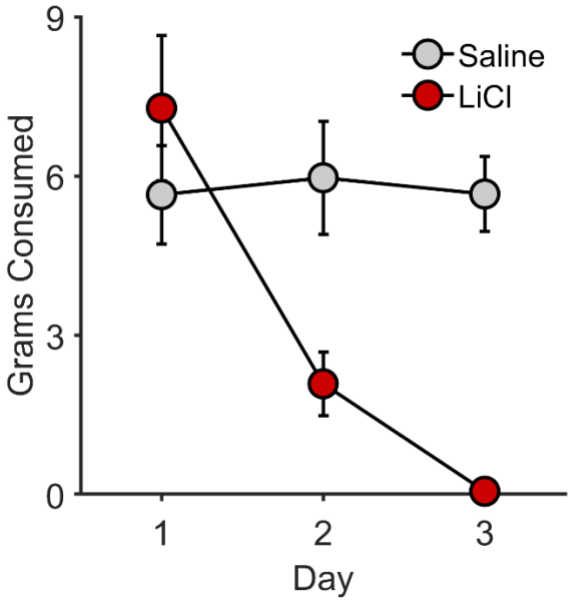

Figure 3. **Conditioned taste aversion in Experiment 2**. Grams consumed of the food reward paired with saline or lithium chloride (LiCl) as a function of day (trial). Error bars reflect ± 1 between-subjects standard error of the mean.

Pavlovian-instrumental transfer (PIT)

After CTA training, rats received 2 sessions of instrumental retraining and 1 session of instrumental extinction before undergoing a PIT test. Figure 3 shows cue-elicited

changes in lever pressing and food-port entry behavior relative to the 30-s pre-CS periods, plotted separately by CS type based on whether or not that cue signaled a devalued reward. For lever-pressing (Figure 4A), there was a main effect of CS type, $F(1, 89) = 6.23$, $p = .014$, with the cue that signaled the nondevalued reward suppressing the rate of lever pressing more than the cue that signaled the devalued reward. There was no effect of CS time, $F(1, 89) = 2.61$, $p = .110$, nor was there a CS Time × CS Type interaction, $F(1, 89) = 0.06$, $p = .809$.

In addition to being less effective in suppressing instrumental behavior, the cue that signaled the devalued reward was less effective in increasing food-port entries versus the cue that signaled the nondevalued reward, $F(1, 92) = 4.41$, $p = .038$ (Figure 4B). There was a main effect of CS time on cue-induced food-port entries, $F(1, 92) = 14.77$, $p < .001$, but not a CS time × CS Type interaction, $F(1, 92) = 0.02$, $p = .902$.

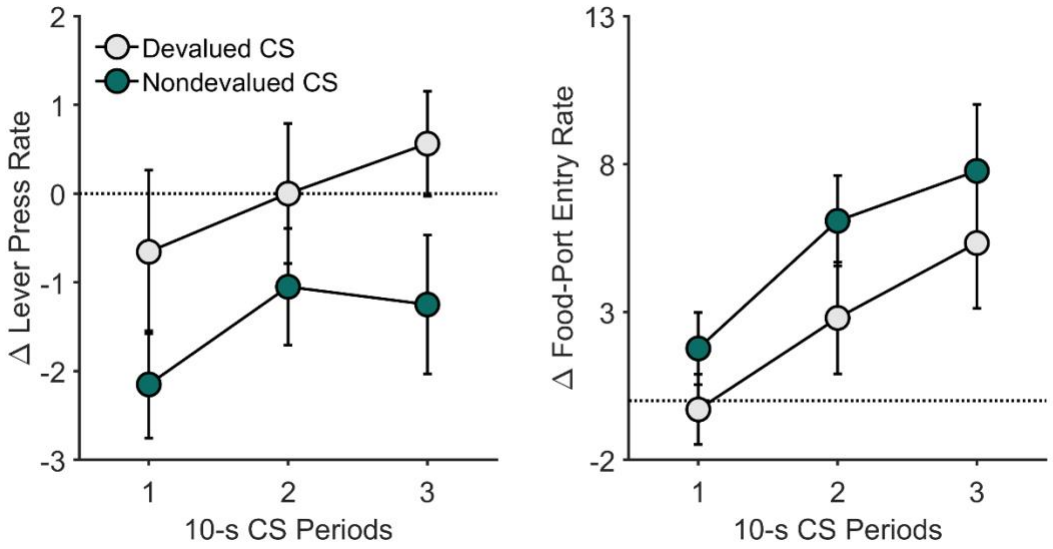

Figure 4. **Pavlovian-instrumental transfer in Experiment 2**. CS-induced changes in (A) lever-press rates and (B) food-port entries relative to the pre-CS baseline on the Pavlovian-

instrumental transfer (PIT) test. CS = conditioned stimulus. Error bars reflect ± 1 between-subjects standard error of the mean.

Discussion

These results show that Pavlovian reward-predictive cues suppress instrumental behavior and increase food-port activity in a flexible manner that depends on the current value of the expected reward. It is worth noting that the devalued CS had little influence on instrumental behavior in either direction. We had hypothesized that even cues that normally suppress behavior may acquire latent motivational properties which can be expressed under certain circumstances. For instance, a cue that signals a devalued reward might become an effective motivator of instrumental behavior if the devaluation treatment removes the inhibitory influence of reward expectancy without disrupting that cue's motivational properties. As noted above, previous studies have shown that reward devaluation typically does not disrupt the motivational influence of reward-paired cues (Colwill & Rescorla, 1990; Holland, 2004; Rescorla, 1994). However, such studies have explored a limited range of parameters, typically involving cues that are loosely correlated with reward delivery. Thus, it is possible that cues that signal precise information about the time of reward delivery, such as the CSs used here, acquire reward-specific motivational properties that are gated by expected reward value. Interestingly, Colwill and Rescorla (1990) found that while cues generally retain their ability to motivate instrumental behavior following reward devaluation, this was not the case for cues that signal forced intra-oral reward delivery. While we did not explore this factor in the current study, this finding shows that Pavlovian incentive motivation is indeed modulated by reward value under certain circumstances.

It is also possible that highly predictive cues, such as our CSs, do not acquire significant motivational properties, or at least do not acquire properties that allow them to motivate instrumental behavior. Yet another possibility is that the devaluation procedure used here was incomplete or for some other reason failed to completely disrupt the suppressive influence of reward expectancy on instrumental behavior. In this case, the motivational potential of the devalued CS would remain unexpressed. While future studies will be needed to determine which of these accounts is correct, the current findings provide strong evidence that the suppressive influence of reward-predictive cues is flexibly modulated by expected reward value.

General Discussion

The current study shows that reward expectancy plays an important role in regulating the expression of PIT performance. In Experiment 1, we found that cues signaling a relatively small reward (1 pellet) elicit an immediate and sustained increase in lever pressing while only modestly increasing food-port activity. In contrast, cues signaling either intermediate (3 pellets) or large (9 pellets) magnitude rewards suppressed lever pressing, particularly at the end of the cue period, near the expected time of reward delivery. These decrements in lever pressing were replaced with high levels of activity at the food port.

Experiment 2 provided more definitive evidence that reward-predictive cues suppress instrumental behavior in a flexible, goal-directed manner. Specifically, we found that the

tendency for rats to withhold lever pressing when presented with such a cue was attenuated if that cue signaled a devalued reward. A corresponding reduction in conditioned food-port activity was also observed on trials when a devalued reward was predicted. These findings show that reward-predictive cues do not exert reflexive control over behavior but instead prompt a shift from the lever to the food port that is contingent on an assessment of expected reward value.

While expected reward value appears to govern the suppressive influence of reward-predictive cues on instrumental behavior, this is usually not the case for the excitatory motivational influence of cues. As noted earlier, cues that motivate instrumental behavior tend to be relatively poor predictors of reward delivery. Previous studies have found that the motivational influence of such cues is surprisingly insensitive to post-training reward devaluation (Colwill & Rescorla, 1990; Holland, 2004; Rescorla, 1994). This has been shown for both the nonspecific, response-invigorating effect of such cues, referred to as general PIT (Holland, 2004), as well as the specific form of PIT, in which cues selectively increase the performance of instrumental actions with which they share a common outcome (Colwill & Rescorla, 1990; Holland, 2004; Rescorla, 1994). Given such findings, it appears that the excitatory PIT effect represents an implicit motivational response which compels instrumental behavior in an automatic or habitual manner, as opposed to prompting a more deliberative process of goal-directed decision making. This conclusion is bolstered by reports that instrumental training protocols that promote habit formation also render performance more sensitive to the motivational influence of reward-paired cues (Holland, 2004; Wiltgen et al., 2012).

The current findings show that this implicit motivational response to reward-paired cues is susceptible to top-down cognitive control, which is flexibly deployed based on the current value of an expected reward. Reward timing is another important factor. For instance, lever pressing is suppressed by cues signaling imminent reward and stimulated by cues signaling a long delay in reward delivery (Crombag et al., 2008; Lovibond, 1981; Marshall & Ostlund, 2018). Moreover, we have shown that cues signaling a relatively short, predictable delay in reward delivery provoke a gradual decrease in instrumental behavior that is most pronounced at the expected reward delivery time (Marshall & Ostlund, 2018), a trend that was also apparent in Experiment 1. PIT expression is also negatively regulated by expected reward probability. For instance, we have shown that cues signaling that reward delivery is unlikely tend to increase instrumental behavior, whereas cues signaling a high probability of reward tend to decrease instrumental behavior (Marshall et al., 2020). In addition to signaling information about reward likelihood, timing, magnitude, and incentive value, cues also shape instrumental behavior by signaling qualitative details about the flavor and texture of expected rewards, as is evident in the specific PIT effect (Balleine, 2016; Balleine & Ostlund, 2007; Delamater, 2011). This ability for cues to bias action selection to promote the pursuit of expected rewards is dissociable from their more general motivational effects (Corbit & Balleine, 2016; Ostlund & Maidment, 2012).

Our conceptual framework is generally compatible with the behavior systems theory (Timberlake et al., 1982), which assumes that cues shape conditioned behavior by

relaying information about the timing and location of food. According to this account, cues that signal uncertain or delayed reward trigger a general search mode (e.g., sign-tracking), whereas cues that signal imminent reward trigger focal search (e.g., goal-tracking). This theory readily explains the effects of expected reward timing and probability on PIT performance if one makes the straightforward assumptions that instrumental behavior is a general search activity and food-port entry is a focal search activity. However, the behavior systems account does not specify how rewards are represented during learning and therefore does not make useful predictions about whether treatments such as reward devaluation should differentially affect general versus focal search behavior.

Our framework is more directly aligned with Konorski's theory (1967, p. 276-280) that Pavlovian conditioning involves independent motivational and cognitive learning processes. According to this model, reward-paired cues rapidly become associated with the emotional properties of reward, allowing them to elicit preparatory conditioned responses and instrumental behavior. This motivational capacity extends to all reward-paired cues, regardless of their temporal relationship with reward. However, short cues that signal imminent reward delivery are unique in their tendency to form a separate association with a sensorily detailed representation of reward. Activating this representation results in what is essentially a cognitive reward expectancy, which is assumed to both evoke reward-specific consummatory behaviors and inhibit concurrent activity in the emotional-motivational pathway, thereby suppressing the impulse to act. This model readily accounts for the opposing effects of reward imminence and reward

probability on lever pressing and food-port approach behavior during PIT, but also accounts for the opposing effects of reward magnitude and incentive value on these behaviors, at least if one assumes that activation of the sensory reward representation is greater when the predicted reward is large and desirable than when it is small or undesirable.

We suggest that this inhibitory influence of reward-predictive cues represents a form of top-down cognitive control that allows for flexible and efficient foraging behavior. Seeking out new rewards through exploratory instrumental behavior is useful when rewards are scarce, unpredictable, or difficult to obtain. However, when reward delivery is imminent, such behavior becomes wasteful and may interfere with reward retrieval and consumption. We suggest that rapidly shifting from reward seeking to reward retrieval is particularly advantageous when a highly desirable reward is expected, because there is more to lose in such situations. Conversely, there is less need to avoid reward seeking when a small or devalued reward is expected since the potential loss is relatively low. In this case, it may be more appropriate to delay searching for the expected reward, perhaps until ongoing instrumental behavior has been completed.

While the current paper focuses on how Pavlovian learning influences instrumental behavior, similar processes appear to shape the expression of conditioned approach behavior. Rats' tendency to sign-track (i.e., approaching the cue) is commonly assumed to reflect a nonspecific motivational response (or attribution of incentive salience), whereas goal-tracking (i.e., checking the food port) is assumed to be mediated by a

cognitive expectancy of reward (Flagel, Akil, & Robinson, 2009; Robinson, Carr, & Kawa, 2018; Sarter & Phillips, 2018). This distinction is supported in part by findings that goal-tracking tends to be more sensitive to reward devaluation than sign-tracking behavior (Amaya, Stott, & Smith, 2020; but see Derman, Schneider, Juarez, & Delamater, 2018; Keefer, Bacharach, Kochli, Chabot, & Calu, 2020; Morrison, Bamkole, & Nicola, 2015; Patitucci, Nelson, Dwyer, & Honey, 2016). There is now a large body of studies that have adopted this framework to elucidate the neural mechanisms of motivation and cognitive control and explore how individual variability in conditioned approach behavior relate to addiction vulnerability (Anselme & Robinson, 2020; Flagel et al., 2009; Robinson et al., 2018; Sarter & Phillips, 2018). We suggest that the PIT paradigm provides a complementary approach that more directly models cognitive control over voluntary instrumental reward-seeking behavior.

The current findings are less compatible with existing theories of incentive learning (Dayan & Balleine, 2002; McClure et al., 2003; Zhang et al., 2009) that assume that the motivational properties of reward-predictive cues are directly linked to their state value – i.e., the total delay-discount reward expected based on a cue, as determined by temporal difference learning. These theories make the straightforward and intuitive prediction that cues signaling that an upcoming reward will be large and desirable should be more motivating than cues that predict small or undesirable rewards. However, the current study found evidence for the opposite relationship; that the motivational effects of a cue vary inversely with the value of the reward that it predicts. This theoretical approach is also difficult to reconcile with findings that the motivational

effects of a cue are negatively regulated by reward probability (Marshall et al., 2020) and reward imminence or timing (Marshall & Ostlund, 2018).

However, our interpretation of the current data has some caveats. For instance, we have assumed that the motivational properties of reward-predictive cues are primarily reflected in their ability to stimulate performance of a separately trained instrumental response. While this is indeed the central assumption of the PIT paradigm (Corbit & Balleine, 2016; Holmes, Marchand, & Coutureau, 2010; Rescorla & Solomon, 1967), it is possible that a cue's motivational properties also stimulate the expression of other reward-related behaviors, such as attempts to retrieve reward from the food port. This would have important implications given that this behavior tends to correlate positively with the expected probability, timing, magnitude, and value of reward (Marshall et al., 2020; Marshall & Ostlund, 2018). However, this inclusive view of cue-motivated behavior is not consistent with the available data. First, the PIT literature suggests that these behavioral responses are neurally dissociable. For instance, manipulations of the mesolimbic dopamine and cholinergic systems alter expression of cue-motivated lever pressing without affecting concomitant food-port activity (Collins, Aitken, Greenfield, Ostlund, & Wassum, 2016; Collins et al., 2019; Halbout et al., 2019). Second, as discussed earlier, cue-elicited food-port entry behavior tends to be suppressed by reward devaluation (Holland, 1998; Holland & Rescorla, 1975; Keefer et al., 2020; Morrison et al., 2015), indicating that its expression depends on expected reward value, in contrast to the motivational influence of reward-paired cues on instrumental behavior (Holland, 2004).

Another caveat is that cue-elicited food-port entry behavior may interfere with the expression of cue-motivated instrumental behavior in a way that complicates data interpretation. Put another way, PIT does not provide a pure readout of the potential for a cue to motivate instrumental behavior, but instead reflects the net motivational effect that is observed. Thus, cues predicting large rewards may indeed acquire a greater potential to motivate behavior than cues for small rewards, but this relationship may be obscured in PIT studies due to competing food-port activity. It is therefore premature to make any firm conclusions about theories that relate Pavlovian incentive learning to reward prediction (Dayan & Balleine, 2002; McClure et al., 2003; Zhang et al., 2009). A more critical test of this hypothesis may be possible using a low-conflict PIT task, such as by employing rewards that do not require a competing consummatory response (Baxter & Zamble, 1982; LeBlanc et al., 2012; Lovibond, 1983), or by extinguishing or otherwise discouraging the execution of such behavior (Holmes et al., 2010).

That said, we propose that the high-conflict PIT task used here and in our previous publications (Marshall et al., 2020; Marshall & Ostlund, 2018) is particularly useful for assaying cognitive control over cue-motivated behavior and how it can go awry to produce maladaptive behavior. For instance, we recently reported that adolescent male rats are impaired in using expected reward probability to shift from instrumental behavior to food-port entry during PIT (Marshall et al., 2020). Whereas these rats increased their instrumental behavior when presented with a cue that signaled a high probability of reward, a control group of adult male rats showed the opposite effect,

suppressing their instrumental behavior in order to check the food port. Interestingly, cues that signaled a low probability of reward elicited a similar increase in instrumental behavior in both adolescent and adult groups, which indicates that the adolescent rats were not simply more motivated by reward-paired cues. Using a similar approach, we have shown that rats with a history of repeated cocaine exposure are impaired in regulating their cue-motivated behavior, in this case based on the expected time of reward delivery (Marshall & Ostlund, 2018).

The current study suggests that such deficits reflect a failure of cognitive control over cue-motivated behavior that is normally expressed in a flexible and goal-directed manner. Further work is needed to validate this account. For instance, more should be done to determine if cues normally promote instrumental behavior through an implicit motivational process and whether this depends on Pavlovian conditioning parameters such as CS-reward interval. More should also be done to characterize the tendency for reward-predictive cues to interrupt instrumental behavior. Is this effect a simply byproduct of the competing conditioned response, or does it reflect an upstream process, such as the inhibition of a motivational impulse or the application of a higher-order action selection scheme? It is also important to investigate how the suppressive influence of reward expectancy relates to other cognitive features of PIT such as the opposing effects on action selection that are produced by cues that predict either the presence (Kruse, Overmier, Konz, & Rokke, 1983; Ostlund & Maidment, 2012) or absence (Delamater, Sosa, & LoLordo, 2003; Laurent & Balleine, 2015; Laurent, Wong, & Balleine, 2017) of specific rewards.

References

Amaya, K. A., Stott, J. J., & Smith, K. S. (2020). Sign-tracking behavior is sensitive to outcome devaluation in a devaluation context-dependent manner: implications for analyzing habitual behavior. *Learning & Memory, 27*(4), 136-149.

Anselme, P., & Robinson, M. J. F. (2020). From sign-tracking to attentional bias: Implications for gambling and substance use disorders. *Progress in Neuro-Psychopharmacology and Biological Psychiatry, 99*, 109861.

Azrin, N. H., & Hake, D. F. (1969). Positive Conditioned Suppresion - Conditioned Suppression Using Positive Reinforcers as Unconditioned Stimuli. *Journal of the Experimental Analysis of Behavior, 12*(1), 167-173. doi:DOI 10.1901/jeab.1969.12-167

Balleine, B. W. (2016). The cognitive control of goal-directed action: how predictive learning affects choice. In *Advances in Cognitive Neurodynamics (V)* (pp. 27-33): Springer.

Balleine, B. W., & Dickinson, A. (1992). Signalling and incentive processes in instrumental reinforcer devaluation. *The Quarterly Journal of Experimental Psychology Section B, 45*(4), 285-301. doi:10.1080/14640749208401007

Balleine, B. W., & Ostlund, S. B. (2007). Still at the choice-point - Action selection and initiation in instrumental conditioning. In B. W. Balleine, K. Doya, J. Odoherty, & M. Sakagami (Eds.), *Reward and Decision Making in Corticobasal Ganglia Networks* (Vol. 1104, pp. 147-171).

Baxter, D. J., & Zamble, E. (1982). Reinforcer and response specificity in appetitive transfer of control. *Animal Learning & Behavior, 10*(2), 201-210. doi:10.3758/BF03212271

Boisgontier, M. P., & Cheval, B. (2016). The anova to mixed model transition. *Neurosci Biobehav Rev, 68*, 1004-1005. doi:10.1016/j.neubiorev.2016.05.034

Bolker, B. M., Brooks, M. E., Clark, C. J., Geange, S. W., Poulsen, J. R., Stevens, M. H. H., & White, J.-S. S. (2009). Generalized linear mixed models: a practical guide for ecology and evolution. *Trends in Ecology & Evolution, 24*(3), 127-135. doi:10.1016/j.tree.2008.10.008

Botvinick, M., & Braver, T. (2015). Motivation and Cognitive Control: From Behavior to Neural Mechanism. *Annual Review of Psychology, 66*(1), 83-113. doi:10.1146/annurev-psych-010814-015044

Bouton, M. E., Broomer, M. C., Rey, C. N., & Thrailkill, E. A. (2020). Unexpected food outcomes can return a habit to goal-directed action. *Neurobiology of Learning and Memory, 169*, 107163. doi:https://doi.org/10.1016/j.nlm.2020.107163

Braver, T. S. (2012). The variable nature of cognitive control: a dual mechanisms framework. *Trends in Cognitive Sciences, 16*(2), 106-113. doi:https://doi.org/10.1016/j.tics.2011.12.010

Cartoni, E., Balleine, B. W., & Baldassarre, G. (2016). Appetitive Pavlovian-instrumental Transfer: A review. *Neurosci Biobehav Rev, 71*, 829-848. doi:10.1016/j.neubiorev.2016.09.020

Collins, A. L., Aitken, T. J., Greenfield, V. Y., Ostlund, S. B., & Wassum, K. M. (2016). Nucleus Accumbens Acetylcholine Receptors Modulate Dopamine and Motivation. *Neuropsychopharmacology, 41*(12), 2830-2838. doi:10.1038/npp.2016.81

Collins, A. L., Aitken, T. J., Huang, I. W., Shieh, C., Greenfield, V. Y., Monbouquette, H. G., . . . Wassum, K. M. (2019). Nucleus Accumbens Cholinergic Interneurons Oppose Cue-Motivated Behavior. *Biol Psychiatry, 86*(5), 388-396. doi:10.1016/j.biopsych.2019.02.014

Colwill, R. M., & Rescorla, R. A. (1990). Effect of reinforcer devaluation on discriminative control of instrumental behavior. *Journal of Experimental Psychology: Animal Behavior Processes, 16*(1), 40-47. doi:http://dx.doi.org/10.1037/0097-7403.16.1.40

Corbit, L. H., & Balleine, B. W. (2016). Learning and Motivational Processes Contributing to Pavlovian-Instrumental Transfer and Their Neural Bases: Dopamine and Beyond. *Curr Top Behav Neurosci, 27*, 259-289. doi:10.1007/7854_2015_388

Corbit, L. H., & Janak, P. H. (2007). Ethanol-associated cues produce general pavlovian-instrumental transfer. *Alcohol Clin Exp Res, 31*(5), 766-774. doi:10.1111/j.1530-0277.2007.00359.x

Corbit, L. H., & Janak, P. H. (2016). Changes in the Influence of Alcohol-Paired Stimuli on Alcohol Seeking across Extended Training. *Front Psychiatry, 7*, 169. doi:10.3389/fpsyt.2016.00169

Crombag, H. S., Galarce, E. M., & Holland, P. C. (2008). Pavlovian influences on goal-directed behavior in mice: The role of cue-reinforcer relations. *Learning & Memory, 15*(5), 299-303. doi:10.1101/lm.762508

Dayan, P., & Balleine, B. W. (2002). Reward, motivation, and reinforcement learning. *Neuron, 36*(2), 285-298. doi:10.1016/s0896-6273(02)00963-7

Delamater, A. R. (2011). At the interface of learning and cognition: An associative learning perspective. *International Journal of Comparative Psychology, 24*(4).

Delamater, A. R., & Holland, P. C. (2008). The influence of CS-US interval on several different indices of learning in appetitive conditioning. *Journal of experimental psychology. Animal behavior processes, 34*(2), 202-222. doi:10.1037/0097-7403.34.2.202

Delamater, A. R., & Oakeshott, S. (2007). Learning about multiple attributes of reward in Pavlovian conditioning. *Annals of the New York Academy of Sciences, 1104*(1), 1-20.

Delamater, A. R., Sosa, W., & LoLordo, V. M. (2003). Outcome-specific conditioned inhibition in Pavlovian backward conditioning. *Animal Learning & Behavior, 31*(4), 393-402.

Derman, R. C., Schneider, K., Juarez, S., & Delamater, A. R. (2018). Sign-tracking is an expectancy-mediated behavior that relies on prediction error mechanisms. *Learning & Memory, 25*(10), 550-563.

Dickinson, A., Smith, J., & Mirenowicz, J. (2000). Dissociation of Pavlovian and instrumental incentive learning under dopamine antagonists. *Behavioral Neuroscience, 114*(3), 468-483. doi:http://dx.doi.org/10.1037/0735-7044.114.3.468

Ellison, G. D., & Konorski, J. (1965). An investigation of the relations between salivary and motor responses during instrumental performance. *Acta biologiae experimentalis, 25*(4), 297-315.

Estes, W. K. (1948). Discriminative conditioning; effects of a Pavlovian conditioned stimulus upon a subsequently established operant response. *J Exp Psychol, 38*(2), 173-177. doi:10.1037/h0057525

Flagel, S. B., Akil, H., & Robinson, T. E. (2009). Individual differences in the attribution of incentive salience to reward-related cues: Implications for addiction. *Neuropharmacology, 56*, 139-148.

Garbusow, M., Schad, D. J., Sebold, M., Friedel, E., Bernhardt, N., Koch, S. P., . . . Heinz, A. (2016). Pavlovian-to-instrumental transfer effects in the nucleus accumbens relate to relapse in alcohol dependence. *Addiction Biology, 21*(3), 719-731. doi:10.1111/adb.12243

Genauck, A., Andrejevic, M., Brehm, K., Matthis, C., Heinz, A., Weinreich, A., . . . Romanczuk-Seiferth, N. (2020). Cue-induced effects on decision-making distinguish subjects with gambling disorder from healthy controls. *Addict Biol, 25*(6), e12841. doi:10.1111/adb.12841

Halbout, B., Marshall, A. T., Azimi, A., Liljeholm, M., Mahler, S. V., Wassum, K. M., & Ostlund, S. B. (2019). Mesolimbic dopamine projections mediate cue-motivated reward seeking but not reward retrieval in rats. *Elife, 8*, e43551. doi:10.7554/eLife.43551

Hoffman, L., & Rovine, M. J. (2007). Multilevel models for the experimental psychologist: Foundations and illustrative examples. *Behavior Research Methods, 39*(1), 101-117. doi:10.3758/bf03192848

Holland, P. C. (1998). Amount of training affects associatively-activated event representation. *Neuropharmacology, 37*(4-5), 461-469. doi:10.1016/s0028-3908(98)00038-0

Holland, P. C. (2004). Relations between Pavlovian-instrumental transfer and reinforcer devaluation. *J Exp Psychol Anim Behav Process, 30*(2), 104-117. doi:10.1037/0097-7403.30.2.104

Holland, P. C., & Rescorla, R. A. (1975). The effect of two ways of devaluing the unconditioned stimulus after first- and second-order appetitive conditioning. *Journal of Experimental Psychology: Animal Behavior Processes, 1*(4), 355-363. doi:10.1037/0097-7403.1.4.355

Holland, P. C., & Straub, J. J. (1979). Differential effects of two ways of devaluing the unconditioned stimulus after Pavlovian appetitive conditioning. *Journal of Experimental Psychology: Animal Behavior Processes, 5*(1), 65.

Holmes, N. M., Marchand, A. R., & Coutureau, E. (2010). Pavlovian to instrumental transfer: a neurobehavioural perspective. *Neurosci Biobehav Rev, 34*(8), 1277-1295. doi:10.1016/j.neubiorev.2010.03.007

Keefer, S. E., Bacharach, S. Z., Kochli, D. E., Chabot, J. M., & Calu, D. J. (2020). Effects of limited and extended Pavlovian training on devaluation sensitivity of sign-and goal-tracking rats. *Frontiers in Behavioral Neuroscience, 14*, 3.

Konorski, J. (1967). *Integrative Activity of the Brain: An Interdisciplinary Approach*: University of Chicago Press.

Konorski, J., & Wyrwicka, W. (1950). Researches Into Conditioned Reflexes of the Second Type. 1. Transformation of Conditioned Reflexes of the First Type into Conditioned Reflexes of the Second Type. *Acta biologiae experimentalis, 15*, 193-208.

Kruse, J. M., Overmier, J. B., Konz, W. A., & Rokke, E. (1983). Pavlovian conditioned stimulus effects upon instrumental choice behavior are reinforcer specific. *Learning and motivation, 14*(2), 165-181.

Laurent, V., & Balleine, B. W. (2015). Factual and Counterfactual Action-Outcome Mappings Control Choice between Goal-Directed Actions in Rats. *Current Biology, 25*(8), 1074-1079. doi:https://doi.org/10.1016/j.cub.2015.02.044

Laurent, V., Wong, F. L., & Balleine, B. W. (2017). The lateral habenula and its input to the rostromedial tegmental nucleus mediates outcome-specific conditioned inhibition. *Journal of Neuroscience, 37*(45), 10932-10942.

LeBlanc, K. H., Maidment, N. T., & Ostlund, S. B. (2013). Repeated cocaine exposure facilitates the expression of incentive motivation and induces habitual control in rats. *Plos One, 8*(4), e61355. doi:10.1371/journal.pone.0061355

LeBlanc, K. H., Maidment, N. T., & Ostlund, S. B. (2014). Impact of repeated intravenous cocaine administration on incentive motivation depends on mode of drug delivery. *Addict Biol, 19*(6), 965-971. doi:10.1111/adb.12063

LeBlanc, K. H., Ostlund, S. B., & Maidment, N. T. (2012). Pavlovian-to-instrumental transfer in cocaine seeking rats. *Behavioral Neuroscience, 126*(5), 681-689. doi:10.1037/a0029534

Leys, C., Ley, C., Klein, O., Bernard, P., & Licata, L. (2013). Detecting outliers: Do not use standard deviation around the mean, use absolute deviation around the median. *Journal of Experimental Social Psychology, 49*(4), 764-766. doi:https://doi.org/10.1016/j.jesp.2013.03.013

Lovibond, P. F. (1981). Appetitive Pavlovian-Instrumental Interactions - Effects of Inter-Stimulus Interval and Baseline Reinforcement Conditions. *Quarterly Journal of Experimental Psychology Section B-Comparative and Physiological Psychology, 33*(Nov), 257-269. doi:Doi 10.1080/14640748108400811

Lovibond, P. F. (1983). Facilitation of instrumental behavior by a Pavlovian appetitive conditioned stimulus. *J Exp Psychol Anim Behav Process, 9*(3), 225-247. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/6153052

Marshall, A. T., Munson, C. N., Maidment, N. T., & Ostlund, S. B. (2020). Reward-predictive cues elicit excessive reward seeking in adolescent rats. *Dev Cogn Neurosci, 45*, 100838. doi:10.1016/j.dcn.2020.100838

Marshall, A. T., & Ostlund, S. B. (2018). Repeated cocaine exposure dysregulates cognitive control over cue-evoked reward-seeking behavior during Pavlovian-to-instrumental transfer. *Learn Mem, 25*(9), 399-409. doi:10.1101/lm.047621.118

Matell, M. S., & Della Valle, R. B. (2018). Temporal specificity in Pavlovian-to-instrumental transfer. *Learning & Memory, 25*(1), 8-20.

McClure, S. M., Daw, N. D., & Montague, P. R. (2003). A computational substrate for incentive salience. *Trends Neurosci, 26*(8), 423-428. doi:10.1016/s0166-2236(03)00177-2

Miller, E. K., & Cohen, J. D. (2001). An Integrative Theory of Prefrontal Cortex Function. *Annual Review of Neuroscience, 24*(1), 167-202. doi:10.1146/annurev.neuro.24.1.167

Morrison, S. E., Bamkole, M. A., & Nicola, S. M. (2015). Sign tracking, but not goal tracking, is resistant to outcome devaluation. *Frontiers in Neuroscience, 9*, 468.

Ostlund, S. B., LeBlanc, K. H., Kosheleff, A. R., Wassum, K. M., & Maidment, N. T. (2014). Phasic mesolimbic dopamine signaling encodes the facilitation of incentive motivation produced by repeated cocaine exposure. *Neuropsychopharmacology, 39*(10), 2441-2449. doi:10.1038/npp.2014.96

Ostlund, S. B., & Maidment, N. T. (2012). Dopamine Receptor Blockade Attenuates the General Incentive Motivational Effects of Noncontingently Delivered Rewards and Reward-Paired Cues Without Affecting Their Ability to Bias Action Selection. *Neuropsychopharmacology, 37*(2), 508-519. doi:10.1038/npp.2011.217

Ostlund, S. B., & Marshall, A. T. (2021). Probing the role of reward expectancy in Pavlovian-instrumental transfer. *Current Opinion in Behavioral Sciences, 41*, 106-113. doi:https://doi.org/10.1016/j.cobeha.2021.04.021

Patitucci, E., Nelson, A. J. D., Dwyer, D. M., & Honey, R. C. (2016). The origins of individual differences in how learning is expressed in rats: A general-process perspective. *Journal of Experimental Psychology: Animal Learning and Cognition, 42*(4), 313.

Pinheiro, J. C., & Bates, D. M. (2000). *Mixed-effects models in S and S-PLUS*. New York: Springer.

Rescorla, R. A. (1994). Transfer of instrumental control mediated by a devalued outcome. *Animal Learning & Behavior, 22*(1), 27-33. doi:10.3758/bf03199953

Rescorla, R. A., & Solomon, R. L. (1967). Two-process learning theory: Relationships between Pavlovian conditioning and instrumental learning. *Psychol Rev, 74*(3), 151-182. doi:10.1037/h0024475

Robinson, T. E., Carr, C., & Kawa, A. B. (2018). The propensity to attribute incentive salience to drug cues and poor cognitive control combine to render sign-trackers susceptible to addiction. *Sign-tracking and drug addiction (Vol. A), 10*.

Saddoris, M. P., Stamatakis, A., & Carelli, R. M. (2011). Neural correlates of Pavlovian-to-instrumental transfer in the nucleus accumbens shell are selectively potentiated following cocaine self-administration. *European Journal of Neuroscience, 33*(12), 2274-2287. doi:10.1111/j.1460-9568.2011.07683.x

Sarter, M., & Phillips, K. B. (2018). The neuroscience of cognitive-motivational styles: Sign-and goal-trackers as animal models. *Behavioral Neuroscience, 132*(1), 1.

Schielzeth, H., Nakagawa, S., & Freckleton, R. (2013). Nested by design: model fitting and interpretation in a mixed model era. *Methods in Ecology and Evolution, 4*(1), 14-24. doi:10.1111/j.2041-210x.2012.00251.x

Sebold, M., Garbusow, M., Cerci, D., Chen, K., Sommer, C., Huys, Q. J., . . . Friedel, E. (2021). Association of the OPRM1 A118G polymorphism and Pavlovian-to-instrumental transfer: Clinical relevance for alcohol dependence. *J Psychopharmacol*, 269881121991992. doi:10.1177/0269881121991992

Shiflett, M. W. (2012). The effects of amphetamine exposure on outcome-selective Pavlovian-instrumental transfer in rats. *Psychopharmacology (Berl), 223*(3), 361-370. doi:10.1007/s00213-012-2724-y

Shiflett, M. W., Riccie, M., & DiMatteo, R. (2013). The effects of amphetamine sensitization on conditioned inhibition during a Pavlovian-instrumental transfer task in rats. *Psychopharmacology (Berl), 230*(1), 137-147. doi:10.1007/s00213-013-3144-3

Soltysik, S., Konorski, J., Holownia, A., & Rentoul, T. (1976). The effect of conditioned stimuli signalling food upon the autochthonous instrumental responses in dogs. *Acta Neurobioliae Experimentalis, 36*, 277-310.

Tatham, T. A., & Zurn, K. R. (1989). The MED-PC experimental apparatus programming system. *Behavior Research Methods, Instruments, & Computers, 21*(2), 294-302. doi:10.3758/bf03205598

Timberlake, W., Wahl, G., & King, D. A. (1982). Stimulus and response contingencies in the misbehavior of rats. *Journal of Experimental Psychology: Animal Behavior Processes, 8*(1), 62.

Vandyne, G. C. (1971). Conditioned Suppression with a Positive Us in Rat. *Journal of Comparative and Physiological Psychology, 77*(1), 131-&. doi:DOI 10.1037/h0031576

Vogel, V., Kollei, I., Duka, T., Snagowski, J., Brand, M., Muller, A., & Loeber, S. (2018). Pavlovian-to-instrumental transfer: A new paradigm to assess pathological mechanisms with regard to the use of Internet applications. *Behav Brain Res, 347*, 8-16. doi:10.1016/j.bbr.2018.03.009

Wiltgen, B. J., Sinclair, C., Lane, C., Barrows, F., Molina, M., & Chabanon-Hicks, C. (2012). The effect of ratio and interval training on Pavlovian-instrumental transfer in mice. *Plos One, 7*(10), e48227.

Wyvell, C. L., & Berridge, K. C. (2001). Incentive sensitization by previous amphetamine exposure: increased cue-triggered "wanting" for sucrose reward. *J Neurosci, 21*(19), 7831-7840. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/11567074

Zhang, J., Berridge, K. C., Tindell, A. J., Smith, K. S., & Aldridge, J. W. (2009). A neural computational model of incentive salience. *PLoS Comput Biol, 5*(7), e1000437. doi:10.1371/journal.pcbi.1000437