Sign-tracking behavior is sensitive to outcome devaluation in a devaluation context-dependent manner: implications for analyzing habitual behavior

Kenneth A. Amaya,[1,2] Jeffrey J. Stott,[1,2] and Kyle S. Smith[1]

[1]Department of Psychological and Brain Sciences, Dartmouth College, Hanover, New Hampshire 03755, USA

Motivationally attractive cues can draw in behavior in a phenomenon termed incentive salience. Incentive cue attraction is an important model for animal models of drug seeking and relapse. One question of interest is the extent to which the pursuit of motivationally attractive cues is related to the value of the paired outcome or can become unrelated and habitual. We studied this question using a sign-tracking (ST) paradigm in rats, in which a lever stimulus preceding food reward comes to elicit conditioned lever-interaction behavior. We asked whether reinforcer devaluation by means of conditioned taste aversion, a classic test of habitual behavior, can modify ST to incentive cues, and whether this depends upon the manner in which reinforcer devaluation takes place. In contrast to several recent reports, we conclude that ST is indeed sensitive to reinforcer devaluation. However, this effect depends critically upon the congruence between the context in which taste aversion is learned and the context in which it is tested. When the taste aversion successfully transfers to the testing context, outcome value strongly influences ST behavior, both when the outcome is withheld (in extinction) and when animals can learn from outcome feedback (reacquisition). When taste aversion does not transfer to the testing context, ST remains high. In total, the extent to which ST persists after outcome devaluation is closely related to the extent to which that outcome is truly devalued in the task context. We believe this effect of context on devaluation can reconcile contradictory findings about the flexibility/inflexibility of ST. We discuss this literature and relate our findings to the study of habits generally.

[Supplemental material is available for this article.]

Environmental stimuli that predict rewards can become motivationally attractive, drawing in behavior in a psychological phenomenon termed incentive salience. This incentive motivation for cues occurs naturally in "sign-tracking" (ST), in which animals engage in appetitive behavior that is directed toward the cue itself (Brown and Jenkins 1968), instead of toward the site of reward delivery ("goal-tracking" (GT); (Boakes 1977). For example, when rats are presented with the extension of a physical lever as a conditioned stimulus (CS) that precedes food delivery, some animals sign-track to the CS lever by approaching and interacting with it rather than going to the source of impending food delivery.

The motivational attraction to reward-predictive cues, which occurs in ST, has drawn considerable interest by neuroscientists working to understand the brain basis of incentive salience. It is thought that the process is normally adaptive, such that animals become drawn to cues predictive of valuable resources, like food. There is also evidence that ST can become maladaptive and inflexible in certain circumstances, leading to compulsive-like behaviors. A striking example of maladaptive behavior occurs in models of addiction. Notably, the acquisition and maintenance of ST behavior, but not GT behavior, depends on dopamine receptor activation in the nucleus accumbens core (Flagel et al. 2011; Saunders and Robinson 2012), and accumbens dopamine signaling has long been linked to the reinforcing properties of drugs and as well as the maintenance of drug-seeking behaviors. For example, the propensity for rodents to exhibit ST behavior correlates with heightened locomotor sensitization to dopaminergic drugs of abuse

(Flagel et al. 2008), primed reinstatement to drug seeking as a model of relapse (Saunders and Robinson 2010, 2011), and the likelihood that an animal will endure punishment (crossing an electrified floor) to seek cocaine during a cued-reinstatement test (Saunders et al. 2013).

Beyond the link of ST to addiction, ST for an appetitive reward itself can develop signs of inflexibility and habit-like behavior. For instance, pigeons exposed to an illuminated key light followed by food delivery at the other end of a chamber will peck at the key light at the expense of being able to retrieve the food reward at the opposite end (Hearst and Jenkins 1974). ST can also persist in rodents when doing so results in food reward omission (Locurto et al. 1976; Chang and Smith 2016). Such behavioral and neural markers of ST behavior have given rise to the notion that GT is a behavior that relies more heavily on a goal-directed, "model-based" strategy, while ST relies more on a stimulus-driven, "model-free" and dopamine-dependent strategy (Lesaint et al. 2014, 2015), akin to a habit. This idea—that ST behavior is model-free, stimulus-driven, and habit-like—carries a clear experimental implication.

A key feature of habitual or model-free behavior is that it persists despite changes to the value of the outcome that it precedes (Dickinson and Balleine, 1994). *Thus, ST behavior should also occur independently of outcome value.* A series of recent studies have examined the persistence of ST/GT behavior using a method whereby a

[2]These authors contributed equally to this work.
Corresponding author: kenneth.a.amaya.gr@dartmouth.edu

previously rewarding outcome like food is made less rewarding (typically by pairing it with injections of a nauseagenic drug, for example, lithium chloride (LiCl)). Animals are then tested to determine whether they continue to perform the behavior related to this now-devalued outcome in extinction conditions. In some studies, ST behaviors persist and appear to be habitual. Several recent studies support the conclusion that ST behavior in rodents, and ST-like behavior in humans, is resistant to change and is generally unaltered by changes in outcome value (Morrison et al. 2015; Nasser et al. 2015; Patitucci et al. 2016; De Tommaso et al. 2017; Vandaele et al. 2017; Smedley and Smith 2018).

However, there are notable exceptions. Davey and Cleland (1982) and Robinson and Berridge (2013) have demonstrated robust and even immediate effects of reward revaluation on ST behavior, showing that ST can exhibit considerable flexibility more in line with a model-based behavioral or motivational system (Dayan and Berridge 2014). Of particular note, (Derman et al. 2018) recently found evidence directly in conflict with the above studies. They showed that ST behavior can in fact display sensitivity to outcome devaluation. Collectively, there appears to be evidence for ST to be both flexible and inflexible, and both outcome-sensitive and outcome-insensitive. Understanding which is true, or, more specifically, why opposing outcome devaluation effects are found, carries importance for interpreting ST. Since variations in outcome devaluation sensitivity also occur in more traditional instrumental tasks such as lever pressing or maze running, this question is potentially a broadly important one.

Focusing on the outcome-devaluation assay for behavioral flexibility/inflexibility, we have noted in our own work and in the literature that the devaluation procedure itself might play a major role in the test outcome. In other words, variation in how (more specifically, *where*) the reward-LiCl pairings are done can lead to an impression that ST is either inflexible and outcome-insensitive, or that ST is flexible and outcome-guided. Thus, we undertook a series of experiments aimed to clarify the effect of devaluation on ST behavior, and the role the reward-LiCl pairing environment has on subsequent devaluation sensitivity in ST. Attention was paid to how devaluation in or out of the task chamber transferred to in-task ST behavior as well as to the in-task value of the reward itself. In Experiment 1, we examined the effect of a novel outcome devaluation procedure on postdevaluation ST rates using a discriminative two-lever CS conditioning design. We found that ST was devaluation-sensitive with this protocol. In Experiment 2, we determined whether LiCl injections alone in the task context could account for the efficacy of this procedure: they could not. Experiment 3 extended the findings from Experiment 1 to nondiscriminative single-lever CS preparations. Finally, in Experiment 4, we directly compared the effects of devaluation when it was administered inside or outside of the task environment, which turns out to play a critical role. Altogether, we provide evidence that ST can show sensitivity to reinforcer devaluation and conclude that ST is mediated a great deal by its relationship with the value of the rewarded outcome. These results carry broad implications in the sense that the location of devaluation procedures can carry great consequence for determining whether behaviors are outcome-sensitive (model-based, goal-directed) or outcome-insensitive (model-free, habitual).

Results

Experiment 1

Experiment 1 directly assessed outcome-devaluation sensitivity of ST behavior. See Figure 1A for a schematic of the experimental procedures. Rats ($n = 16$) were given a magazine training session followed by 12 daily sessions of ST training where a lever cue (CS+)

was presented for 10 sec followed by the delivery of two grain pellets. We used two pellets as reward because on a rare occasion a pellet does not get delivered by the device, so using just one can lead to the odd trial where a pellet is not delivered. Further, we adopted the same procedure that we have used previously, in order to more directly compare our results to past studies (Chang et al. 2015, 2018; Smedley and Smith 2018; Smedley et al. 2019). A second lever was presented on separate trials but was not predictive of reward (CS−). 25 CS+ and 25 CS− trials were administered during each session with an average intertrial interval (ITI) length of 60 sec. Predevaluation probe sessions consisting of a brief extinction session (5 CS+, 5 CS− presentations) and a fully rewarded reacquisition session (25 CS+, 25 CS−) were conducted to establish baseline response levels before proceeding to outcome devaluation. After training, rats were behaviorally matched as determined by mean response rates and split into two groups, Group LiCl-Pellet ($n = 8$), which received pellet reinforcers and then LiCl injections during devaluation, and Group Saline-Pellet ($n = 8$), which received pellet reinforcers and then saline injections during devaluation. A hybrid devaluation approach was used where outcome devaluation was conducted in transport chambers for two sessions then conducted in operant chambers for subsequent sessions. Postdevaluation probe sessions, structured similarly to predevaluation probes, were then conducted. Magazine entry data throughout the experiment was examined and all figures and statistics can be found in Supplemental Figure 1 and Supplemental Table 1. If ST behavior is guided by outcome value, then rats given LiCl-reward pairings in this multienvironment way should show significantly decreased lever-pressing behavior when compared to saline controls.
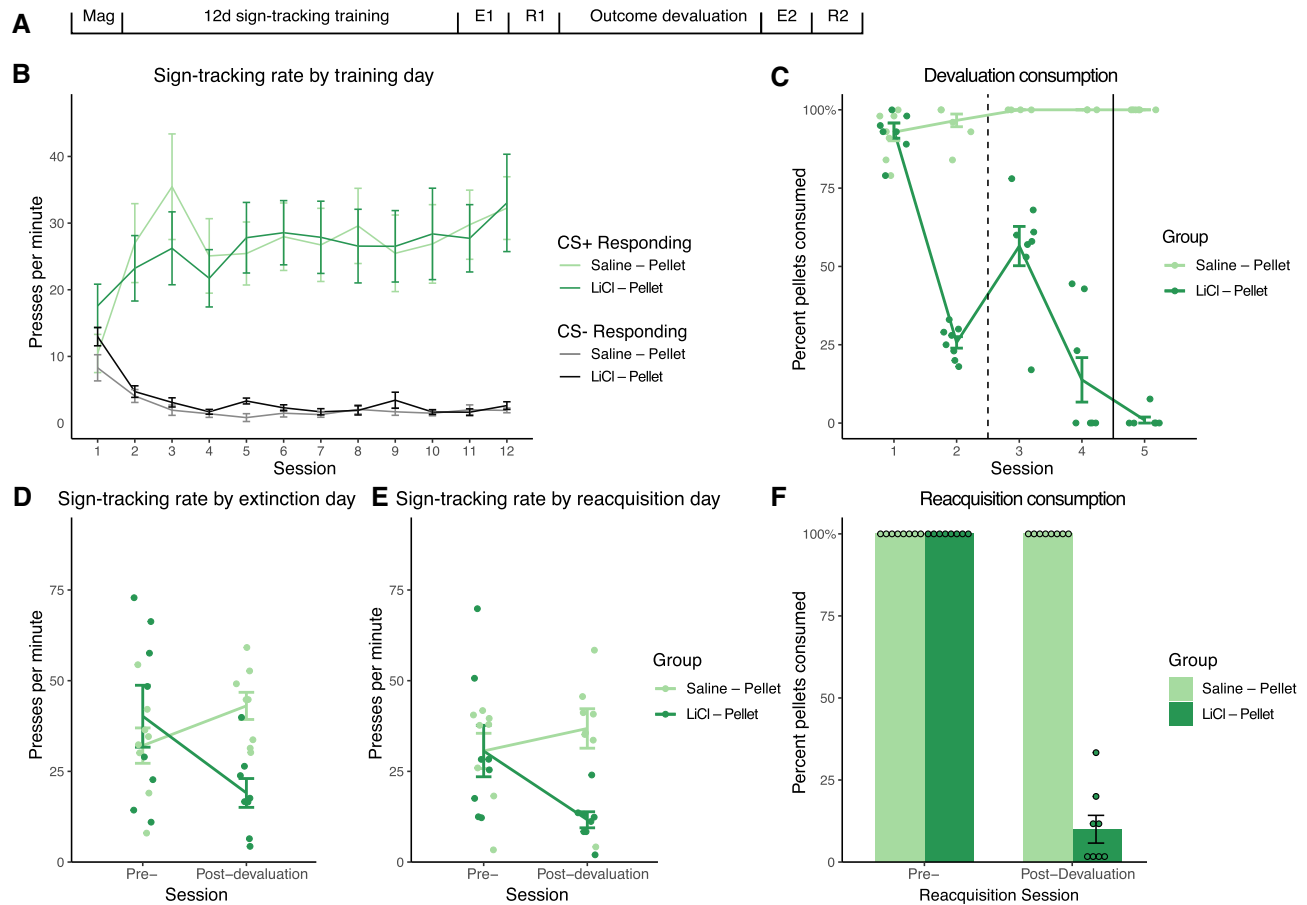
Acquisition

The mean presses per minute (ppm) over the course of training is presented in Figure 1B. To compare rates of responding, a linear mixed model using ppm as the dependent variable and fixed effects of CS type, group, and logarithmic session (sessions 1–12; logSession) as well as an interaction between CS type, group, and session with random intercepts for individual animals and learning curves were included. The logarithmic fit of session produced a model with a lower Akaike information criterion (AIC, 2824.5) than a model using linear session (2837.0).

There was no significant effect of Group (estimate: 0.79 ppm; confidence interval (CI): −2.63–4.22; $P = 0.66$), showing that Group Saline-Pellet and Group LiCl-Pellet did not differ on average in ppm. There was a no effect of logSession (estimate: 0.94 ppm; CI: −0.97–2.85; $P = 0.35$), showing no overall increase in lever presses over training. Importantly, there was a significant effect of CS-type (estimate: 5.75; CI: 3.57–7.92; $P < 0.001$) and a significant interaction between CS-type and logSession (estimate: 3.64; CI: 2.44–4.84; $P < 0.001$). Further, there was not a significant Group by logSession interaction (estimate: −0.40; CI: −2.32–1.51; $P = 0.68$) nor Group by CS-type interaction (estimate: −0.77; CI: 2.94–1.41; $P = 0.49$). There was not a significant CS-type by Group by logSession interaction (estimate: 0.22; CI: −0.98–1.41; $P = 0.724$). Magazine entries made by groups similarly decreased to very low levels over the course of the training sessions (Supplemental Fig. 1A).

Together, these results indicate that animals readily discriminated between the CS+ and CS−, as manifested in the number of lever presses made during respective presentations of these stimuli throughout training, and that there was no difference between the groups in how they interacted with the CS+ and CS− as a function of training sessions.

Outcome devaluation

Rats received two devaluation pairings in transport boxes, and the following three pairings in operant chambers. This was done to

Figure 1. ST behavior was sensitive to outcome devaluation. (*A*) Timeline of experimental procedures: magazine training (Mag), predevaluation extinction session (E1), predevaluation reacquisition session (R1), postdevaluation extinction session (E2), and postdevaluation extinction session (R2). (*B*) Sign-tracking behavior did not differ between groups on a discriminative sign-tracking protocol over the course of 12 training sessions. Groups readily discriminated between the CS+ and CS− and adjusted responding (presses per minute, ppm) accordingly. (*C*) Outcome devaluation conducted over the course of five sessions. Sessions 1 and 2, to the *left* of the dashed line, were conducted in external holding chambers distinct from operant chambers. Sessions 3 and 4 were conducted in operant chambers. Session 5, to the *right* of the solid line, was conducted in operant chambers but no lithium chloride was administered following pellet presentation as minimal pellets were consumed. (*D*) Sign-tracking rates (ppm) during 5-min extinction tests pre- and postdevaluation show that Group LiCl-Pellet significantly decreased sign-tracking to the lever following devaluation. (*E*) Sign-tracking rates (ppm) during fully rewarded sessions show the same drop following outcome devaluation in Group LiCl-Pellet. (*F*) Pellet consumption during the pre- and postdevaluation reacquisition sessions. Predevaluation, all animals consumed all pellets, as expected. Postdevaluation, animals in Group Saline-Pellet continued to consume all pellets while animals in Group LiCl-Pellet significantly decreased pellet consumption, indicating that outcome devaluation was successful.

encourage generalization of the learned aversion to the testing context (operant chambers). The mean percentage of pellets consumed during each day of outcome devaluation is presented in Figure 1C. A generalized linear mixed model was created to analyze fixed effects of Session, Group, and Group by Session interaction with random intercepts for individual rats.

Here, there were effects of Session (odds ratio (OR): 0.65, CI: 0.55–0.76, $P < 0.001$), Group (OR: 598.90, CI: 306.84–1168.97, $P < 0.001$), and a significant Group by Session interaction (OR: 4.31, CI: 3.67–5.07, $P < 0.001$). Thus, Group Saline-Pellet was far more likely to consume pellets than Group LiCl-Pellet on Day 5, the final day of outcome devaluation. Additionally, pellet consumption across all animals became less likely with each successive session and animals in different groups changed their pellet consumption differently over sessions, as indicated by the significant interaction between Group and Session. Interestingly, there was a relative increase in the proportion of pellets consumed from Day 2 to Day 3 of outcome devaluation in Group LiCl-Pellet, which was the first session of outcome devaluation conducted in the operant chambers. A Wilcoxon signed rank test revealed that animals

in Group LiCl-Pellet significantly increased their pellet consumption between these 2 d ($V = 1$, $P = 0.021$), suggestive of poor taste-aversion generalization from transport boxes back to the operant chambers.

Devaluation sensitivity in extinction

Responding during pre- and postdevaluation extinction sessions was compared. Predevaluation ST rates by group and session are presented in Figure 1D. A linear mixed model using response rates as the dependent variable and fixed effects of Group, Session, and the interaction between Group and Session with random effects for individual rat starting points was created. There was a significant effect of Group (estimate: 40.20, CI: 8.44–71.96, $P = 0.022$), but no effect of Session (estimate: 10.95, CI: −3.16–25.06, $P = 0.148$). However, there was a significant interaction between Group and Session (estimate: −32.10, CI: −52.05–(−12.15), $P = 0.006$). There was slight increase in magazine entries made over extinction sessions, but there was no effect of group, nor an interaction between group and session, was observed (Supplemental Fig. 1B).

Magazine entries remained very low. Together, these data indicate that, across the pre- and postdevaluation extinction test days, Group Saline-Pellet displayed more ST behavior than Group LiCl-Pellet but there were equivalent levels of ST between sessions. Ultimately, the interaction between Group and Session shows the differential drop in ST behavior between Group LiCl-Pellet and Group Saline-Pellet over sessions that is indicative of sensitivity to outcome devaluation.

Devaluation sensitivity in reacquisition

Next, responding during pre- and postdevaluation reacquisition sessions was compared. Predevaluation ST rates in reacquisition are presented in Figure 1E. A linear mixed model was created as above. The results here were similar to those seen for extinction test days, as there was a significant effect of Group (estimate: 25.11, CI: 1.95–48.28, $P = 0.043$), but no effect of Session (estimate: 6.19, CI: −3.22–15.61, $P = 0.215$). However, there was a significant interaction between Group and Session (estimate: −25.14, CI: −38.46–(−11.83), $P = 0.002$). The two groups differed over time in how they changed their ST rates, with Group LiCl-Pellet showing less ST than Group Saline-Pellet after devaluation. Magazine entries were also examined; they were low and we did not observe any effects (Supplemental Fig. 1C).

To confirm the thoroughness of the outcome devaluation procedure, the number of pellet reinforcers remaining in magazine food cups following the postdevaluation reacquisition test was recorded and are presented in Figure 1F. As these data are not normally distributed, a Wilcoxon rank-sum test with continuity correction was performed and revealed a significant effect of Group (W = 64, $P < 0.001$), meaning that animals in Group LiCl-Pellet ate significantly less than animals in Group Saline-Pellet during the postdevaluation reacquisition probe session.

Overall, these data show that the outcome devaluation protocol used here significantly decreased ST rates of animals in the LiCl-paired condition. This was true both in extinction and during rewarded reacquisition sessions and is further supported by animals' rejection of pellets during the postdevaluation reacquisition test.

Experiment 2

In the previous experiment, rats given LiCl injections decreased their ST rate compared to animals given saline. The possibility remains that animals experiencing LiCl injections in the conditioning chambers could have developed a general aversion or fear response to the chamber context. If animals developed a conditioned aversion to the chambers due to the LiCl experience alone, it could manifest as decreased ST during the test sessions. To test this possibility, rats in Experiment 2 ($n = 16$) underwent training as before. For the devaluation procedure, all rats received LiCl injections in the holding boxes for the first two injections and conditioning chambers for the last three injections (as in Experiment 1). Half of the subjects were given pellet reinforcers before each injection (Group LiCl-Pellet, $n = 8$), while the other half of the subjects received nothing before injection (Group LiCl only, $n = 8$). Magazine entry data throughout the experiment was examined and all figures and statistics can be found in Supplemental Figure 1 and Supplemental Table 1. If an aversion to the devaluation context (conditioning chambers included) was responsible for the decreased ST we observed in Experiment 1, then in Experiment 2 both groups should show a similar decrease in ST after extinction.

Acquisition

The mean ppm over the course of training are presented in Figure 2B. The same linear mixed model structures used in Experiment 1

were used here to compare CS+ and CS− responding between groups over time.

There was no effect of Group (estimate: 0.95 ppm; CI: −1.71–3.62; $P = 0.49$) but there was a significant effect of logSession (estimate: 2.62 ppm; CI: 0.69–4.56; $P = 0.015$). Importantly, there was a significant effect of CS-type on responding (estimate: 3.99; CI: 1.40–6.57; $P = 0.003$). There were not significant interactions between Group and logSession (estimate: −0.42 ppm; CI: −2.36–1.52; $P = 0.674$) nor Group and CS-type (estimate: 1.01; CI: −1.58–3.59; $P = 0.45$), meaning that Group LiCl-Pellet and Group LiCl-Only did not differ in how they increased lever interaction rates over training sessions nor in their preferential interaction with the CS+ over the CS−. However, there was a significant interaction between CS-type and logSession (estimate: 4.60; CI: 3.18–6.02; $P < 0.001$), showing that CS+ interactions increased over training sessions. Finally, there was no significant interaction between CS-type, Group, and logSession (estimate: −0.94; CI: −2.36–0.48; $P = 0.20$), showing that both groups similarly increased responding on the CS+ over sessions. Additionally, magazine entries made by animals in both groups similarly decreased over sessions (see Supplemental Fig. 1D).

Outcome devaluation

Devaluation procedures were the same as in Experiment 1, with the exception that Group LiCl-Only received no pellets before LiCl injections. The mean percentage of pellets consumed on each devaluation day by group is presented in Figure 2C. Because animals in Group LiCl-Only did not receive any pellets, their consumption could not be analyzed. Therefore, analysis of consumption for Group LiCl-Pellet alone was conducted. A generalized linear mixed model with fixed effect of Session and random intercepts for individual rats was created.

This model revealed an effect of Session (OR: 0.34, CI: 0.30–0.38, $P < 0.001$), indicating that, as sessions progressed, the likelihood of pellet consumption decreases. Similar to Experiment 1, there is a relative increase in the proportion of pellets consumed from Day 2 to Day 3, the first day that animals were given pellet access in operant chambers during outcome devaluation (V = 0, $P = 0.014$), indicating that animals may not have fully generalized the taste aversion from one devaluation context to the other.
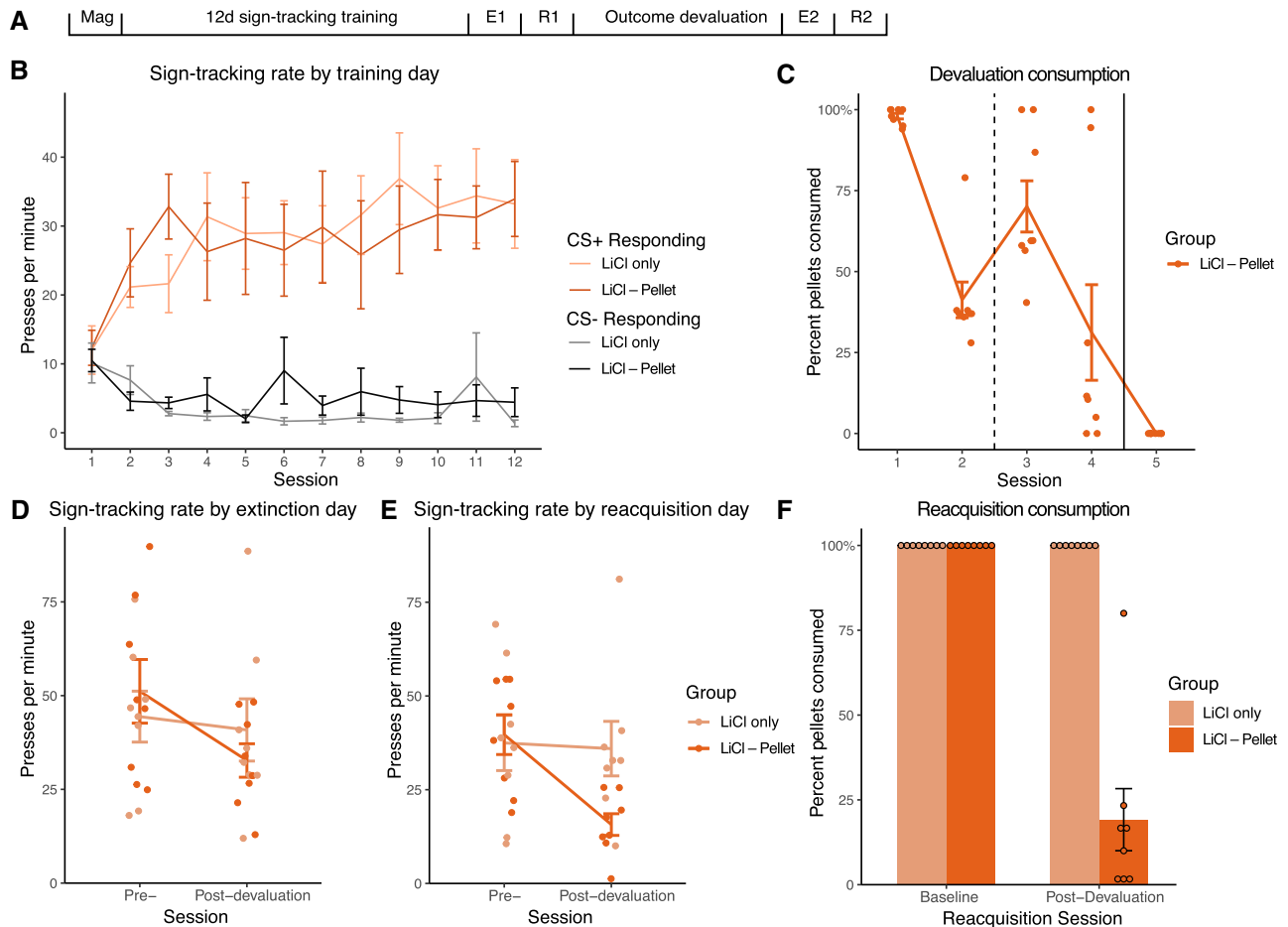
Devaluation sensitivity in extinction

CS+ responding during pre- and postdevaluation extinction sessions are compared in Figure 2D. A linear mixed model was created as above. There was no effect of Group (estimate: 21.65, CI: −4.57–47.87, $P = 0.115$) or Session (estimate: −3.55, CI: −12.77–5.67, $P = 0.461$), but there was a significant effect of the interaction between Group and Session (estimate: −14.90, CI: −27.93–(−1.87), $P = 0.040$). Magazine entries were also examined with no effects observed (see Supplemental Fig. 1E). These data indicate that, surprisingly, there was equivalent levels of ST both across days and between groups overall. However, importantly, the significant interaction is evidence that devaluation was sufficient to drive a differential change in ST between groups over time.

Devaluation sensitivity in reacquisition

Next, CS+ responding during pre- and postdevaluation reacquisition sessions are compared in Figure 2E. There was an effect of Group (estimate: 24.69, CI: 2.89–46.49, $P = 0.034$), but no effect of Session (estimate: −1.52, CI: −9.16–6.12, $P = 0.702$), and a significant interaction between Group and Session (estimate: −22.47, CI: −33.28–(−11.66), $P = 0.001$). There was slight increase in magazine entries made during reacquisition sessions over sessions, but

Figure 2. ST sensitivity to outcome devaluation was not a product of generalized negative affect following outcome devaluation. (*A*) Timeline of experimental procedures: magazine training (Mag), predevaluation extinction session (E1), predevaluation reacquisition session (R1), postdevaluation extinction session (E2), and postdevaluation extinction session (R2). (*B*) Sign-tracking behavior did not differ between groups on a discriminative sign-tracking protocol over the course of 12 training sessions. Groups readily discriminated between the CS+ and CS− and adjusted responding (presses per minute, ppm) accordingly. (*C*) Outcome devaluation conducted over the course of five sessions. Sessions 1 and 2, to the *left* of the dashed line, were conducted in external holding chambers distinct from operant chambers. Sessions 3 and 4 were conducted in operant chambers. Session 5, to the *right* of the solid line, was conducted in operant chambers but no lithium chloride was administered following pellet presentation as no pellets were consumed. Note that there is no line denoting the control group consumption as no pellets were presented to Group LiCl only during this phase of the experiment. (*D*) Sign-tracking rates (ppm) during 5-min extinction tests pre- and postdevaluation show that group LiCl-Pellet significantly decreased sign-tracking to the lever following devaluation over sessions. (*E*) Sign-tracking rates (ppm) during fully rewarded sessions show the same drop following outcome devaluation in Group LiCl-Pellet. (*F*) Pellet consumption during the pre- and postdevaluation reacquisition sessions. Predevaluation, all animals consumed all pellets, as expected. Postdevaluation, animals in Group LiCl only continued to consume all pellets while animals in group LiCl-Pellet significantly decreased pellet consumption, indicating that outcome devaluation was successful.

no effect of group, nor an interaction between group and session, was observed (and entries were low in number; Supplemental Fig. 1F). These results show that Group LiCl-Pellet decreased ST in reacquisition to a greater extent than Group LiCl-Only.

Figure 2F shows pellet consumption during reacquisition. A Wilcoxon rank-sum test with continuity correction shows a significant effect of Group on pellets consumed (W = 64, $P < 0.001$), reflecting a far greater rejection in Group LiCl-Pellet compared to Group LiCl-Only.

Overall, these data show that the outcome devaluation protocol used significantly decreased ST rates of animals selectively in the reinforcer-paired condition.

Experiment 3

Experiments 1–2 used a discriminative stimulus training paradigm, which included the presentation of CS+ (food-paired) and

CS− (nonpaired) levers. Our group and others have used this discriminative ST procedure to study the neural and behavioral basis of ST (Chang et al. 2012, 2015; Holland et al. 2014; DeAngeli et al. 2017). Another line of research to study incentive motivation has used a single-lever CS+ paradigm, where no CS− is presented (Day et al. 2006; Flagel et al. 2007; Robinson and Flagel 2009; Tomie et al. 2012; Fitzpatrick et al. 2013). For this experiment, we sought to determine whether the sensitivity of ST behavior to reward devaluation described above also extends to the single lever CS+ paradigm. To do so, we compared the sensitivity of ST to reward devaluation in a manner identical to Experiment 1 with a single CS+ lever design between groups LiCl-Pellet (n = 8) and Saline-Pellet (n = 8). For these sessions, there were 25 trials in which only the CS+ lever was inserted and the ITI was the same as before (60 sec). Magazine entry data throughout the experiment was examined and all figures and statistics can be found in Supplemental Figure 1 and Supplemental Table 1.

Acquisition

The mean ppm for each group over the course of training are presented in Figure 3B. A linear mixed model was used to analyze CS+ responding by group and session.

There was no significant effect of group (estimate: 2.89; CI: −5.90–11.67; $P = 0.523$), indicating that, overall, Group Saline-Pellet and Group LiCl-Pellet did not differ in their response rates. There was a significant effect of logSession (estimate: 5.94; CI: 1.89–9.98; $P = 0.009$), showing that over the course of training, animals significantly increased their lever interactions. However, there was no effect of the interaction between Group and logSession (estimate: 0.39; CI: −5.32–6.11; $P = 0.894$), indicating that the two groups did not differ in how they increased their lever presses over the course of training. Curiously, magazine entries did not significantly decrease over sessions in the same manner as they did in the other experiments (Supplemental Fig. 1G). We found this was the result of just one rat that both interacted with the lever and entered the magazine during its presentation.

Outcome devaluation

The mean percentage of pellets consumed on each devaluation day is presented in Figure 3C. Here, there were effects of Session (OR: 1.43, CI: 1.21–1.67, $P < 0.001$), Group (OR 265.07, CI: 94.06–747.02, $P < 0.001$), and a significant Group by Session interaction (OR: 3.71, CI: 3.16–4.35, $P < 0.001$). Again, there was a significant rebound in pellet consumption in LiCl-Pellet from Day 2 to Day 3 of outcome devaluation (V = 0, $P = 0.022$), suggestive of poor taste-aversion generalization. These effects show that rats in Group Saline-Pellet consumed more pellets than rats in Group LiCl-Pellet.

Devaluation sensitivity in extinction

Responding during pre- and postdevaluation extinction sessions are compared in Figure 3D. Again, there was no effect of Group (estimate: 24.17, CI: −6.19–54.54, $P = 0.130$) or Session (estimate: −4.92, CI: −16.84–6.99, $P = 0.430$). However, there was an effect of the interaction between Group and Session (estimate: −19.82,
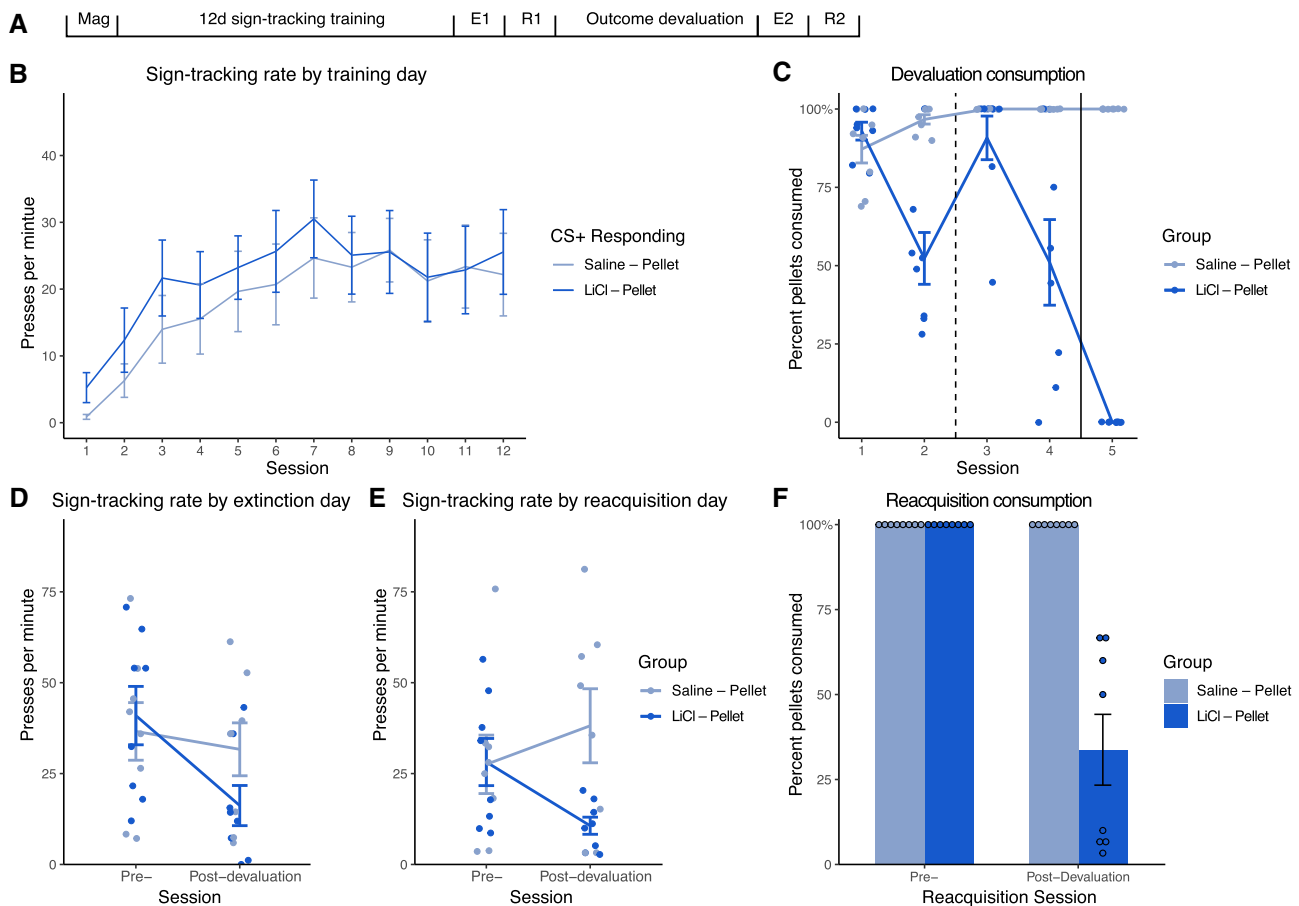
Figure 3. Sign-tracking behavior sensitivity to outcome devaluation extends to single-lever CS preparations. (*A*) Timeline of experimental procedures: magazine training (Mag), predevaluation extinction session (E1), predevaluation reacquisition session (R1), postdevaluation extinction session (E2), and postdevaluation extinction session (R2). (*B*) Sign-tracking behavior did not differ between groups on a single-cue sign-tracking protocol over the course of 12 training sessions. Groups increased responding over training days (presses per minute, ppm). (*C*) Outcome devaluation conducted over the course of five sessions. Sessions 1 and 2, to the *left* of the dashed line, were conducted in external holding chambers distinct from operant chambers. Sessions 3 and 4 were conducted in operant chambers. Session 5, *right* of the solid line, was conducted in operant chambers but no lithium chloride was administered following pellet presentation as no pellets were consumed. (*D*) Sign-tracking rates (ppm) during 5-min extinction tests pre- and postdevaluation show that Group LiCl-Pellet significantly decreased sign-tracking to the lever following devaluation over sessions. (*E*) Sign-tracking rates (ppm) during fully rewarded sessions show the same drop after outcome devaluation in Group LiCl-Pellet. (*F*) Pellet consumption during the pre- and postdevaluation reacquisition sessions. Predevaluation, all animals consumed all pellets, as expected. Postdevaluation, animals in Group Saline-Pellet continued to consume all pellets while animals in Group LiCl-Pellet significantly decreased pellet consumption, indicating that outcome devaluation was successful.

CI: −36.67–(−2.98), $P = 0.035$). Magazine entries were also examined with no effects observed (Supplemental Fig. 1H). Together, these data indicate that groups differed in ST behavior changes over sessions; namely, Group LiCl-Pellet decreased their ST in response to devaluation between the two probe sessions.

Devaluation sensitivity in reacquisition

Next, pre- and postdevaluation reacquisition sessions were compared. ST rates during these reacquisition sessions are presented in Figure 3E. There were effects of Group (estimate: 28.76, CI: 1.52–56.00, $P = 0.047$) and Session (estimate: 10.61, CI: 0.88–20.34, $P = 0.048$) and a significant interaction between Group and Session (estimate: −28.13, CI: −41.90–(−14.36), $P = 0.001$). Magazine entries were also examined with no effects observed (Supplemental Fig. 1I). Reacquisition consumption is presented in Figure 3F. A Wilcoxon rank-sum test with continuity correction showed a significant effect of Group on pellets consumed ($W = 64$, $P < 0.001$). Overall, these data show that the outcome devaluation protocol used here significantly decreased ST rates for LiCl-Pellet animals during the reacquisition period.

Experiment 4

As noted, one factor that differed in prior studies on this topic was the context in which reinforcer-LiCl pairings were administered. However, the effect of devaluation context on the ST response has not been explicitly examined. In the above Experiments, we generally found a sensitivity of ST behavior to reward devaluation when the devaluation procedure was done in both the transport boxes and operant chamber to encourage generalization of the learned aversion. In parallel with those experiments, data were acquired for a separate study using designer receptors exclusively activated by designer drugs (DREADD) manipulations to examine the neural basis of flexibility in ST behavior. Control subjects in the DREADD study are included here in Experiment 4 and are not published elsewhere. See Table 1 for a list of subjects.

Experiment 4 explicitly addressed the context question by asking whether the ST response is more sensitive to devaluation when the devaluation takes place in the training context (IN group, $n = 10$) than when the devaluation is done in a different environment (OUT group, $n = 10$). Acquisition of discriminative ST took place as in Experiments 1–3. After acquisition, animals in group IN underwent reinforcer-LiCl pairings inside the operant chambers. Animals in group OUT underwent reinforcer-LiCl pairings in empty plastic cages in a separate location. If the association between LiCl and the reinforcer is generalized equally well in the test sessions for both groups, then the amount of ST after devaluation should be similar between conditions. If instead, generalization works best for animals given LiCl-reward pairings in the operant chamber, then the Group IN ought to show outcome devaluation sensitivity. Magazine entry data throughout the experiment was examined and all figures and statistics can be found in Supplemental Figure 1 and Supplemental Table 1.

Acquisition

The mean ppm over sessions by group and by CS-type is presented in Figure 4B. A linear mixed model was created as above. There was

Table 1. A summary table of surgery designations for Group IN and Group OUT animals used in Experiment 4

	Control virus injected	No surgery	Total animals
Group IN	4	6	10
Group OUT	8	2	10

no effect of Group (estimate: 2.51; CI: −0.34–5.37; $P = 0.09$) nor logSession (estimate: 1.27; CI: −0.28–2.82; $P = 0.12$) on responding. Importantly, there was a significant effect of CS-type (estimate: 2.57; CI: 0.30–4.84; $P = 0.027$), showing that animals discriminated between the CS+ and CS− and adjusted responding appropriately. Additionally, there were no significant interactions between Group and logSession (estimate: −0.50; CI: −2.05–1.05; $P = 0.53$) nor Group and CS-type (estimate: 0.89; CI: −1.38–3.16; $P = 0.45$). There was a significant interaction between CS-type and logSession (estimate: 3.51; CI: 2.25–4.76; $P < 0.001$) but no significant interaction between CS-type, Group, and logSession (estimate: 0.05; CI: −1.20–1.30; $P = 0.93$). Magazine entries made by groups similarly decreased over training sessions (Supplemental Fig. 1J). These results indicate that there was no overall difference between groups during acquisition and that animals readily discriminated between the CS+ and CS− overall and over training sessions.
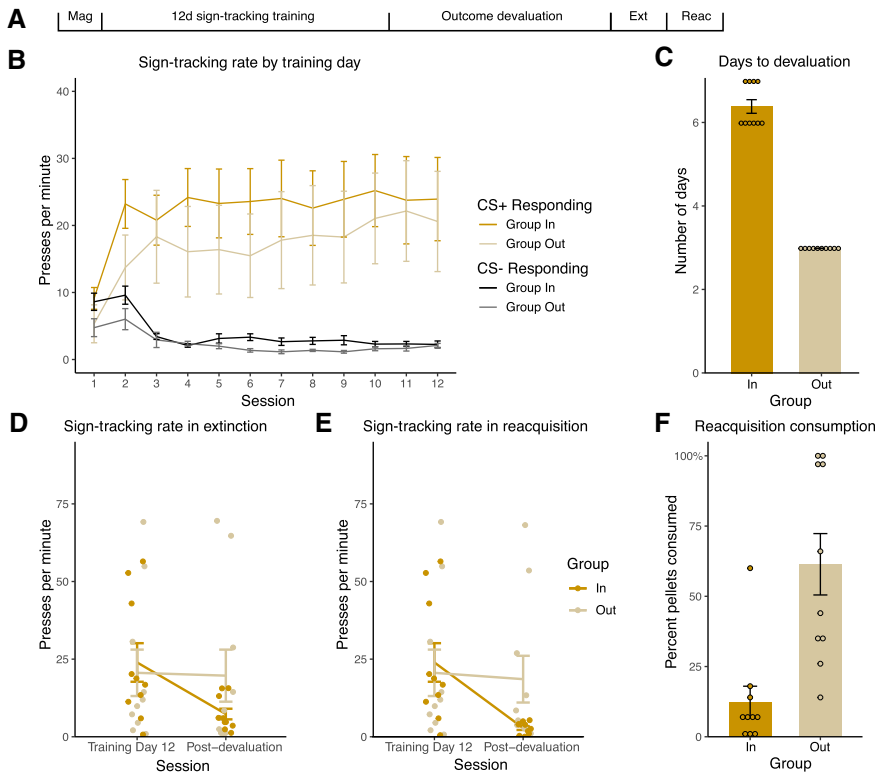
Outcome devaluation

The mean number of days required for animals in the two groups to reject all but one pellet is presented in Figure 4C. Interestingly, it took animals in Group IN significantly more sessions to reject pellets ($t_{(9)} = −20.82$, $P < 0.001$). There are a couple of possibilities to explain this difference in time to reach complete devaluation. First, the animals in Group OUT received a larger amount of food to consume than the animals in Group IN received. Group IN rats were given a smaller amount of total food to avoid clogging the food delivery tube in later pairings. This difference was most pronounced for the first pairing, in which rats in Group OUT ate close to 10 g on average, whereas rats in Group IN ate ∼3.5 g. Therefore, greater consumption during the first pairing may have resulted in stronger taste aversion learning, and a faster decline to zero consumption. A second possibility, not exclusive of the first, is that taste aversion learning in the operant chambers (Group IN) may have been slowed due to latent inhibition. Repeated prior experience of the rewarded outcome with the conditioning chambers may have interfered with new learning that the grain pellets, when delivered to the food magazine, now result in illness. We emphasize that both groups reached a point of essentially zero consumption, reflecting a complete devaluation of the pellet reinforcer in their respective contexts.

Devaluation sensitivity in extinction

Mean lever press rates by group and day are presented in Figure 4D. Note that lever press rates shown are for the final training session, Training Day 12, and the extinction probe session. While there was no effect of Group (estimate: −19.03, CI: −43.30–5.24, $P = 0.132$), there were significant effects of Session (estimate: −16.61, CI: −25.26–(−7.96), $P = 0.001$) and the interaction between Group and Session (estimate: 15.70, CI: 3.46–27.93, $P = 0.021$). Magazine entries were also examined with no effects observed (Supplemental Fig. 1K). This shows that, overall, the animals in the two groups did not differ in their ST rates but there was a drop in ST following outcome devaluation, showing that across all subjects, devaluation reduced incentive lever pressing. Importantly, the significant interaction between Group and Session shows that animals in Group IN decrease ST behavior *more* than those animals in Group OUT.

Devaluation sensitivity in reacquisition

Mean lever press rates by group and day are presented in Figure 4E. Again, note that lever press rates shown are for the final training session (Training Day 12) and the postdevaluation reacquisition session. Overall, there was no effect of Group (estimate: −22.42,

Figure 4. ST sensitivity to outcome devaluation is dependent on outcome devaluation context. (*A*) Timeline of experimental procedures: magazine training (Mag), postdevaluation extinction session (Ext), and postdevaluation extinction session (Reac). (*B*) Sign-tracking behavior did not differ between groups on a discriminative sign-tracking protocol over the course of 12 training sessions. Groups readily discriminated between the CS+ and CS− and adjusted responding (presses per minute, ppm) accordingly. (*C*) Average number of sessions needed to reduce individual rat pellet consumption to less than or equal to one pellet, shown by group. (*D*) Sign-tracking rates (ppm) during brief 5-min extinction tests on Training Day 12 and postdevaluation shows that Group IN significantly decreased sign-tracking to the lever following devaluation over sessions. (*E*) Sign-tracking rates (ppm) during fully rewarded sessions show the same drop following outcome devaluation in Group IN. (*F*) Pellet consumption during the postdevaluation reacquisition session. Animals in Group OUT consumed a greater proportion of pellets than animals in Group IN, indicating that outcome devaluation was more thorough for Group IN.

CI: −46.17–1.34, $P = 0.072$). However, there were significant effects of Session (estimate: −21.12, CI: −29.82–(−12.42), $P < 0.001$) and the interaction between Group and Session (estimate: 19.08, CI: 6.78–31.38, $P = 0.006$). While there was a drop in ST following outcome devaluation across all animals, the animals in Group IN decreased their ST behavior more than animals in Group OUT over sessions Indicating that devaluation in the operant chambers was more effective. Magazine entries were also examined, and there were no effects (Supplemental Fig. 1L).

The mean percentage of pellets consumed by each group is shown in Figure 4F. A Wilcoxon rank-sum test shows that animals in Group IN ate significantly less than animals in Group OUT ($W = 6.5$, $P = 0.001$). Thus, the qualitatively equal aversion established to the reward following LiCl pairings in both groups did not last or generalize as well to the task if pairings were done outside of the operant chamber.

Discussion

Cues that are reliably paired with a reward can acquire motivational properties, and this phenomenon is exemplified in incentive ST behavior. In ST, animals approach and interact with the reward-predicting cue, often licking, gnawing, and biting it, as

though it possessed sensory and motivational properties of the reward itself. There has been a surge of interest in characterizing the behavioral profile of ST animals (Hughson et al. 2019) in part because the ST phenotype is thought to indicate vulnerability to addiction, and it can be used as a preclinical model for the excessive motivational pull of drug-predicting stimuli (Saunders and Robinson 2013; Tomie and Sharma 2013; Huys et al. 2014). For these reasons, it is important to characterize the mechanisms of ST, and the conditions under which ST may respond to environmental or behavioral manipulations.

Here we conducted a series of experiments that investigated the effects of LiCl outcome devaluation on ST performance. Our results show that when devaluation includes multiple pairings in the training context, ST behavior is indeed sensitive to reinforcer devaluation. In Experiments 1, 3, and 4, ST behavior significantly decreased in extinction for animals that received LiCl + reinforcer pairings that included pairings in the training context. In Experiment 2, there was a trend toward lower responding in Group LiCl-Pellet, but this trend did not reach significance. For control animals who did not receive LiCl paired with the reinforcer, or who received those pairings exclusively outside of the conditioning context (Group OUT in Experiment 4), the average ST rate was largely unchanged in extinction (Experiments 2–4) or was even slightly elevated (Experiment 1). Overall, these results agree with classic demonstrations of reinforcer devaluation effects on Pavlovian conditioned responses (Holland and Straub 1979; Hatfield et al. 1996; Gallagher et al. 1999) and on goal-directed actions (Dickinson 1985). They indicate that ST behavior, at least when it comes to posttraining outcome manipulations, is not different from other forms of Pavlovian conditioning. The fact that ST was devaluation-sensitive—even after extensive training (12 d)—supports the notion that ST behavior is governed, at least in part, by an expectation of the outcome.

Devaluation sensitivity is often talked about in an absolute manner. However, it should be kept in mind that any test of outcome sensitivity is conducted using a specific devaluation protocol, and the parameters for conducting outcome devaluation vary considerably. The most common techniques used to manipulate the desirability of the outcome are selective satiety (Holland and Rescorla 1975; Gremel and Costa 2013; De Tommaso et al. 2017) and reinforcer devaluation with LiCl (Adams 1982; Smith et al. 2012), although there are others, including high-speed rotation to induce nausea (Holland and Straub 1979) or outcome inflation by means of extended food deprivation (Quinn et al. 2013) or nutrient depletion (Berridge et al. 1984; Tordoff 2001). Just considering devaluation with LiCl, protocols differ along a number of dimensions, including the type of reinforcer used in training, the number of devaluation pairings, the concentration of LiCl, the method of delivery of the reinforcer, the type of control group used, and the context in which devaluation takes place. Previous work has shown

that these factors can result in meaningful differences in task performance under postdevaluation extinction conditions. For example, a liquid reinforcer delivered directly into the oral cavity is better able to engender a complete devaluation effect than is a traditionally consumed reinforcer (Colwill and Rescorla 1990). In a different example, simple reexposure to the devalued outcome before testing produced a greater devaluation effect than a single pairing alone (Lopez et al. 1992). Findings like these indicate that it is important to take into account the specific procedure when comparing outcome devaluation effects across studies.

There are at least a couple of steps that need to take place for devaluation to affect behavior. First, the aversion to the outcome learned during the devaluation phase must transfer to the testing context. This updated outcome value must be represented or retrieved in some way from memory and be "online" during the extinction test if behavior is to be sensitive. Even if devaluation is performed, say, in the same environment as the task, the devaluation experience is subtly different than task performance and could result in less-than-perfect outcome sensitivity as a result. Second, knowledge about the outcome must influence the ST response itself. Even if the outcome is recognized to be aversive in that situation, that information needs to then influence behavior. This is similar to the distinction between goal-directed actions and habits. The lever cue might trigger a stimulus-response association that overrides any consideration of outcome value. Importantly, the presence of the behavioral response (e.g., ST) does not unequivocally mean that responding is habitual or inflexible. It could be due to a failure of generalization (first step above). This can be measured by offering animals access to the aversive outcome in a reacquisition test conducted in the same context as extinction. If the now-aversive outcome is rejected, then one can infer that devaluation successfully transferred to the testing context, whether or not the conditioned response was suppressed. In principle, given that the aversion must generalize to the task environment as well as to the behavior leading to the outcome, it may be important to consider generalization to the behavior itself as a factor involved in interpreting the behavior as goal-directed vs. habitual (beyond the underlying associative structure of behavior).

Historically, investigators have come to varying conclusions about the effectiveness of outcome devaluation tests. Colwill and Rescorla (1985) point out that numerous studies, even at that time, found diverse evidence either *for* (Chen and Amsel 1980; Adams and Dickinson 1981; Adams 1982; Dickinson et al. 1983; St. Claire-Smith and MacLaren 1983) or *against* (Garcia et al. 1970; Morrison and Collyer 1974; Holman 1975; Adams 1980; Wilson et al. 1981) outcome devaluation effects on behavior. Some variables, like the type of reinforcement schedule (e.g., a variable interval encourages habits more rapidly than a fixed ratio schedule) and amount of training, could account for some of these differences (Dickinson et al. 1983; Colwill 1988), but not all (e.g., Colwill and Rescorla 1985). Our central result, that ST is sensitive to outcome devaluation, stands in contrast to several recent reports to the contrary (Morrison et al. 2015; Nasser et al. 2015; Patitucci et al. 2016). One important procedural difference between these recent reports and our experiments (as well as that of Derman et al. 2018, below) is the context of the devaluation pairings.

In the study by Nasser et al. (2015), the authors examined the relationship between Pavlovian outcome devaluation and ST tendency. Animals learned to respond to a light cue with standard Pavlovian conditioning. They then received two pairings of food reward with LiCl in their home cage. In an extinction session, rats showed a modest but significant difference in the time spent in the food cup between LiCl paired and unpaired groups (Fig. 1C in that paper). Later, rats were trained in a discriminative autoshaping procedure. Their terminal degree of ST was correlated with their response to the Pavlovian light cue (devaluation

insensitivity). The authors concluded from this study that ST rats have a general difficulty displaying flexible behavior. However, the devaluation sensitivity of ST specifically was not measured. In addition, it may be that a stronger devaluation protocol (in the test chambers) would have resulted in ST rats significantly reducing their responding. We can conclude from this study that non-ST rats were responsible for a modest devaluation effect, and that they showed more flexibility than ST rats with this protocol.

The study by Morrison et al. (2015) measured the devaluation sensitivity of ST directly. Rats underwent 7–15 d of autoshaping training for sucrose reward and then received a single pairing of sucrose with LiCl in their home cages. Rats then underwent a probe session in extinction and their sucrose consumption was measured again afterward in the home cage. The authors found that ST tendency actually increased after reward devaluation, as least in comparison with GT tendency. In the probe session, goal-trackers emitted fewer entries into the reward magazine, with a concomitant increase in ST behavior that was perhaps compensatory. Sign-trackers showed little change in either behavior, although the subject pool had few, if any, conventionally defined ST rats (i.e., Pavlovian Conditioned Approach index between 0.5 and 1.0). The devaluation procedure here was relatively weak, involving one pairing in the home cage. Postdevaluation consumption of the devalued reward was not measured in the test chambers, so it is unknown if the subjects generalized their conditioned taste aversion (CTA) to the testing context. These results are somewhat in conflict with the study by Nasser et al. in that here, GT rats showed more lever-directed behavior in the test session.

Overall, both the studies by Nasser and Morrison found a distribution of responses to outcome devaluation, with ST animals showing little change in behavior postdevaluation. However, both studies used relatively mild devaluation protocols; either one or two pairings with LiCl in the home cage environment. Therefore, the persistence of ST in these studies may be due to a failure for animals to transfer devaluation learning to the testing context. In the case of Morrison et al. consumption was reduced, but not eliminated in a postdevaluation test in the home cage, and not measured in the testing context. It is important to confirm that there is a strong rejection of the devalued food reward in the testing context, as any residual value of the food reinforcer could support conditioned responding in extinction (Colwill and Rescorla 1985).

This latter point is also relevant for a study conducted by our group (Smedley and Smith 2018), in which it was found that ST behavior in extinction was unchanged after extensive LiCl devaluation (3–5 pairings). These pairings took place in a nontask environment (the transport boxes). Surprisingly, ST rate during a reacquisition session was also largely unchanged; that is, even with the opportunity to sample the devalued food after its associated action in the task context. In this study, consumption of the devalued food reward was measured in the devaluation context and not in the operant chambers. In a reacquisition test, rats that received two lever stimuli in sequence ate nearly all pellets (∼45/50 pellets consumed, Fig. 4D of that paper). However, animals that received only one lever stimulus consumed fewer pellets, but still far more than half (∼35/50 pellets consumed, Fig. 2D of that paper). The difference in consumption between groups that received different cues indicates that task differences can affect how food aversion transfers to the task.

Overall, these results agree with the prior two studies in that devaluation pairings outside of the test chambers results in relatively little change in ST behavior in extinction, but this may be attributable to a failure of the conditioned food aversion to transfer to the task context. In a different study, using a conditioned lever stimulus—similar to autoshaping—devaluation likewise resulted in little change in ST behavior in extinction (Vandaele et al. 2017).

The amount of consumption of the devalued food was high (∼50%, measured in the operant chamber), and the LiCl pairings were performed in the home cage. These results match our data from Experiment 4, in which rats that underwent devaluation in a different environment (Group OUT) ate well over 50% of the pellets in reacquisition. Unsurprisingly, rats in Group OUT also showed no change in ST during either extinction or reacquisition. These considerations pose a problem for the interpretation that ST is devaluation insensitive and inflexible in the classical sense (Morrison et al. 2015; Nasser et al. 2015). Instead, the insensitivity of ST to devaluation is better understood as a failure to generalize learning—in this case, CTA—when learning takes place in a different context.

Our results more closely agree with those of Derman et al. (2018). In that study, rats were trained on a two-lever ST task in which every lever was paired with a different food outcome. After 12 d of training, rats underwent five cycles of LiCl pairing *in the operant chambers*. ST was significantly lower for the lever paired with the devalued reward as compared to the lever paired with the nondevalued reward (Fig. 1A; Derman et al. 2018). In a between-subjects experiment using a discriminative autoshaping task (similar to ours) with the same method of devaluation as before, ST rate was reduced by ∼75% in the extinction sessions. This result matches our data presented here. When the devaluation protocol contains several pairings in the operant chamber, ST behavior in extinction was significantly reduced (Experiments 1, 3, and 4). This is a consequence of the fact that taste aversion successfully transferred to the task context, as measured by reacquisition consumption. Likewise, when taste aversion transfers, ST during *reacquisition* is further reduced over and above the reduction in extinction (Experiments 1–4), perhaps reflecting incentive learning processes (Lopez et al. 1992). When instead, the taste aversion does not transfer to the task context, because of a too-great dissimilarity in the conditioning and testing environments, then ST remains high. Thus, transfer of the taste aversion to the task and the effect of the aversion on ST seem to be highly related. When rats generalize taste aversion to the task, ST behavior is reduced, a signature of goal-directed or model-based behavioral control. This highlights the importance of taking context effects into account when making interpretations about the underlying associative structure of the behavior (i.e., habit vs. goal-directed).

While it is well known that taste-nausea associations are relatively easy to condition (in comparison to other classes of events—the "Garcia effect"), the role of contextual cues in CTA learning is not well understood. There is extensive evidence that the ability of new reward-related information to reduce responding in tasks can be remarkably context-dependent, such as for extinction of instrumental learning (Bouton and Todd 2014). Interestingly, there is evidence that contextual factors (also called "exteroceptive stimuli") are important for CTA learning as well. A learned aversion to drinking saccharin solution was greatly attenuated when animals were tested in a dissimilar context: importantly, the aversion they initially learned was still active when returned to the original context (Archer et al. 1979, 1984; Sjödén and Archer 1981). This effect parallels that of some of the studies mentioned here and our data from Experiment 4 (Fig. 4F), in which rejection of the devalued food pellets in reacquisition is nearly complete when tested "in-context," but greatly diminished or absent when tested "out-of-context."

An important consideration to note in Experiment 4 lies in how many reinforcer pairings were given during the devaluation protocol. Group OUT received 3 pairings, and Group IN received 6–7 pairings. While both groups reached the same endpoint of rejecting the devalued pellets during taste-aversion training, it remains possible that the greater number of pairings given to group IN could have resulted in a longer-lasting taste aversion memory, and this could in part account for the greater pellet consumption in reacquisition among rats in Group OUT. Group IN rats took longer to reach the devaluation threshold, and therefore they received more pairings. This may be due to a process such as latent inhibition (Turgeon and Reichstein 2002), whereby prior experience with pellet consumption in the conditioning chambers (i.e., the ST acquisition phase) slowed down the process of acquiring the new pellet-nausea association. Further experiments are needed to isolate number of pairings and the terminal degree of devaluation as factors that contribute to the longevity of conditioned taste aversion.

Devaluation sensitivity is a graded measure and it is most accurate to imagine it on a continuum, rather than categorically. In instrumental behavior, responding after devaluation can be divided into a sensitive, goal-directed component, and a habitual, insensitive component (Thrailkill and Bouton 2015). The ability of devaluation learning to affect performance depends on the nature of the underlying association (goal directed vs. habitual) as well as the ability for the aversion learning to transfer to the testing context, perhaps analogous to a generalization gradient. It may be that the relative inability to generalize outcome value learning itself is a signature of habitual behavior. For example, Thrailkill and Bouton (2015) found that the habitual component of instrumental behavior was more sensitive to a switch in context than the goal-directed component. Additionally, another way to conduct outcome devaluation is via sensory-specific satiety and this, too, shows signs of context dependency (Parkes et al. 2016). Although ST is considered as having a major Pavlovian learning component, there may be parallels between the context dependence of instrumental learning and extinction and what we review here for ST. Changes to outcome value, when divorced from the learning context, have relatively little impact on postdevaluation ST (Morrison et al. 2015; Vandaele et al. 2017; Smedley and Smith 2018). When outcome value is strongly reduced in the training context, ST behavior shows flexibility and sensitivity to it (our data here; (Derman et al. 2018). Importantly, in these data and in much of the literature on outcome sensitivity in behavior, the level of devaluation in the task rarely reaches 100% (i.e., animals exhibit some intake of the devalued outcome in the task, even if minimal, when given sufficient trials). Residual responding in such cases, whether ST behavior here or instrumental behavior in other studies, could reflect either (1) a habit component, or (2) the fact that the outcome is still not *completely* devalued. In other words, it will be important in studies of habits using the outcome devaluation procedure to consider to what extent habit-like responding after devaluation reflects a habit or rather a less-than-complete transfer of outcome devaluation knowledge to the task conditions.

Materials and Methods

Experiment 1

Subjects

Subjects were experimentally naïve male Long Evans rats obtained from Charles River ($n = 16$; Charles River), which weighed 250–300 g upon arrival. Rats were pair-housed in a climate-controlled colony room illuminated from 7:00 A.M. to 7:00 P.M. Following an acclimation period of 7 d, animals were individually housed and put on a food-restriction schedule to maintain body weights at 85% of their *ad libitum* weights for the duration of the experiment. The experiments were performed in accordance with the National Institute of Health's *Guide for the Care and Use of Laboratory Animals*; protocols were approved by the Dartmouth College Institutional Animal Care and Use Committee.

Apparatus

ST training and testing was carried out in eight identical operant conditioning chambers ($24 \times 30.5 \times 29$ cm; Med Associates),

enclosed in sound-attenuating chambers outfitted with an exhaust fan to provide airflow and background noise (~ 68 dB). The chambers were illuminated by a house light on the back wall of the chamber. Each chamber contained a recessed food magazine in the center of the front wall. Retractable levers (Med Associates model: ENV-112CM) were positioned on either side of the food magazine. Lever deflections were automatically recorded, and magazine entries were recorded through breaks of an infrared beam. Data were acquired through MED-PC software (Med Associates).

Behavioral training

The sequence of training phases is presented in Figure 1A. All rats first received a single 30-min acclimation session of magazine training where grain pellets (Bio Serv, Product #F0165, 45 mg dustless precision pellets: Protein 21.3%, Fat 3.8%, Carbohydrate 54.0%) were delivered freely on a random-time 30 sec (RT30) schedule. Next, rats underwent 12 daily, 60-min sessions of discriminative training. During each training session, rats received 25 conditioned stimulus (CS+) trials and 25 CS– trials, such that no more than two of the same trial type occurred in a sequence. Each trial consisted of a 10 sec lever presentation, but only CS+ trials were followed by delivery of two food pellets upon lever retraction. The assignment of left and right levers to CS+ and CS– identities was counterbalanced within groups of animals but held constant per animal. Training was followed by one abbreviated predevaluation test session (5 CS+, 5 CS– presentations) conducted in extinction conditions to establish a baseline level of responding. The next day, rats were given one rewarded reacquisition session (25 CS+, 25 CS– presentations).

Outcome devaluation and postdevaluation testing

Rats were split into two groups based on mean lever press rates and standard error of the mean such that groups had matched responding levels by Day 12 of training. After group assignment, rats were exposed to an outcome devaluation procedure. Devaluation of the grain pellets was carried out in two phases; rats received up to five pairings. The first and second pairings took place in plastic holding boxes, as previously described (Smedley and Smith 2018), while subsequent pairings took place in the Med Associates conditioning chambers. This procedure was chosen because it was thought that a variety of contexts for devaluation would allow animals to best generalize the CTA beyond the holding chambers. In separate experiments in our laboratory, we found that doing devaluation exclusively in the holding boxes or in empty cages resulted in many animals consuming many, if not most, of the pellets during reacquisition (data not shown). We speculated that adding pairings inside the conditioning chambers would strengthen the association between the aversive food outcome (pellets) and the response (lever pressing) that had previously coincided with food delivery.

For the first and second pairings, animals were given 10 g of pellets in a plastic dish in clear plastic holding boxes normally used for transport between the colony and testing room. Rats were allowed 20 min to consume pellets. The plastic dishes were then removed from the holding boxes, rats were injected with either lithium chloride, termed Group LiCl-Pellet (LiCl; 0.3 M; 10 mg/kg; in deionized water) or 0.9% saline, termed Group Saline-Pellet. Then, the remainder of the pellets were weighed, and weights of the rats were recorded. Rats stayed in the boxes for an additional 20 min following the injection and were then returned to their home cages. After 48 h, this devaluation procedure was repeated.

The third and subsequent pairings were conducted in the conditioning chambers. Again, these pairings were spaced 48 h apart. For these days, pellets were delivered on the RT30 schedule previously used during magazine training. Levers were not extended during these sessions. To avoid clogging of the magazine with pellets, pairings 3–5 were successively shorter in length, as animals in Group LiCl-Pellet rejected more pellets over time which increased the likelihood of rejected pellets backing up within the delivery tube. At the conclusion of these sessions, animals were removed from the conditioning chambers, held briefly in the plastic holding boxes, and the number of pellets consumed was recorded. Then, animals were injected with either LiCl or saline, based on their group assignment, and allowed to rest for 20 min in the conditioning chambers. Once an animal consumed 1 or fewer pellets during devaluation, it was advanced to postdevaluation probe sessions. These consisted of a brief extinction session (5 CS+, 5 CS– presentations) followed by an abbreviated, fully rewarded, reacquisition session (15 CS+, 15 CS– presentations) to assess ST persistency in the face of devalued reward.

Behavioral measures and analyses

Lever deflections, magazine entries, and time spent in the magazine area were recorded through MED-PC. During outcome devaluation, pellets were weighed before and after consumption to calculate the percentage of grams consumed. All statistical tests were carried out using R (R Core Team 2016). All graphs were created through R (R: "ggplot2") and designed with Adobe Illustrator.

Zero-sum contrasts were made for categorical variables where appropriate (e.g., Group Saline-Pellet vs. Group LiCl-Pellet). Individual linear mixed models (R; "lme4") were used to analyze effects of dependent variable responding (e.g., lever presses per minute, ppm) by fixed effects of experimental group, logSession, CS-type, and the interactions between these variables while accounting for random effects of differences in individual starting press rates and individual learning rates over sessions. LogSession, created by logarithmically transforming session, was used to model training data as models using logSession were fit data statistically better than models using linear session alone as determined by Akaike information criterion for nonnested model comparison (see Smedley et al. 2019 for a similar application). Linear mixed models were fit by maximum likelihood and *t*-tests use Satterthwaite approximations of degrees of freedom (R; "lmerMod"). Linear models were analyzed with package lme4 from CRAN (Bates et al. 2015). Reported statistics include parameter estimates (β values), confidence intervals (95% confidence intervals) and *P*-values (R; "lmerTest," Kuznetsova et al. 2017).

As percentage data are not normally distributed, a generalized linear mixed model was used to analyze effects of devaluation and session on pellet consumption during outcome devaluation. Session was recentered to assess group differences in consumption on Day 5, the final day of outcome devaluation. Odds ratios, confidence intervals, and *P*-values are reported. Post-hoc Wilcoxon signed rank tests with continuity corrections were used to assess whether animals in Group LiCl-Pellet increased the proportion of pellets consumed on Day 2 and Day 3 of outcome devaluation. Notably, Day 3 of outcome devaluation is the first day that animals are reintroduced to the operant chamber context and given access to pellets.

Responding in the postdevaluation extinction session (ppm) was compared with responding in the predevaluation extinction session by creating individual linear mixed models to assess response rates by fixed effects of group, session, and the interaction between group and session while accounting for random effects of individual starting points. Responding in the postdevaluation reacquisition session was similarly compared with responding in the predevaluation reacquisition session.

Pellet consumption during postdevaluation reacquisition sessions was recorded. These data were not normally distributed. Therefore, a Wilcoxon rank-sum test with continuity correction was used to determine whether animals differed in pellet consumption based on their group treatments.

Experiment 2

Subjects and behavioral training

Subjects were 16 experimentally naïve male Long Evans rats obtained from the same vendor as in Experiment 1 and maintained under the same conditions. The apparatus was the same as in Experiment 1. Rats underwent 1 d of magazine training and 12 d of autoshaping training as in Experiment 1. Data collection methods and analyses were the same as in Experiment 1.

Outcome devaluation and postdevaluation testing

Outcome devaluation procedures were similar to those described for Experiment 2. Rats in Group LiCl-Pellet ($n = 8$) were given grain pellets (as described above) in the holding boxes (days 1–2) and operant chambers (days 3–5) before being injected with LiCl solution. Rats in Group LiCl-Only ($n = 8$) spent the same amount of time in the boxes and conditioning chambers but received no food before receiving injections of LiCl. Once animals in Group LiCl-Pellet consumed no more than 1 pellet each, no further injections of LiCl were administered, and animals were advanced to a brief postdevaluation extinction session (5 CS+, 5 CS− presentations). Following this extinction session, a reacquisition session (15 CS+, 15 CS− presentations) was administered.

Experiment 3

Subjects and behavioral training

Subjects were 16 experimentally naïve male Long Evans rats obtained from the same vendor as in Experiment 1 and maintained under the same conditions. The apparatus was the same as in Experiment 1. Data collection methods and analyses were the same as in Experiment 1, with the exception that there were no CS− trials. Sessions lasted for 30 minutes and consisted of 25 CS+ trials (left and right lever positions were counterbalanced across subjects). The intertrial interval was the same as in the other experiments (average 60 sec).

Outcome devaluation and postdevaluation testing

Outcome devaluation took place in the same manner as in Experiment 1. After pellet consumption, subjects in Group LiCl-Pellet ($n = 8$) received injections of LiCl solution, whereas subjects in Group Saline-Pellet ($n = 8$) received saline injections. Once animals in Group LiCl-Pellet rejected all pellets, no further injections of LiCl were administered, and animals were advanced to a postdevaluation extinction session (5 CS+ presentations). Following this extinction session, an abbreviated reacquisition session (15 CS+ presentations) was administered.

Experiment 4

Subjects and behavioral training

The subjects were 20 experimentally naïve male Long Evans rats obtained from the same vendor as in Experiment 1 and maintained under the same conditions. The apparatus was the same as in Experiment 1. The sequence of training phases is presented in Figure 4A. Rats underwent 1 d of magazine training and 12 d of autoshaping training in Experiment 1. Data collection methods and analyses are the same as in Experiment 1. 12 subjects underwent surgery and were injected with a control virus as part of a separate study (see above). In addition, eight control rats that did not receive surgery were added to the previous 12, for a total of 20 subjects (Table 1). Training for the autoshaping task was the same as in the above experiments.

Surgical procedures for control virus subjects

All surgeries were performed under aseptic conditions with isoflurane anesthesia, and all infusions were made with a 10 µL syringe equipped with a 36-gauge beveled needle (World Precision Instruments) and a Quintessential Stereotaxic Injector (Stoelting). Bilateral infusions of *CAV2-Cre* were made into the amygdala at −3.0 mm posterior from bregma, 5.0 mm from the midline, and 5.0 mm ventral from the skull surface. Injection volume at each site was 0.5 µL. Additionally, rats received bilateral infusions into the orbitofrontal cortex, at 3.9 mm posterior to bregma, 3.0 mm from the midline, and 4.4 mm ventral from the skull surface, of the control virus *AAV8-hSyn-DIO-mCherry* (Addgene). Injection volume at each site was 0.5 µL. Expression of these transgenes was allowed to take place over the course of at least 3 wk before the commencement of behavioral training.

Outcome devaluation and postdevaluation testing

For rats in Group IN ($n = 10$), subjects were placed in the same conditioning chambers as in training. Pellets were delivered on the RT30 schedule previously used during magazine training. The levers were not extended during these sessions. After 30 min, animals were removed from the conditioning chambers and briefly held in plastic transport boxes while the leftover pellets were removed (and later counted). Rats were injected with LiCl and returned to the conditioning chambers for at least 20 min. This process was performed again after 48 h. This consisted of up to 7 d of pellet pairings, determined by whether animals were at or below a threshold of pellet consumption defined by experimenters a priori as each group averaging less than one pellet consumed each day

For rats in Group OUT ($n = 10$), subjects were placed in clear plastic tubs, with a metal food cup that contained 10 g of grain pellets. After 20 min, the food cups were removed, and the weight of leftover pellets later measured. Rats were injected with LiCl, as above, and allowed to rest for at least 20 min before being returned to the animal colony. This process was performed again after 48 h, with animals receiving a total of three LiCl pairings.

In order to measure the persistence of ST behavior after devaluation, animals were returned to the conditioning chambers and ran the same autoshaping task in extinction. In the analyses below, only the first 5 CS+ and first 5 CS− trials were analyzed. The following day, animals ran an additional reacquisition session with pellet delivery as normal. The reacquisition session was used to test the degree to which animals were able adjust their behavior when exposed the now devalued food outcome. Animals that underwent surgery with control virus received injections of clozapine-n-oxide (CNO; 1 mg/kg; National Institute of Mental Health's Chemical Synthesis and Drug Supply Program) 30 min prior to testing on the extinction and reacquisition sessions. CNO is an inert ligand in the absence of DREADD receptors (Armbruster et al. 2007), and is expected to have no effect on behavior in control animals lacking DREADD receptors. Anecdotally, we have seen no difference in ST between control animals treated with CNO and naïve animals.

Acknowledgments

References

Adams CD. 1980. Post-conditioning devaluation of an instrumental reinforcer has no effect on extinction performance. *Q J Exp Psychol* **32:** 447–458. doi:10.1080/14640748008401838

Adams CD. 1982. Variations in the sensitivity of instrumental responding to reinforcer devaluation. *Q J Exp Psychol B* **34:** 77–98. doi:10.1080/14640748208400878

Adams CD, Dickinson A. 1981. Instrumental responding following reinforcer devaluation. *Q J Exp Psychol B Comp Physiol Psychol* **33:** 109–121. doi:10.1080/14640748108400816

Archer T, Sjödén PO, Nilsson LG, Carter N. 1979. Role of exteroceptive background context in taste-aversion conditioning and extinction. *Anim Learn Behav* **7:** 17–22. doi:10.3758/BF03209650

Archer T, Sjödén PO, Nilsson LG. 1984. The importance of contextual elements in taste-aversion learning. *Scand J Psychol* **25:** 251–257. doi:10.1111/j.1467-9450.1984.tb01016.x

Armbruster BN, Li X, Pausch MH, Herlitze S, Roth BL. 2007. Evolving the lock to fit the key to create a family of G protein-coupled receptors potently activated by an inert ligand. *Proc Natl Acad Sci* **104:** 5163–5168. doi:10.1073/pnas.0700293104

Bates D, Mächler M, Bolker B, Walker S. 2015. Fitting linear mixed-effects models using lme4. *J Stat Softw* **67:** 1–48. doi:10.18637/jss.v067.i01

Berridge KC, Flynn FW, Schulkin J, Grill HJ. 1984. Sodium depletion enhances salt palatability in rats. *Behav Neurosci* **98:** 652–660. doi:10.1037/0735-7044.98.4.652

Boakes R. 1977. Performance on learning to associate a stimulus with positive reinforcement. In *Operant-Pavlovian Interactions* (ed. Davis HH, Hurwitz H), pp. 67–97. Lawrence Erlbaum Associates, Hillsdale, NJ.

Bouton ME, Todd TP. 2014. A fundamental role for context in instrumental learning and extinction. *Behav Processes* **104:** 13–19. doi:10.1016/j.beproc.2014.02.012

Brown PL, Jenkins HM. 1968. Auto-shaping of the pigeon's key-peck. *J Exp Anal Behav* **11:** 1–8. doi:10.1901/jeab.1968.11-1

Chang SE, Smith KS. 2016. An omission procedure reorganizes the microstructure of sign-tracking while preserving incentive salience. *Learn Mem* **23:** 151–155. doi:10.1101/lm.041574.115

Chang SE, Wheeler DS, Holland PC. 2012. Effects of lesions of the amygdala central nucleus on autoshaped lever pressing. *Brain Res* **1450:** 49–56. doi:10.1016/j.brainres.2012.02.029

Chang SE, Todd TP, Bucci DJ, Smith KS. 2015. Chemogenetic manipulation of ventral pallidal neurons impairs acquisition of sign-tracking in rats. *Eur J Neurosci* **42:** 3105–3116. doi:10.1111/ejn.13103

Chang SE, Todd TP, Smith KS. 2018. Paradoxical accentuation of motivation following accumbens-pallidum disconnection. *Neurobiol Learn Mem* **149:** 39–45.

Chen JS, Amsel A. 1980. Recall (versus recognition) of taste and immunization against aversive taste anticipations based on illness. *Science* **209:** 831–833. doi:10.1126/science.7403850

Colwill RM. 1988. Associations between the discriminative stimulus and the reinforcer in instrumental learning. *J Exp Psychol Anim Behav Process* **14:** 155–164. doi:10.1037/0097-7403.14.2.155

Colwill RM, Rescorla RA. 1985. Instrumental responding remains sensitive to reinforcer devaluation after extensive training. *J Exp Psychol Anim Behav Process* **11:** 520–536. doi:10.1037/0097-7403.11.4.520

Colwill RM, Rescorla RA. 1990. Effect of reinforcer devaluation on discriminative control of instrumental behavior. *J Exp Psychol Anim Behav Process* **16:** 40–47. doi:10.1037/0097-7403.16.1.40

Davey GC, Cleland GG. 1982. Topography of signal-centered behavior in the rat: effects of deprivation state and reinforcer type. *J Exp Anal Behav* **38:** 291–304.

Day JJ, Wheeler RA, Roitman MF, Carelli RM. 2006. Nucleus accumbens neurons encode Pavlovian approach behaviors: evidence from an autoshaping paradigm. *Eur J Neurosci* **23:** 1341–1351. doi:10.1111/j.1460-9568.2006.04654.x

Dayan P, Berridge KC. 2014. Model-based and model-free Pavlovian reward learning: revaluation, revision, and revelation. *Cogn Affect Behav Neurosci* **14:** 473–492. doi:10.3758/s13415-014-0277-8

DeAngeli NE, Miller SB, Meyer HC, Bucci DJ. 2017. Increased sign-tracking behavior in adolescent rats. *Dev Psychobiol* **59:** 840–847. doi:10.1002/dev.21548

Derman RC, Schneider K, Juarez S, Delamater AR. 2018. Sign-tracking is an expectancy-mediated behavior that relies on prediction error mechanisms. *Learn Mem* **25:** 550–563. doi:10.1101/lm.047365.118

De Tommaso M, Mastropasqua T, Turatto M. 2017. The salience of a reward cue can outlast reward devaluation. *Behav Neurosci* **131:** 226–234. doi:10.1037/bne0000193

Dickinson A. 1985. Actions and Habits: the development of behavioural autonomy. *Philos Trans R Soc Lond B Biol Sci* **308:** 67–78. doi:10.1098/rstb.1985.0010

Dickinson A, Balleine B. 1994. Motivational control of goal-directed action. *Anim Learn Behav* **22:** 1–18. doi:10.3758/BF03199951

Dickinson A, Nicholas DJ, Adams CD. 1983. The effect of the instrumental training contingency on susceptibility to reinforcer devaluation. *Q J Exp Psychol B Comp Physiol Psychol* **35:** 35–51. doi:10.1080/14640748308400912

Fitzpatrick CJ, Gopalakrishnan S, Cogan ES, Yager LM, Meyer PJ, Lovic V, Saunders BT, Parker CC, Gonzales NM, Aryee E, et al. 2013. Variation in the form of Pavlovian conditioned approach behavior among outbred male Sprague-Dawley rats from different vendors and colonies: sign-tracking vs. goal-tracking. *PLoS ONE* **8:** e75042. doi:10.1371/journal.pone.0075042

Flagel SB, Watson SJ, Robinson TE, Akil H. 2007. Individual differences in the propensity to approach signals vs goals promote different adaptations in the dopamine system of rats. *Psychopharmacology (Berl)* **191:** 599–607. doi:10.1007/s00213-006-0535-8

Flagel SB, Watson SJ, Akil H, Robinson TE. 2008. Individual differences in the attribution of incentive salience to a reward-related cue: influence on cocaine sensitization. *Behav Brain Res* **186:** 48–56. doi:10.1016/j.bbr.2007.07.022

Flagel SB, Clark JJ, Robinson TE, Mayo L, Czuj A, Willuhn I, Akers CA, Clinton SM, Phillips PE, Akil H. 2011. A selective role for dopamine in stimulus-reward learning. *Nature* **469:** 53–57. doi:10.1038/nature09588

Gallagher M, McMahan RW, Schoenbaum G. 1999. Orbitofrontal cortex and representation of incentive value in associative learning. *J Neurosci* **19:** 6610–6614. doi:10.1523/JNEUROSCI.19-15-06610.1999

Garcia J, Kovner R, Green KF. 1970. Cue properties vs palatability of flavors in avoidance learning. *Psychon Sci* **20:** 313–314. doi:10.3758/BF03329085

Gremel CM, Costa RM. 2013. Orbitofrontal and striatal circuits dynamically encode the shift between goal-directed and habitual actions. *Nat Commun* **4:** 2264. doi:10.1038/ncomms3264

Hatfield T, Han JS, Conley M, Gallagher M, Holland P. 1996. Neurotoxic lesions of basolateral, but not central, amygdala interfere with Pavlovian second-order conditioning and reinforcer devaluation effects. *J Neurosci* **16:** 5256–5265. doi:10.1523/JNEUROSCI.16-16-05256.1996

Hearst E, Jenkins HM. 1974. *Sign-tracking : the stimulus-reinforcer relation and directed action*. Psychonomic Society, Austin, Texas.

Holland PC, Rescorla RA. 1975. The effect of two ways of devaluing the unconditioned stimulus after first- and second-order appetitive conditioning. *J Exp Psychol Anim Behav Process* **1:** 355–363. doi:10.1037/0097-7403.1.4.355

Holland PC, Straub JJ. 1979. Differential effects of two ways of devaluing the unconditioned stimulus after Pavlovian appetitive conditioning. *J Exp Psychol Anim Behav Process* **5:** 65–78. doi:10.1037/0097-7403.5.1.65

Holland PC, Asem JS, Galvin CP, Keeney CH, Hsu M, Miller A, Zhou V. 2014. Blocking in autoshaped lever-pressing procedures with rats. *Learn Behav* **42:** 1–21. doi:10.3758/s13420-013-0120-z

Holman EW. 1975. Some conditions for the dissociation of consummatory and instrumental behavior in rats. *Learn Motiv* **6:** 358–366. doi:10.1016/0023-9690(75)90015-6

Hughson AR, Horvath AP, Holl K, Palmer AA, Solberg Woods LC, Robinson TE, Flagel SB. 2019. Incentive salience attribution, "sensation-seeking" and "novelty-seeking" are independent traits in a large sample of male and female heterogeneous stock rats. *Sci Rep* **9:** 2351. doi:10.1038/s41598-019-39519-1

Huys QJ, Tobler PN, Hasler G, Flagel SB. 2014. The role of learning-related dopamine signals in addiction vulnerability. *Prog Brain Res* **211:** 31–77. doi:10.1016/B978-0-444-63425-2.00003-9

Kuznetsova A, Brockhoff PB, Christensen RHB. 2017. lmerTest package: tests in linear mixed effects models. *J Stat Softw* **82:** 26. doi:10.18637/jss.v082.i13

Lesaint F, Sigaud O, Flagel SB, Robinson TE, Khamassi M. 2014. Modelling individual differences in the form of Pavlovian conditioned approach responses: a dual learning systems approach with factored representations. *PLoS Comput Biol* **10:** e1003466. doi:10.1371/journal.pcbi.1003466

Lesaint F, Sigaud O, Clark JJ, Flagel SB, Khamassi M. 2015. Experimental predictions drawn from a computational model of sign-trackers and goal-trackers. *J Physiol* **109:** 78–86. doi:10.1016/j.jphysparis.2014.06.001

Locurto C, Terrace HS, Gibbon J. 1976. Autoshaping, random control, and omission training in the rat. *J Exp Anal Behav* **26:** 451–462. doi:10.1901/jeab.1976.26-451

Lopez M, Balleine B, Dickinson A. 1992. Incentive learning following reinforcer devaluation is not conditional upon the motivational state during re-exposure. *Q J Exp Psychol B* **45:** 265–284. doi:10.1080/14640749208401327

Morrison GR, Collyer R. 1974. Taste-mediated conditioned aversion to an exteroceptive stimulus following LiCl poisoning. *J Comp Physiol Psychol* **86:** 51–55. doi:10.1037/h0035947

Morrison SE, Bamkole MA, Nicola SM. 2015. Sign tracking, but not goal tracking, is resistant to outcome devaluation. *Front Neurosci* **9:** 468. doi:10.3389/fnins.2015.00468

Nasser HM, Chen YW, Fiscella K, Calu DJ. 2015. Individual variability in behavioral flexibility predicts sign-tracking tendency. *Front Behav Neurosci* **9:** 289. doi:10.3389/fnbeh.2015.00289

Parkes SL, Marchand AR, Ferreira G, Coutureau E. 2016. A time course analysis of satiety-induced instrumental outcome devaluation. *Learn Behav* **44:** 347–355. doi:10.3758/s13420-016-0226-1

Patitucci E, Nelson AJ, Dwyer DM, Honey RC. 2016. The origins of individual differences in how learning is expressed in rats: a general-process perspective. *J Exp Psychol Anim Learn Cogn* **42:** 313–324. doi:10.1037/xan0000116

Quinn JJ, Pittenger C, Lee AS, Pierson JL, Taylor JR. 2013. Striatum-dependent habits are insensitive to both increases and decreases in reinforcer value in mice. *Eur J Neurosci* **37:** 1012–1021. doi:10.1111/ejn.12106

R Core Team. 2016. R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

Robinson MJ, Berridge KC. 2013. Instant transformation of learned repulsion into motivational "wanting". *Curr Biol* **23:** 282–289. doi:10.1016/j.cub.2013.01.016

Robinson TE, Flagel SB. 2009. Dissociating the predictive and incentive motivational properties of reward-related cues through the study of individual differences. *Biol Psychiatry* **65:** 869–873. doi:10.1016/j.biopsych.2008.09.006

Saunders BT, Robinson TE. 2010. A cocaine cue acts as an incentive stimulus in some but not others: implications for addiction. *Biol Psychiatry* **67:** 730–736. doi:10.1016/j.biopsych.2009.11.015

Saunders BT, Robinson TE. 2011. Individual variation in the motivational properties of cocaine. *Neuropsychopharmacology* **36:** 1668–1676. doi:10.1038/npp.2011.48

Saunders BT, Robinson TE. 2012. The role of dopamine in the accumbens core in the expression of Pavlovian-conditioned responses. *Eur J Neurosci* **36:** 2521–2532. doi:10.1111/j.1460-9568.2012.08217.x

Saunders BT, Robinson TE. 2013. Individual variation in resisting temptation: implications for addiction. *Neurosci Biobehav Rev* **37:** 1955–1975. doi:10.1016/j.neubiorev.2013.02.008

Saunders BT, Yager LM, Robinson TE. 2013. Cue-evoked cocaine "craving": role of dopamine in the accumbens core. *J Neurosci* **33:** 13989–14000. doi:10.1523/JNEUROSCI.0450-13.2013

Sjödén P-O, Archer T. 1981. Associative and nonassociative effects of exteroceptive context in taste-aversion conditioning with rats. *Behav Neural Biol* **33:** 74–92. doi:10.1016/S0163-1047(81)92254-8

Smedley EB, Smith KS. 2018. Evidence of structure and persistence in motivational attraction to serial Pavlovian cues. *Learn Mem* **25:** 78–89. doi:10.1101/lm.046599.117

Smedley EB, DiLeo A, Smith KS. 2019. Circuit directionality for motivation: lateral accumbens-pallidum, but not pallidum-accumbens, connections regulate motivational attraction to reward cues. *Neurobiol Learn Mem* **162:** 23–35. doi:10.1016/j.nlm.2019.05.001

Smith KS, Virkud A, Deisseroth K, Graybiel AM. 2012. Reversible online control of habitual behavior by optogenetic perturbation of medial prefrontal cortex. *Proc Natl Acad Sci* **109:** 18932–18937. doi:10.1073/pnas.1216264109

St. Claire-Smith R, MacLaren D. 1983. Response preconditioning effects. *J Exp Psychol Anim Behav Process* **9:** 41–48. doi:10.1037/0097-7403.9.1.41

Thrailkill EA, Bouton ME. 2015. Contextual control of instrumental actions and habits. *J Exp Psychol Anim Learn Cogn* **41:** 69–80. doi:10.1037/xan0000045

Tomie A, Sharma N. 2013. Pavlovian sign-tracking model of alcohol abuse. *Curr Drug Abuse Rev* **6:** 201–219. doi:10.2174/18744737113069990023

Tomie A, Lincks M, Nadarajah SD, Pohorecky LA, Yu L. 2012. Pairings of lever and food induce Pavlovian conditioned approach of sign-tracking and goal-tracking in C57BL/6 mice. *Behav Brain Res* **226:** 571–578. doi:10.1016/j.bbr.2011.10.021

Tordoff MG. 2001. Calcium: taste, intake, and appetite. *Physiol Rev* **81:** 1567–1597. doi:10.1152/physrev.2001.81.4.1567

Turgeon SM, Reichstein DA. 2002. Decreased striatal c-Fos accompanies latent inhibition in a conditioned taste aversion paradigm. *Brain Res* **924:** 120–123. doi:10.1016/S0006-8993(01)03245-0

Vandaele Y, Pribut HJ, Janak PH. 2017. Lever insertion as a salient stimulus promoting insensitivity to outcome devaluation. *Front Integr Neurosci* **11:** 23. doi:10.3389/fnint.2017.00023

Wilson CL, Sherman JE, Holman EW. 1981. Aversion to the reinforcer differentially affects conditioned reinforcement and instrumental responding. *J Exp Psychol Anim Behav Process* **7:** 165–174. doi:10.1037/0097-7403.7.2.165