Rats that sign-track are resistant to Pavlovian but not instrumental extinction

Allison M. Ahrens[a*], Bryan F. Singer[a], Christopher J. Fitzpatrick[b], Jonathan D. Morrow[b,c] and

Terry E. Robinson[a,b]

Department of Psychology[a], Neuroscience Graduate Program[b], and Department of Psychiatry[c],

The University of Michigan, Ann Arbor, MI 48109, USA

*Corresponding author:

Allison M. Ahrens, Ph.D.
Department of Psychology
University of Michigan
530 Church St.
Ann Arbor, MI 48109
Tel.: +1 512 7731315
 E-mail: ahrensam@umich.edu

ABSTRACT

Individuals vary in the extent to which they attribute incentive salience to a discrete cue (conditioned stimulus; CS) that predicts reward delivery (unconditioned stimulus; US), which results in some individuals approaching and interacting with the CS (sign-trackers; STs) more than others (goal-trackers; GTs). Here we asked how periods of non-reinforcement influence conditioned responding in STs vs. GTs, in both Pavlovian and instrumental tasks. After classifying rats as STs or GTs by pairing a retractable lever (the CS) with the delivery of a food pellet (US), we introduced periods of non-reinforcement, first by simply withholding the US (i.e., extinction training; experiment 1), then by signaling alternating periods of reward (R) and non-reward (NR) within the same session (experiments 2 and 3). We also examined how alternating R and NR periods influenced instrumental responding for food (experiment 4). STs and GTs did not differ in their ability to discriminate between R and NR periods in the instrumental task. However, in Pavlovian settings STs and GTs responded to periods of non-reward very differently. Relative to STs, GTs very rapidly modified their behavior in response to periods of non-reward, showing much faster extinction and better and faster discrimination between R and NR conditions. These results highlight differences between Pavlovian and instrumental extinction learning, and suggest that if a Pavlovian CS is strongly attributed with incentive salience, as in STs, it may continue to bias attention toward it, and to facilitate persistent and relatively inflexible responding, even when it is no longer followed by reward.

KEYWORDS: Extinction; Pavlovian conditioning; Sign tracking; Goal tracking; Incentive motivation**;** Occasion setting; Discriminative stimuli

1. INTRODUCTION

In appetitive Pavlovian conditioning an initially neutral cue is paired with the delivery of a reward (the unconditioned stimulus; US), and as they become associated, the cue (conditioned stimulus; CS) may come to elicit a variety of conditioned responses (CRs) that anticipate reward delivery. CRs can take many forms, depending not only on the nature of the US, but also the properties of the CS (Holland 1977), including whether the CS is attributed with incentive salience (Anselme et al. 2013; Flagel et al. 2009; Meyer et al. 2014; Robinson et al. 2014a; Robinson and Flagel 2009; Saunders and Robinson 2013). For example, if presentation of a discrete, localizable CS is consistently followed by food reward in a response-independent manner some animals learn to approach the CS itself (termed "sign-trackers"; STs) (Hearst 1974), whilst others approach the site of food delivery (termed "goal trackers"; GTs) (Boakes 1977; Zener 1937). Although both STs and GTs learn the predictive value of the CS – it comes to evoke an anticipatory CR in both – it has been suggested that the form of the CR differs in part because STs attribute greater incentive salience to the CS than GTs (Flagel et al. 2009; Meyer et al. 2012a; Robinson and Flagel 2009; Saunders and Robinson 2013; Tomie et al. 2000; Beckmann et al. 2015; Anselme 2014). Not only are STs more attracted to the CS, but it is a more effective conditioned reinforcer and more effective in renewing reward-seeking behavior in STs than GTs (Robinson and Flagel 2009; Saunders et al. 2013; Yager and Robinson 2010). Furthermore, performance on the Pavlovian conditioned approach (PCA) task using food reward predicts the extent to which discrete drug cues acquire control over motivated behavior (Meyer et al. 2012b; Robinson et al. 2014b; Saunders and Robinson 2013). Thus, it appears that the cue is an equally effective *CS* in STs and GTs, but it acts as a more powerful, motivating *incentive stimulus* in STs.

Both food and drugs cues engage so-called brain 'motive circuits' to a greater extent in STs and GTs (Flagel et al. 2011a, 2011b; Yager et al. 2015), which may contribute to STs being more susceptible to develop impulse control disorders, such as addiction (Flagel et al. 2011b; Saunders and Robinson 2013; Tunstall and Kearns 2015; Beckmann et al. 2011). Addiction-like behaviors tend to persist despite non-reinforcement (Deroche-Gamonet et al. 2004), therefore, one might expect sign-tracking behavior to be more resistant to extinction than goal-tracking. However, in previous studies involving the extinction of *instrumental responding* for food (Yager and Robinson 2010) or cocaine (Saunders et al. 2014; Saunders and Robinson 2011; Yager and Robinson 2013), there was no difference between STs and GTs in the rate of instrumental extinction. However, in all of these cases extinction was conducted in the absence of a reward cue (CS), or an explicit reward cue was not used in training. In another study presentation of an increasingly aversive outcome progressively decreased instrumental responding for cocaine at the same rate in STs and GTs, despite continued presence of a cue associated with cocaine delivery (Saunders et al. 2013). On the other hand, in one study in which extinction of cocaine self-administration behavior was conducted in the presence of the reward cue, STs were somewhat more resistant to extinction than GTs, but the effect was very small and was only evident on the first day (out of 28 days) of extinction training (Saunders and Robinson 2010). Thus, there is no compelling evidence that STs and GTs differ in extinction learning in an instrumental task, but we are not aware of any study comparing Pavlovian extinction in STs and GTs.

It is important to examine Pavlovian extinction because, although there are many similarities in the processes involved in instrumental and Pavlovian extinction, there are also differences (e.g., Todd et al. 2014). Beckmann and Chow (2015) did examine Pavlovian

extinction using a 2-CS procedure, in which rats developed a ST CR to a lever-CS and a GT CR to a tone-CS, when both CSs were presented within the same session. They reported that when the lever and tone were no longer reinforced the GT CR extinguished faster than the ST CR (Beckmann and Chow 2015). However, in this study all rats showed a ST CR during lever-CS presentation; i.e., no rats showed a GT CR in response to the lever-CS. Therefore, it is still not known whether STs and GTs differ in Pavlovian extinction, when the same CS evokes two different CRs in different individuals.

Thus, we conducted a series of studies on extinction in STs and GTs. In the first study, after initial training, the US was simply withheld over 4-8 days of extinction training. Next, we examined within-session extinction and recovery of ST and GT CRs during alternating blocks of rewarded (R) and non-rewarded (NR) trials. The R and NR periods were signaled by either a diffuse and continuous change in illumination (experiment 2) or by spatially and temporally discrete light cues (experiment 3). These signals were intended to facilitate discrimination between R and NR phases by acting as occasion setters, indicating that the CS would be rewarded when accompanied by one signal but would not be rewarded when accompanied by the other (Bouton 2004; Bouton et al. 2014; Crombag et al. 2008; Holland 1992). Finally, in a fourth experiment changes in illumination were used as discriminative stimuli (Bouton et al. 2014; Cardinal et al. 2002) signaling whether an instrumental response would be reinforced or not. STs were highly resistant to changes in reward omission, relative to GTs, but only in the Pavlovian setting, when the CS was present.

2. METHODS

2.1. Subjects

A total of 133 male Sprague-Dawley rats purchased from Charles River or Harlan were used (Fitzpatrick et al. 2013). All rats weighed 250-275g on arrival, and were singly housed in standard polycarbonate cages with food and water available *ad libitum* throughout experiments (i.e., the rats were *not food deprived* at any time). Rats were housed on a reverse 12-h light/12-hr dark cycle, with all testing taking place during the dark phase of the cycle. Testing began after one week of acclimation and handling. All procedures were approved by the University of Michigan Committee on the Use and Care of Animals (UCUCA).

2.2. Apparatus

All experiments took place in standard test chambers (30.5 x 24.1 x 21 cm) located inside sound-attenuating cabinets with ventilating fans, which masked outside noise (Med Associates, Inc.; St Albans, VT, USA). Each chamber contained a pellet dispenser connected to a recessed pellet magazine, which was mounted in the center of the front wall and equipped with an infrared sensor to record magazine head entries. During Pavlovian conditioned approach (PCA) training a single retractable lever was mounted to the right or left of the magazine (the position of the lever was counterbalanced across rats). An LED behind the lever was lit any time the lever was inserted into the cage and illuminated the slot through which the lever protruded. During instrumental sessions, the lever was removed and active and inactive nose-poke ports were installed on either side of the magazine, with the left or right position of the active port counterbalanced across rats. Chambers also contained a red house light mounted at the top of the back wall, and two white stimulus lights, one mounted on the back wall of the chamber, and one mounted on the front wall, with the left or right position relative to the magazine

counterbalanced. Behavioral responses (lever deflections, magazine entries, and nose pokes) were recorded with MED-PC® software (Med Associates, Inc.).

2.3. Pavlovian conditioned approach (PCA) training

All experiments began with PCA training, during which the insertion of the lever (the conditional stimulus; CS) into the cage was paired with the delivery of a single food pellet (unconditional stimulus, US; a 45 mg banana-flavored pellet; #F0059; Bio-Serv; Frenchtown, NJ, USA), using methods described previously (Flagel et al. 2007). For two days prior to the start of training, rats were given 25 pellets in their home cages in order to familiarize them with this food. On a single pre-training day, 25 pellets were delivered on a variable time (VT 30-s) schedule without a CS, in order to train rats to retrieve pellets from the magazine. One rat failed to eat the banana pellets, and was excluded from the study. PCA sessions began the following day. There were 25 trials per session separated by VT-90 (30-150-s) intervals. In each trial, the lever was inserted into the cage for 8 s, then retracted and immediately followed by the response-independent delivery of a banana pellet. The chamber was illuminated by the red house light during PCA training. Lever presses and magazine entries were recorded during the 8-s CS period when the lever was present.

2.4. Quantification of PCA behavior

As in previous studies the lever-CS evoked different CRs in different rats. The relative occurrence of a sign-tracking CR (ST; approach and engagement with the lever) vs. a goal-tracking CR (GT; approach and engagement with the food cup) was quantified by measuring: 1) the number of contacts with the lever or magazine, 2) the latency to first contact the lever or magazine, and 3) the probability of contacting the lever or magazine on any given trial. To classify rats as STs or GTs, we used a composite of these three measures termed the PCA Index

Score (Meyer et al. 2012a). For each of the five training sessions we determined a PCA Score, which was the average of three measures: response bias (lever presses - magazine entries / total contacts), latency score (magazine latency - lever latency / 8 s), and probability difference (lever probability - magazine probability). PCA Scores range from 1 (indicating an animal exclusively made ST CRs in every trial) to -1 (indicating exclusively GT CRs on every trial). PCA Scores from the last two training sessions were averaged to produce a PCA Index Score that was used to classify rats as STs (+0.5 to 1) or GTs (-0.5 to -1). Rats that fell in the middle range (-0.5 to 0.5) were classified as intermediates and were not included in these studies.

2.5. Experiment 1 - Extinction of ST vs. GT CRs

Rats were initially trained for 7-9 days using the PCA procedures described in section 2.3. The day after the last day of training all rats were then subjected to extinction conditions, in which the lever-CS was presented as before (25 trials per session, VT-90 s interval) but was not followed by a food pellet. There was no signal provided to indicate extinction conditions were in effect, other than reward was not provided. One group of 15 rats had four days of extinction training, and another group of 34 rats had eight days of extinction training. In both groups data from the first 4 days were very similar, so the two groups were pooled in the analyses below.

2.6. Experiment 2 – Alternating rewarded (R) and non-rewarded (NR) periods signaled by illumination of the chamber

In experiment 1 there was no explicit signal indicating whether the CS would be followed by reward. In this, and in the following experiment, we asked whether STs would show similar resistance to extinction as seen in experiment 1 if alternating blocks of reward and non-reward periods were explicitly signaled. In this experiment a 'diffuse' signal (illumination of the entire chamber) indicated whether presentation of the CS would or would not be followed by reward.

All rats (N=21) were first trained using the PCA procedure described above for 5 daily sessions. On each of the following 5 days sessions consisted of 60 trials (i.e., 8-s CS presentations) separated by VT-90s intervals. However, the sessions were divided into two blocks in which the CS was followed by presentation of the food reward ('Rewarded'; R) and two blocks in which presentation of the CS was not followed by reward ('Non-Rewarded'; NR). The two R and NR periods lasted for 15 trials each. Thus, during R blocks, the red house light was constantly illuminated (as during PCA training phase), and thus the chamber would appear dimly lit, and a banana pellet followed each presentation of the lever-CS. NR blocks were signaled by the continuous illumination of a white stimulus light, which illuminated the entire chamber, and under this condition the 8-s lever-CS presentations were not followed by a food pellet. The four periods alternated in two different patterns (R, NR, R, NR or NR, R, NR, R), which were counterbalanced across rats and across days.

2.7. Experiment 3 - Alternating rewarded (R) and non-rewarded (NR) periods signaled by a discrete signal

 To extend experiment 2 we repeated the experiment, but further asked whether similar results would be obtained when R and NR periods were signaled by a spatially and temporally discrete cue, rather than illumination of the entire chamber. Illumination of the entire chamber may be perceived as a change in 'context', which could be processed differently than a discrete cue (e.g., (Bouton 2004; Burns and Domjan 2001; Domjan 2003; Holland and Bouton 1999). A separate group of 39 rats was used in this experiment. All rats first received 5 days of PCA training, as above, followed by 5 days of discrimination training, as in experiment 2. However, in this experiment R periods were signaled by illumination of a white stimulus light mounted at the front of the cage, and NR periods by illumination of an identical white stimulus light at the back

of the cage. Unlike in experiment 2, these lights were illuminated for only 6 s immediately prior

to each lever-CS presentation, and the red house light stayed on throughout the session. When

the lever-CS was preceded by the R signal, the trial ended with delivery of a pellet. When the

lever-CS was preceded by the NR signal, no pellet was delivered. During discrimination training

there were 60 trials per session separated by VT-90-s intervals. There were two R and two NR

blocks that consisted of 15 consecutive trials each, alternating in two counterbalanced patterns (R,

NR, R, NR or NR, R, NR, R).

2.8. Experiment 4 – Effects of alternating rewarded (R) and non-rewarded (NR) periods signaled by illumination of the chamber on instrumental responding for food reward

In our previous studies on instrumental extinction, in which STs and GTs did not differ,

the extinction period was not indicated by any explicit signal. The purpose of this experiment

was to again examine instrumental extinction in STs and GTs but using a discrimination

procedure that better matched the procedures used in experiments 2 and 3 above.

2.7.1. Acquisition: Following PCA training, a separate group of 24 rats were trained to nose

poke in the active port to receive a food pellet (responses in the inactive port had no

consequences). In sessions 1-3 rats were trained on a fixed ratio (FR1) schedule with a 10-s time-

out period after each pellet delivery. For the next three sessions rats were trained on variable

interval (VI) schedules that increased the time between rewarded responses: day 4 VI-20 (10-30

s), day 5 VI-40 (20-60 s), day 6 VI-60 (30-90 s). These first six sessions lasted until rats self-

administered 30 pellets, to ensure that all animals experienced the same number of reinforced

responses. On days 7 and 8, rats were trained on the VI-60 schedule, but sessions had a fixed

duration of 30 min to facilitate comparison with discrimination sessions. A red house light was

constantly illuminated throughout all acquisition sessions.

2.7.2. Discrimination: For the next 10 days rats were trained on discrimination sessions that consisted of four alternating 15-min blocks, much like in experiment 2. During the R periods, the red house light was constantly illuminated and responses in the active port were rewarded on a VI-60 schedule of reinforcement. During NR periods, the red light was extinguished and a white stimulus light mounted above the inactive port (the same as used in experiment 2) was constantly illuminated, and responses in the active port were never reinforced. The order of R and NR phases (R, NR, R, NR or NR, R, NR, R) was counterbalanced across rats and across days.

2.9. Statistics

Between-group comparisons were analyzed with repeated-measures (RM) ANOVAs comparing ST and GT groups across sessions. Significant interaction results were followed by post hoc independent *t* tests (Bonferroni corrected) comparing STs and GTs in each of the sessions included in the ANOVA. In experiments 2 and 3, within-group comparisons were analyzed with RM ANOVAs comparing R and NR blocks across the five days of discrimination training, with both session and block as within-subjects variables. Significant interaction results were followed by post hoc paired *t* tests (Bonferroni corrected) comparing R and NR results across all sessions included in the ANOVA. Corrected *p* values are reported if data did not pass Levene's test of equality of variance or Mauchly's test of sphericity. All statistical procedures were performed with SPSS for Windows (Version 21).

3. RESULTS

3.1. PCA training: ST and GT groups

As described previously, individual rats differed in the form of conditioned approach behavior they developed during the PCA training phase, with some rats primarily approaching the lever and others the food cup during the 8-s CS periods (Flagel et al. 2009; Meyer et al. 2012a; Robinson and Flagel 2009). Across the four experiments, 72 rats were classified as STs, and 61 were classified as GTs. Additional rats were classified as intermediates, but these rats were not used further because we wanted to directly compare groups that varied markedly in their propensity to attribute incentive salience to the lever-CS (Meyer et al. 2012a). Measures of probability, contacts, and latency are shown for lever-directed behavior and food cup-directed behavior during the first five days of PCA training for all rats used in subsequent experiments (**Fig. 1**). For all measures the behavior of STs and GTs diverged with repeated days of training, and for each of the six comparisons there were significant interaction effects [Fs(4, 328) = 48.51 - 75.47, ps < .001].

During discrimination training in experiments 2 and 3, the majority of animals maintained the ST or GT classification that was established during PCA training. However, in some rats discrimination training disrupted their initial preference for the lever or food cup, and these rats switched from sign-tracking to goal-tracking or from goal-tracking to sign-tracking by the end of the discrimination training phase (**Fig. 2**). Only data from rats that maintained PCA scores in the direction of their original classification (STs > 0 and GTs < 0) are included in the analyses below.

3.2. Experiment 1 - Extinction of ST vs. GT CRs

In this experiment, non-reinforced exposure to the lever-CS (without any signal) resulted in faster extinction of PCA behavior in GTs ($n = 21$) than STs ($n = 28$). Within-session changes in probability of approach are shown for the last day of PCA training (baseline), and for

extinction days 1 and 4. GTs showed a rapid decrease in approach toward the food cup within the first extinction session, as well as a further decrease by the fourth extinction session (**Fig. 3A**). In STs, approach toward the lever did not change from baseline within the first extinction session; however, by the fourth extinction session the STs did show a significant decrease in sign-tracking behavior (**Fig. 3B**).

Fig. 3C shows probability of approach averaged across all 25 trials per session, analyzed with a 2 x 5 RM ANOVA (including baseline plus extinction sessions 1-4). Responding during extinction was significantly lower in GTs compared to STs [group $F(1, 47) = 14.57, p < .001$]. Repeated days of extinction training decreased approach in both groups [session $F(4, 188) = 171.79, p < .001$]; however, the rate of decrease differed in GTs and STs, with a significant group x session interaction effect [$F(4, 188) = 13.97, p < .001$]. Paired t tests comparing extinction sessions to baseline (Bonferroni corrected) showed that changes from baseline were significant for GTs in all four sessions, but were only significant for STs in the last three sessions (**Fig. 3C**).

In **Fig. 3D** the first 3 trials of each session were examined separately to evaluate the recall of extinction from one session to the next. GTs showed more immediate recall of extinction than STs, with a significant main effect of group [$F(1, 47) = 32.97, p < .001$]. Though recall improved in both groups with repeated days of extinction training [session $F(4, 188) = 32.32, p < .001$], this improvement was better in GTs than STs, with a significant group x session interaction effect [$F(4, 188) = 7.25, p < .001$]. Paired t tests comparing extinction sessions to baseline (Bonferroni corrected), indicated that GTs showed significant recall of extinction in sessions 2-4, whereas STs showed significant, but weaker, recall in sessions 3 and 4 (**Fig. 3D**). We also examined the first 3 trials of each session in the subset of rats that received 8 days of

extinction training instead of just 4 (9 GTs and 25 STs), and found that significant differences persisted throughout the additional sessions [session $F(8, 256) = 24.28$, $p < .001$; group $F(1, 32) = 30.17$, $p < .001$; interaction $F(8, 256) = 2.77$, $p < .01$; **Fig. 3E**].

3.3. Experiment 2 - Alternating rewarded and non-rewarded periods signaled by illumination of the chamber

Fig. 4 shows the effect of alternating rewarded (R) and non-rewarded (NR) periods, signaled by chamber illumination, on the probability of sign-tracking and goal-tracking behavior, in STs and GTs, on the first and fifth day of R/NR training. It is clear that even in the very first session the behavior of GTs ($n = 11$) changed markedly between R and NR periods, but this was not the case for STs ($n = 7$). GTs rapidly decreased responding during NR phases, and rapidly resumed making GT CRs during R phases (**Fig. 4A-B**). In contrast, in the first session the behaviors of STs did not differ between the R and NR periods (**Fig. 4C-D**). By the fifth session GTs continued to show excellent discrimination between the R and NR periods, and by this time STs did so as well, based on probability of making a ST CR (see below). In addition, **Fig. 4** shows that the introduction of discrimination training had essentially no effect on the very low levels of food cup directed behavior in STs or on lever directed behavior in GTs.

For lever-directed behavior in STs and food cup-directed behavior in GTs, measures of probability, contacts, and latency were averaged across the 30 R trials and 30 NR trials per session, and analyzed for each group separately with 2 x 5 (block x session) RM ANOVAs. STs initially showed no discrimination between R and NR periods, but with repeated training sign-tracking responses began to diverge between the two blocks. For all three measures, there were significant effects of block [$Fs(1, 6) = 7.78 - 30.24$; $ps < .05 - .01$] and significant interaction effects [$Fs(4, 24) = 4.46 - 9.21$; $ps < .01 - .001$; **Fig. 5A-C**]. In STs, post hoc tests comparing R

and NR blocks were significant in session 5 for measures of probability and latency, but not contacts (**Fig. 5A-C**). In addition, during the discrimination sessions sign-tracking behavior was somewhat disrupted during R periods relative to the last day of PCA training. When lever-directed behavior during PCA training session 5 was compared to the reinforced blocks of R/NR session 5 (paired t tests), significant reductions were seen for probability [$t(6) = 2.60, p < .05$] and contacts [$t(6) = 2.51, p < .05$], but not for latency [$t(6) = 1.55$, n.s.] (**Fig. 5A-C**).

In contrast to STs, GTs showed excellent discrimination between R and NR blocks in all five sessions. For probability, latency, and contacts there were significant differences between R and NR blocks [$Fs(1, 10) = 29.93 - 298.21, ps < .001$]. The GTs also showed some improvement in discrimination across the five discrimination sessions, with significant interaction effects for all three measures [$Fs(4, 40) = 3.51 - 4.50, ps < .05 - .01$]. Post hoc tests showed that the differences between R and NR periods were significant in all of the five sessions (**Fig. 5D-F**). In addition, and unlike sign-tracking (above), goal-tracking behavior was maintained in the R periods at the same level as during the last day of PCA training, with no significant differences when PCA training session 5 was compared to R/NR session 5 [$ts(10) = 1.07 - 1.53$, n.s.] (**Fig. 5D-F**).

To more directly compare the performance of STs and GTs, we calculated difference scores for each measure (R responses minus NR responses), and performed 2 x 5 (group x session) RM ANOVAs. For probability, the difference scores of GTs were significantly higher than STs [group $F(1, 16) = 11.01, p < .01$]. Although the scores of both groups increased across sessions [session $F(4, 64) = 15.07, p < .001$], the STs improved to the point where they matched the GTs on the last day, with a significant interaction effect [interaction $F(4, 64) = 3.63, p < .01$]. Post hoc tests showed significant differences between STs and GTs only in the first two sessions

(**Fig. 5G**). For contacts, difference scores were significantly higher in GTs than STs, and for both groups improved across sessions at similar rates, with significant main effects but no interaction [group $F(1, 16) = 9.32$, $p < .01$; session $F(4, 64) = 7.01$, $p < .001$; interaction $F(4, 64) = 0.76$, n.s.] (**Fig. 5H**). For latency, there was a significant difference between STs and GTs and an improvement in discrimination across sessions, but no interaction effect [group $F(1, 16) = 33.74$, $p < .001$; session $F(4, 64) = 10.04$, $p < .001$; interaction $F(4, 64) = 2.27$, n.s.] (**Fig. 5I**). Thus, for these latter two measures STs never achieved the same level of discrimination between R and NR periods, even after 5 days of discrimination training.

3.4. Experiment 3 - Alternating rewarded and non-rewarded periods signaled by a discrete signal

When discrete cues were used to signal whether or not the lever-CS would be followed by reward the results looked very much as in experiment 2 above. GTs ($n = 18$) showed much better discrimination between the R and NR periods than STs ($n = 15$). **Fig. 6** shows changes in lever- and food cup-directed behavior during blocks of R and NR trials. In the very first session, as well as the fifth session, GTs rapidly decreased responding during NR periods, and rapidly resumed making GT CRs during R periods (**Fig. 6A-B**). In contrast, the STs showed no difference in responding between R and NR periods in the first session, though discrimination was evident by the fifth session (**Fig. 6C-D**).

In **Fig. 7**, measures of lever-directed behavior for STs, and food-cup directed behavior for GTs, were averaged across the 30 R trials and 30 NR trials per session, and analyzed with 2 x 5 (phase x session) RM ANOVAs. Similar to experiment 2, STs showed no difference between R and NR periods in the first session, but they showed improvement in discrimination in subsequent sessions. For all three measures (probability, contacts, and latency) there were

significant effects of block [Fs(1, 14) = 9.21 - 23.59; ps < .01 - .001] and significant block x

session interactions [Fs(4, 56) = 5.12 - 5.43; ps < .01 - .001]. Post hoc tests showed significant

differences between R and NR periods emerging in the third session (**Fig. 7A-C**). Although in

later trials STs showed more sign tracking during R trials compared to NR trials, sign-tracking

behavior during the R trials was reduced compared to baseline levels on PCA day 5. For all three

measures, lever-directed behavior was diminished during R trials in session 5 compared to PCA

5 [paired t tests: probability, t(14) = 3.44, p <.01; contacts, t(14) = 4.05, p < .01; latency, t(14) =

2.55, p < .05; **Fig. 7A-C**].

   Fig. 7D-F show probability, contacts, and latency for food-cup directed behavior in GTs.

The GTs showed significant discrimination in all five sessions, and for all three measures there

were significant effects of block [Fs(1, 17) = 95.25 - 467.47, ps < .001]. There were also

significant session effects for latency and probability [Fs(4, 68) = 3.74 - 4.88, ps < .01], but not

for contacts [F(4, 68) = .79, n.s.]; and significant interaction effects for latency and contacts

[Fs(4, 68) = 2.79 - 2.81; ps < .05], but not for probability [F(4, 68) = 1.64, n.s.]. Post hoc tests

showed significant differences between R and NR periods in all five sessions (all ps < .001).

When the last day of PCA training was compared to R/NR session 5 (paired t tests), there were

no significant differences in probability or contacts [ts(17) = .32 - 1.80, n.s.], though there was a

significant decrease in latency to approach the food cup [t(17) = 2.33, p < .05], indicating some

improvement in performance in GTs, unlike STs.

   Difference scores comparing STs and GTs are shown in **Fig. 7G-I**. Similar results were

found for measures of probability, contacts, and latency, with GTs showing greater

discrimination between R and NR periods than STs [group: Fs(1, 31) = 18.06 - 60.65, ps < .001].

Both groups showed improvement in difference scores with repeated days of training, with

significant session effects [session: Fs$(4, 124) = 6.85 - 8.00$, ps $< .001$]; however, discrimination in the STs never reached the level of the GTs, as group x session interactions were not significant for any of the three measures [interaction: Fs$(4, 124) = 1.58 - 1.88$, n.s.].

3.5. Experiments 2 and 3 - Discrimination within the first trial between R and NR blocks.

In experiments 2 and 3 light cues signaled whether the CS would be reinforced; however, the analyses above cannot confirm whether these signals themselves actually influenced the probability of conditioned responding. To try and address this question, we conducted a separate analysis using only the very first trial of each block, because in this first trial the signals were the only indication (other than the number of trials) that an NR block had switched to an R block, or vice versa. Thus, if behavior changed on the first trial it suggests (although does not prove) that the signals may have acted as occasion-setters. Lever and magazine contacts during the first trial of R/NR periods was compared to the average of the last three trials of the preceding block, and difference scores were calculated (R minus NR or NR minus R) to represent the change from the end of one block to the beginning of the next. For each group these difference scores were analyzed with 2 x 5 ANOVAs with both session and block as repeated measures factors. In experiment 2, both STs and GTs showed significant interaction effects [STs, $F(4, 24) = 4.02$, $p < .05$; GTs, $F(4, 40) = 3.496$, $p < .05$], though Bonferroni-corrected post hoc were only significant in session 4 for GTs ($p < .01$) (**Fig 8A-B**), suggesting that the diffuse light cues had a significant influence on behavior in both groups. In experiment 3, STs showed a significant block x session interaction [$F(4, 56) = 3.457$, $p < .05$] (**Fig. 8C**), but GTs did not [$F(4, 68) = 0.249$, n.s.] (**Fig. 8D**).

3.6. Experiment 4 - Effects of alternating rewarded and non-rewarded periods signaled by illumination of the chamber on instrumental responding for food reward

Fig. 9A shows that STs ($n = 14$) and GTs ($n = 10$) acquired stable instrumental responding for food reward at similar rates. The rate of active nose-poke responses increased across the 8 acquisition sessions, with no group difference between STs and GTs, and no interaction effect [2 x 8 RM ANOVA: session $F(7, 147) = 16.804$, $p < .001$; group $F(1, 21) = 2.82$, n.s.; interaction $F(7, 147) = .752$, n.s.]. When R and NR periods were introduced, both ST and GT groups showed rapid discrimination between these two periods. Separate two-way ANOVAs were performed for STs and GTs, with both block and session included as repeated-measures factors. For both groups there were significant block x session interaction effects [STs: $F(9, 117) = 15.09$, $p < .001$; GTs: $F(9, 81) = 11.19$, $p < .001$], and in both groups post hoc comparisons of R and NR responses were significant in each of the 10 sessions (all ps $< .001$; **Fig. 9B**). Although both STs and GTs discriminated between R and NR periods equally well, the overall rate of responding during the R periods was higher for GTs than STs, and 2 x 10 RM ANOVAs comparing STs and GTs showed significant group effects during R blocks [$F(1, 22) = 5.23$, $p < .05$], but not during NR blocks [$F(1, 22) = 1.93$, n.s.] (**Fig. 9B**). For this reason, differences scores (R responses minus NR responses) were also significantly higher in GTs compared to STs [2 x 10 RM ANOVA: group $F(1, 22) = 5.53$, $p < .05$] (**Fig. 9C**). For inactive nose-poke responses, there were no significant block x session interaction effects [STs: $F(9, 117) = 1.19$, n.s.; GTs: $F(9, 81) = 2.3$, n.s], and there were no significant group differences during R periods [$F(1, 22) = .09$, n.s.] or NR periods [$F(1, 22) = .14$, n.s.].

4. DISCUSSION

In appetitive Pavlovian conditioning, when a CS is associated with delivery of a reward, animals may acquire many different anticipatory CRs depending on the form of both the US and

the CS (Delamater 2012b; Holland 1977). However, as seen here, even when subjected to exactly the same training procedures, using the same CS and US, different individuals may acquire quite different CRs (also see Boakes 1977; Flagel et al. 2009; Saunders et al. 2013; Zener 1937). Thus, presentation of a lever-CS, paired with delivery of a food reward, evoked a ST CR (approach and engagement with the lever-CS itself) in some animals and a GT CR (approach and engagement with the food cup) in others. This presumably reflects differences in the learning mechanisms engaged in different animals, a notion that is supported by evidence that different neurobiological systems mediate sign-tracking and goal-tracking CRs (Clark et al. 2013; Chang et al. 2012; Danna and Elmer 2010; Danna et al. 2013; Flagel et al. 2011a, 2011b; Saunders and Robinson 2012; Yager et al. 2015). Here we asked whether such presumed differences in what is learned in STs and GTs also influences the ability of periods of non-reward to decrease conditioned responding; i.e., produce extinction. It did, as there were large differences between STs and GTs in Pavlovian, but not instrumental, extinction.

In the first experiment, after initial acquisition of ST and GT CRs, STs and GTs were exposed to extinction conditions. In this experiment there was no signal indicating extinction conditions were in effect, other than that the CS was no longer followed by the US. GTs showed rapid extinction of their GT CR, even on the first day of extinction training. In contrast, STs showed no evidence of extinction of their ST CR on the first day of training, and showed continued resistance to extinction (relative to GTs) over a further 4-8 days of extinction training. This was especially evident during the first few trials of each day of extinction training, when the retention of extinction memory from one session to the next is often most clearly expressed (Peters and De Vries 2013; Rescorla 2004). During these trials STs continued to respond at

levels close to those seen prior to extinction, whereas GTs greatly decreased responding over days of extinction training (Fig. 3D).

We next examined how established sign- and goal-tracking CRs were affected when alternating periods of reward (CS followed by US) were interspersed with periods of non-reward (CS not followed by US). We found that GTs rapidly altered their behavior between R and NR periods, showing a rapid decrease in goal-tracking behavior during the NR periods and rapid recovery of goal-tracking during the R phases. GTs discriminated between the R and NR periods even on the first day of discrimination training, and improved further over 5 days. On the other hand, STs did not discriminate between R and NR periods on the first day of discrimination training, and although over 5 days of training they significantly improved, on some measures they never discriminated R from NR periods as well as GTs. It is important to note that this situation is quite different from that involving *partial reinforcement*, which has been shown to facilitate sign-tracking (Anselme 2014; Anselme et al. 2013; Boakes 1977; M. Robinson et al. 2014a), possibly because, as suggested by Anselme (2014), uncertainty potentiates incentive salience attribution.

The R and NR blocks were signaled by either diffuse, continuous illumination of the entire test chamber (experiment 2) or by spatially and temporally discrete light stimuli (experiment 3). These signals were intended to function as occasion setters, indicating whether or not the lever-CS would be rewarded (Bouton 2004; Bouton et al. 2014; Holland 1992; Holland and Bouton 1999). In both cases patterns of discrimination between R and NR conditions were very similar, suggesting that the diffuse and discrete cues may have had similar influences on behavior. However, given that the R and NR periods consisted of blocks of 15 trials, it is difficult to confirm whether these signals really influenced behavior because of their action as occasion

setters. With this block design one trial predicted reinforcement on the next trial with high

probability (14/15), and this may have actually controlled behavior. Only on the very first trial of

each block could we directly assess the impact of the signals, as reinforcement during this first

trial was not predicted by the previous trial. In experiment 2, by the end of training, both STs and

GTs discriminated between R and NR periods even on the very first trial between blocks,

suggesting that chamber illumination did influence conditioned responding. However, we cannot

exclude the possibility that the rats kept a running tally of the number of trials. In experiment 3

the discrete light signals had less of an impact on behavior, and only in STs. It is not clear why

diffuse illumination and discrete lights had different effects, but one possibility is that the diffuse

cues were easier to distinguish from one another, as they were likely perceived as a dimly lit vs.

brightly lit chamber; whereas the discrete cues were identical white lights on opposite sides of

the chamber. Another is that the diffuse lights may have acted as contextual features because

they were constantly illuminated throughout the two blocks (Burns and Domjan 2001; Domjan

2003), and previous studies have shown that while STs are more susceptible to cue-induced

behavior than GTs, GTs show more context-induced conditioned behavior than STs (Morrow et

al. 2011; Saunders et al. 2014). In conclusion, although these signals may have influenced

conditioned responding differently in STs and GTs because they are differentially sensitive to

occasion setters, further studies will be required to determine if this is in fact the case.

Consistent with our previous studies of instrumental extinction in STs and GTs following

omission of either a food (Yager and Robinson 2010) or drug reward (Saunders et al. 2014;

Saunders and Robinson 2011; Yager and Robinson 2010, 2013; but see Saunders and Robinson

2010), there were no marked differences in the ability of STs and GTs to discriminate R from

NR periods when they were tested in an instrumental setting (food reinforced responding). GTs

did show a greater number of active responses during R periods (but not NR periods), but STs and GTs acquired the discrimination at the same rate.

There are many similarities between Pavlovian and instrumental extinction, but there are also differences (e.g., Todd et al. 2014). However, the mechanisms responsible for the differences between Pavlovian vs. instrumental extinction reported here, or for why STs and GTs differ in resistance to Pavlovian extinction, are not known. This is in part because the mechanisms responsible for extinction are very complex, are not well understood and there are a number of alternative and not mutually exclusive mechanisms that may contribute to the decrement in responding in extinction (Bouton 2002; Delamater 2012a; Todd et al. 2014). For example, extinction could result in a negative prediction error slowly degrading the CS-US (or action-outcome) association. However, STs and GTs do not differ in the rate of learning either Pavlovian (experiments 1-4; Flagel et al. 2009, Meyer et al. 2012a) or instrumental responses (experiment 4; Yager and Robinson 2010), so it is not clear why they would differ in extinction learning based on prediction errors. Furthermore, the idea that extinction 'erases' the original learning has been challenged by phenomena such as spontaneous recovery, reinstatement and context renewal (Bouton 2002). It has been suggested, therefore, that in extinction the original CS-US association in maintained, but a new inhibitory CS-no US association is learned, which comes to control responding. It is possible that STs do not learn this inhibitory association as well as GTs, although again, it is not clear why this would only be true for Pavlovian extinction learning and not instrumental extinction learning, although the target of extinction does differ in these two cases (a stimulus vs. a response, respectively).

It is also possible attentional factors could contribute; perhaps STs are less able to shift their attention away from the CS, relative to GTs, thus slowing learning about the fact that

conditions have changed. Unexpected reward omission is thought to cause a state of surprise that directs an animal's attention to other potential cues in the environment, and this is thought to explain why, for example, the extinction of a Pavlovian CR is much more context-specific than learning the original CS-US association (Bouton 2004; Rosas et al. 2013). Potential individual differences in attentional control are supported by a study showing that STs exhibit poorer performance on a test of "top-down" attentional control than GTs, and therefore reward omission may engage cognitive control mechanisms to a greater extent in GTs than STs (Paolone et al. 2013). That is, GTs may be able to shift their focus of attention faster than STs after the unexpected omission of reward. However, if the present results were influenced by poor attentional control and delayed learning in STs relative to GTs, it appears this is only the case if the lever-CS is present during a task, as STs do not show delayed instrumental extinction (experiment 4 and Yager and Robinson 2010), nor do they show delayed Pavlovian extinction with a tone-CS (Beckmann and Chow 2015).

This raises another possibility – STs attribute much greater incentive salience to a lever-CS than GTs and this may interfere with the extinction of a sign-tracking CR, either by causing a narrowing of attention that prevents learning about changes in the contingencies, or alternatively, by promoting continued sign-tracking despite understanding that the lever will not be rewarded. The idea that the lever-CS is simply attributed with so much more incentive salience in STs then GTs that it continues to act as a 'motivational magnet', even in the absence of reward (Mahler and Berridge 2009), is supported by studies of conditioned reinforcement showing that for STs, but not GTs, the CS becomes intrinsically rewarding and can support instrumental actions in the absence of the US (Lomanowska et al. 2011; Robinson and Flagel 2009). In addition, for STs, the lever-CS is more effective in evoking a motivational state that reinstates reward-seeking

behavior (Saunders and Robinson 2013), and is salient enough to block learning about an auditory cue that reliably predicts reward (Holland et al. 2014). Therefore, because the CS has greater incentive stimulus properties in STs, this alone may be sufficient to continue to attract STs toward the CS, even under extinction conditions.

It is further possible that impulsivity may have contributed to the behavior of STs, as several studies have reported a link between the tendency to attribute incentive salience to a cue and a tendency toward impulsive behavior (Flagel et al. 2010; Flagel et al. 2014; Olshavsky et al. 2014; Tomie et al. 2008). In particular, tests of impulsive action have shown that STs have a reduced ability to inhibit behavior and withhold inappropriate responses compared to GTs (Lovic et al. 2011), meaning that even if STs had learned that no reward was forthcoming during the NR periods they may have been less able to withhold responding.

It is notable that introduction of discrimination training disrupted performance of STs, but not GTs, during the R periods. Even after STs learned to discriminate between R and NR blocks, sign-tracking behavior during the R periods was reduced compared to the last day of PCA training. Changes in the strength of incentive motivation can be reflected in the vigor of the interaction (i.e., number of contacts) with the lever or food cup (DiFeliceantonio and Berridge 2012; Saunders and Robinson 2012), which may suggest that for STs when the CS eventually was recognized as an ambiguous predictor of reward it lost some (but not all) of its incentive value. For GTs, goal-tracking was not reduced during R phases, and for some measures even improved somewhat. It should also be noted that experiments 2 and 3 included only the rats that maintained their initial preference for the lever or food cup throughout the discrimination sessions, as there was a subpopulation of STs that switched to goal-tracking (as well as a single GT that switched to sign-tracking). Data from rats that switched were too variable, and their

numbers were too few, for their behavior to be examined in further detail. However, this suggests that at least in a subset of STs the CS may lose all incentive value once it ceases to be a reliable predictor of reward.

It is clear that the Pavlovian CS engages different response (behavior) systems in STs and GTs (e.g., Timberlake 1994), and thus they presumably differ in the content of learning. Indeed, as argued by Delamater (Delamater 2012a, p. 17) associations may be formed, "between the CS and multiple components of the US, e.g., its sensory, emotional, and specific response components (e.g., Delamater and Oakeshott 2007; Konorski 1967; Wagner 1989)." "These different aspects of learning might obey different learning rules" and if so " it would not at all be surprising if they would each show different sensitivities to extinction treatments." By this way of thinking the differences between STs and GTs (and between Pavlovian and instrumental settings) may be because in STs and GTs behavior is governed by different associations and/or behavioral systems, and different rules apply. However, exactly what these rules are, and how they differ in STs and GTs is not known (although see Meyer et al. 2014).

In conclusion, there are many unanswered questions regarding mechanism, but it seems clear that it is much more difficult to extinguish a ST CR than a GT CR, further supporting previous suggestions that they are produced by two very different psychological (and neurobiological) learning processes (Chang et al. 2012; Clark et al. 2012; Danna et al. 2013; Flagel et al. 2011a; Meyer et al. 2014; Robinson et al. 2014b; Saunders et al. 2013). This may have implications for the treatment of disorders characterized by hypersensitivity to reward cues, such as obesity and addiction (Berridge 2009; Robinson and Berridge 2008), and perhaps traumatic cues as well (Morrow et al. 2011; Morrow et al. 2014). The ability of cues to trigger excessive motivation for food or drugs is well established (Everitt and Robbins 2005; Robinson

and Berridge 2008; Stewart et al. 1984). However, the success of treatments aimed at mitigating the effect of cues, such as cue-exposure therapy, has been equivocal, with many patients experiencing relapse after treatment (Conklin and Tiffany 2002; Troisi 2013). The results presented here suggest that treatments to extinguish motivation to overeat or to pursue drugs may be less effective in individuals prone to attribute excessive incentive salience to reward cues than in individuals less prone to do so. A better understanding of the psychological and neurobiological bases of this individual variation in resistance to extinction will be important in the development of individualized treatment programs.

ACKNOWLEDGEMENTS

REFERENCES

Anselme, P. (2014). Incentive salience attribution under reward uncertainty: A Pavlovian model. *Behav Processes* 111C: 6-18.

Anselme, P., Robinson, M. J. and Berridge, K. C. (2013). Reward uncertainty enhances incentive salience attribution as sign-tracking. *Behav Brain Res* 238: 53-61.

Beckmann, J. S. and Chow, J. J. (2015). Isolating the incentive salience of reward-associated stimuli: value, choice, and persistence. *Learn Mem* 22(2): 116-127.

Beckmann, J. S., Marusich, J. A., Gipson, C. D. and Bardo, M. T. (2011). Novelty seeking, incentive salience and acquisition of cocaine self-administration in the rat. *Behav Brain Res* 216(1): 159-165.

Berridge, K. C. (2009). 'Liking' and 'wanting' food rewards: brain substrates and roles in eating disorders. *Physiol Behav* 97(5): 537-550.

Boakes, R. A. (1977). Performance on learning to associate a stimulus with positive reinforcement. Operant-Pavlovian Interactions. Davis, H., Hurwitz, H. (Eds.). Hillsdale, NJ; Earlbaum: 67–97.

Bouton, M. E. (2002). Context, ambiguity, and unlearning: sources of relapse after behavioral extinction. *Biol Psychiatry* 52(10): 976-986.

Bouton, M. E. (2004). Context and behavioral processes in extinction. *Learn Mem* 11(5): 485-494.

Bouton, M. E., Todd, T. P. and Leon, S. P. (2014). Contextual control of discriminated operant behavior. *J Exp Psychol Anim Learn Cogn* 40(1): 92-105.

Burns, M. and Domjan, M. (2001). Topography of spatially directed conditioned responding: Effects of context and trial duration. *Journal of Experimental Psychology: Animal Behavior Processes* 27(3): 269-278.

Cardinal, R. N., Parkinson, J. A., Hall, J. and Everitt, B. J. (2002). Emotion and motivation: the role of the amygdala, ventral striatum, and prefrontal cortex. *Neurosci Biobehav Rev* 26(3): 321-352.

Chang, S. E., Wheeler, D. S. and Holland, P. C. (2012). Roles of nucleus accumbens and basolateral amygdala in autoshaped lever pressing. *Neurobiol Learn Mem* 97(4): 441-451.

Clark, J. J., Collins, A. L., Sanford, C. A. and Phillips, P. E. (2013). Dopamine encoding of Pavlovian incentive stimuli diminishes with extended training. *J Neurosci* 33(8): 3526-3532.

Clark, J. J., Hollon, N. G. and Phillips, P. E. (2012). Pavlovian valuation systems in learning and decision making. *Curr Opin Neurobiol* 22(6): 1054-1061.

Conklin, C. A. and Tiffany, S. T. (2002). Applying extinction research and theory to cue-exposure addiction treatments. *Addiction* 97(2): 155-167.

Crombag, H. S., Bossert, J. M., Koya, E. and Shaham, Y. (2008). Review. Context-induced relapse to drug seeking: a review. *Philos Trans R Soc Lond B Biol Sci* 363(1507): 3233-3243.

Danna, C. L. and Elmer, G. I. (2010). Disruption of conditioned reward association by typical and atypical antipsychotics. *Pharmacol Biochem Behav* 96(1): 40-47.

Danna, C. L., Shepard, P. D. and Elmer, G. I. (2013). The habenula governs the attribution of incentive salience to reward predictive cues. *Front Hum Neurosci* 7: 781.

Delamater, A. R. (2012a). Issues in the extinction of specific stimulus-outcome associations in Pavlovian conditioning. *Behav Processes* 90(1): 9-19.

Delamater, A. R. (2012b). On the nature of CS and US representations in Pavlovian learning. *Learn Behav* 40(1): 1-23.

Delamater, A. R. and Oakeshott, S. (2007). Learning about multiple attributes of reward in Pavlovian conditioning. *Ann N Y Acad Sci* 1104: 1-20.

Deroche-Gamonet, V., Belin, D. and Piazza, P. V. (2004). Evidence for addiction-like behavior in the rat. *Science* 305(5686): 1014-1017.

DiFeliceantonio, A. G. and Berridge, K. C. (2012). Which cue to 'want'? Opioid stimulation of central amygdala makes goal-trackers show stronger goal-tracking, just as sign-trackers show stronger sign-tracking. *Behav Brain Res* 230(2): 399-408.

Domjan, M. (2003). Stepping outside the box in considering the C/T ratio. *Behavioural Processes* 62(1-3): 103-114.

Everitt, B. J. and Robbins, T. W. (2005). Neural systems of reinforcement for drug addiction: from actions to habits to compulsion. *Nat Neurosci* 8(11): 1481-1489.

Fitzpatrick, C. J., Gopalakrishnan, S., Cogan, E. S., Yager, L. M., Meyer, P. J., Lovic, V., Saunders, B. T., Parker, C. C., Gonzales, N. M., Aryee, E., Flagel, S. B., Palmer, A. A., Robinson, T. E. and Morrow, J. D. (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(10): e75042.

Flagel, S. B., Akil, H. and Robinson, T. E. (2009). Individual differences in the attribution of incentive salience to reward-related cues: Implications for addiction. *Neuropharmacology* 56 Suppl 1: 139-148.

Flagel, S. B., Cameron, C. M., Pickup, K. N., Watson, S. J., Akil, H. and Robinson, T. E. (2011a). A food predictive cue must be attributed with incentive salience for it to induce c-fos mRNA expression in cortico-striatal-thalamic brain regions. *Neuroscience* 196: 80-96.

Flagel, S. B., Clark, J. J., Robinson, T. E., Mayo, L., Czuj, A., Willuhn, I., Akers, C. A., Clinton, S. M., Phillips, P. E. and Akil, H. (2011b). A selective role for dopamine in stimulus-reward learning. *Nature* 469(7328): 53-57.

Flagel, S. B., Robinson, T. E., Clark, J. J., Clinton, S. M., Watson, S. J., Seeman, P., Phillips, P. E. and Akil, H. (2010). An animal model of genetic vulnerability to behavioral disinhibition and responsiveness to reward-related cues: implications for addiction. *Neuropsychopharmacology* 35(2): 388-400.

Flagel, S. B., Waselus, M., Clinton, S. M., Watson, S. J. and Akil, H. (2014). Antecedents and consequences of drug abuse in rats selectively bred for high and low response to novelty. *Neuropharmacology* 76 Pt B: 425-436.

Flagel, S. B., Watson, S. J., Robinson, T. E. and 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(3): 599-607.

Hearst, E., Jenkins, H. (1974). Sign-Tracking: The Stimulus-Reinforcer Relation and Directed Action. *Monograph of the Psychonomic Society, Austin.*

Holland, P. C. (1977). Conditioned stimulus as a determinant of the form of the Pavlovian conditioned response. *Journal of Experimental Psychology-Animal Behavior Processes* 3: 77-104.

Holland, P. C. (1992). Occasion setting in Pavlovian conditioning. The psychology of learning and motivation. Medin, D. San Diego, CA; Academic Press. 28: 69–125.

Holland, P. C., Asem, J. S., Galvin, C. P., Keeney, C. H., Hsu, M., Miller, A. and Zhou, V. (2014). Blocking in autoshaped lever-pressing procedures with rats. *Learn Behav* 42(1): 1-21.

Holland, P. C. and Bouton, M. E. (1999). Hippocampus and context in classical conditioning. *Curr Opin Neurobiol* 9(2): 195-202.

Konorski, J. (1967). Some new ideas concerning the physiological mechanisms of perception. *Acta Biol Exp (Warsz)* 27: 147-161.

Lomanowska, A. M., Lovic, V., Rankine, M. J., Mooney, S. J., Robinson, T. E. and Kraemer, G. W. (2011). Inadequate early social experience increases the incentive salience of reward-related cues in adulthood. *Behav Brain Res* 220(1): 91-99.

Lovic, V., Saunders, B. T., Yager, L. M. and Robinson, T. E. (2011). Rats prone to attribute incentive salience to reward cues are also prone to impulsive action. *Behav Brain Res* 223(2): 255-261.

Mahler, S. V. and Berridge, K. C. (2009). Which cue to "want?" Central amygdala opioid activation enhances and focuses incentive salience on a prepotent reward cue. *J Neurosci* 29(20): 6500-6513.

Meyer, P. J., Cogan, E. S. and Robinson, T. E. (2014). The form of a conditioned stimulus can influence the degree to which it acquires incentive motivational properties. *PLoS One* 9(6): e98163.

Meyer, P. J., Lovic, V., Saunders, B. T., Yager, L. M., Flagel, S. B., Morrow, J. D. and Robinson, T. E. (2012a). Quantifying individual variation in the propensity to attribute incentive salience to reward cues. *PLoS One* 7(6): e38987.

Meyer, P. J., Ma, S. T. and Robinson, T. E. (2012b). A cocaine cue is more preferred and evokes more frequency-modulated 50-kHz ultrasonic vocalizations in rats prone to attribute incentive salience to a food cue. *Psychopharmacology (Berl)* 219(4): 999-1009.

Morrow, J. D., Maren, S. and Robinson, T. E. (2011). Individual variation in the propensity to attribute incentive salience to an appetitive cue predicts the propensity to attribute motivational salience to an aversive cue. *Behav Brain Res* 220(1): 238-243.

Morrow, J. D., Saunders, B. T., Maren, S. and Robinson, T. E. (2014). Sign-tracking to an appetitive cue predicts incubation of conditioned fear in rats. *Behav Brain Res*.

Olshavsky, M. E., Shumake, J., Rosenthal, A. A., Kaddour-Djebbar, A., Gonzalez-Lima, F., Setlow, B. and Lee, H. J. (2014). Impulsivity, risk-taking, and distractibility in ratsexhibiting robust conditioned orienting behaviors. *J Exp Anal Behav*.

Paolone, G., Angelakos, C. C., Meyer, P. J., Robinson, T. E. and Sarter, M. (2013). Cholinergic control over attention in rats prone to attribute incentive salience to reward cues. *J Neurosci* 33(19): 8321-8335.

Peters, J. and De Vries, T. J. (2013). D-cycloserine administered directly to infralimbic medial prefrontal cortex enhances extinction memory in sucrose-seeking animals. *Neuroscience* 230: 24-30.

Rescorla, R. A. (2004). Spontaneous recovery. *Learn Mem* 11(5): 501-509.

Robinson, M. J., Anselme, P., Fischer, A. M. and Berridge, K. C. (2014a). Initial uncertainty in Pavlovian reward prediction persistently elevates incentive salience and extends sign-tracking to normally unattractive cues. *Behav Brain Res* 266: 119-130.

Robinson, T. E. and Berridge, K. C. (2008). Review. The incentive sensitization theory of addiction: some current issues. *Philos Trans R Soc Lond B Biol Sci* 363(1507): 3137-3146.

Robinson, T. E. and Flagel, S. B. (2009). Dissociating the predictive and incentive motivational properties of reward-related cues through the study of individual differences. *Biol Psychiatry* 65(10): 869-873.

Robinson, T. E., Yager, L. M., Cogan, E. S. and Saunders, B. T. (2014b). On the motivational properties of reward cues: Individual differences. *Neuropharmacology* 76 Pt B: 450-459.

Rosas, J. M., Todd, T. P. and Bouton, M. E. (2013). Context Change and Associative Learning. *Wiley Interdiscip Rev Cogn Sci* 4(3): 237-244.

Saunders, B. T., O'Donnell, E. G., Aurbach, E. L. and Robinson, T. E. (2014). A Cocaine Context Renews Drug Seeking Preferentially in a Subset of Individuals. *Neuropsychopharmacology*.

Saunders, B. T. and Robinson, T. E. (2010). A cocaine cue acts as an incentive stimulus in some but not others: implications for addiction. *Biol Psychiatry* 67(8): 730-736.

Saunders, B. T. and Robinson, T. E. (2011). Individual variation in the motivational properties of cocaine. *Neuropsychopharmacology* 36(8): 1668-1676.

Saunders, B. T. and Robinson, T. E. (2012). The role of dopamine in the accumbens core in the expression of Pavlovian-conditioned responses. *Eur J Neurosci* 36(4): 2521-2532.

Saunders, B. T. and Robinson, T. E. (2013). Individual variation in resisting temptation: implications for addiction. *Neurosci Biobehav Rev* 37(9 Pt A): 1955-1975.

Saunders, B. T., Yager, L. M. and Robinson, T. E. (2013). Cue-evoked cocaine "craving": role of dopamine in the accumbens core. *J Neurosci* 33(35): 13989-14000.

Stewart, J., de Wit, H. and Eikelboom, R. (1984). Role of unconditioned and conditioned drug effects in the self-administration of opiates and stimulants. *Psychol Rev* 91(2): 251-268.

Timberlake, W. (1994). Behavior systems, associationism, and Pavlovian conditioning. *Psychon Bull Rev* 1(4): 405-420.

Todd, T. P., Vurbic, D. and Bouton, M. E. (2014). Behavioral and neurobiological mechanisms of extinction in Pavlovian and instrumental learning. *Neurobiol Learn Mem* 108: 52-64.

Tomie, A., Aguado, A. S., Pohorecky, L. A. and Benjamin, D. (2000). Individual differences in pavlovian autoshaping of lever pressing in rats predict stress-induced corticosterone release and mesolimbic levels of monoamines. *Pharmacol Biochem Behav* 65(3): 509-517.

Tomie, A., Grimes, K. L. and Pohorecky, L. A. (2008). Behavioral characteristics and neurobiological substrates shared by Pavlovian sign-tracking and drug abuse. *Brain Res Rev* 58(1): 121-135.

Troisi, J. R. (2013). Perhaps More Consideration of Pavlovian-Operant Interaction May Improve the Clinical Efficacy of Behaviorally Based Drug Treatment Programs. *Psychological Record* 63(4): 863-893.

Tunstall, B. J. and Kearns, D. N. (2015). Sign-tracking predicts increased choice of cocaine over food in rats. *Behav Brain Res* 281: 222-228.

Wagner, A. R., Brandon, S. E. (1989). Evolution of a structured connectionist model of Pavlovian conditioning (AESOP). . Contemporary Learning Theories: Pavlovian Conditioning and the Status of Traditional Learning Theory. Klein, S. B., Mowrer, R. R. Hillsdale, NJ; Lawrence Erlbaum Associates, Inc.: 149-189.

Yager, L. M. and Robinson, T. E. (2010). Cue-induced reinstatement of food seeking in rats that differ in their propensity to attribute incentive salience to food cues. *Behav Brain Res* 214(1): 30-34.

Yager, L. M. and Robinson, T. E. (2013). A classically conditioned cocaine cue acquires greater control over motivated behavior in rats prone to attribute incentive salience to a food cue. *Psychopharmacology (Berl)* 226(2): 217-228.

Yager, L. M., Pitchers, K. K., Flagel, S. B. and Robinson, T. E. (2015). Individual variation in the motivational and neurobiological effects of an opioid cue. *Neuropsychopharmacology* 40(5): 1269-1277.

Zener, K. (1937). The significance of behavior accompanying conditioned salivary secretion for theories of the conditioned response. *American Journal of Psychology* 50: 384–403.

Figure 1. Acquisition of Pavlovian conditioned approach behavior in STs ($n = 72$) and GTs ($n = 61$) from experiments 1-4. Sign-tracking and goal-tracking behaviors during the CS period are shown in six measures (mean ± SEM): **A)** the probability of deflecting the lever, **B)** the probability of entering the food cup, **C)** the number of lever deflections, **D)** the number of food cup head entries, **E)** the latency to deflect the lever, **F)** and the latency to enter the food cup.

Figure 2. Each symbol represents the PCA Index Score for an individual rat, first averaged across days 4-5 of initial PCA training and then again after days 4-5 of discrimination (R/NR) training (and for each rat these two measures are connected with a line). Most rats maintained PCA Index Scores in the same direction after discrimination training, but in some rats in both experiments 2 **(A)** and 3 **(B)**, the initial preference for the lever or food cup changed as a result of discrimination training. Subsequent analyses include only the rats that maintained scores in the direction of their original classification (STs > 0 and GTs < 0).

Figure 3. In experiment 1, GTs ($n = 21$) showed faster extinction of PCA behavior than STs ($n = 28$). The figures show the mean (+ SEM) probability of food cup approach in GTs **(A)** and lever approach in STs **(B)** during extinction sessions compared to the last PCA session (baseline). Top panels show within-session probability of approach in 3-trial blocks. **A)** GTs showed significant decreases in approach during the first extinction session (Ext 1) and the fourth extinction session (Ext 4) compared to baseline. **B)** STs continued to approach the lever throughout first extinction session (Ext 1), but showed a significant decrease in approach by Ext 4. When behavior was averaged across all 25 trials per session **(C)**, and the first three trials of each session **(D)**, the GTs showed significant decreases from baseline in earlier sessions than the STs; * $p < .05$, *** $p < .001$, change to relative to baseline. **E)** In a subset of animals (25 STs and 9 GTs), the probability of approach during the first 3 trials remained significantly lower in GTs than STs during additional extinction sessions (extinction days 1-8); # $p < .001$, main effect of group.

Figure 4. In experiment 2, rewarded (R) and non-rewarded (NR) periods were indicated a change in chamber illumination (red or white lights illuminated for the duration of 15-trial blocks). The figure shows the mean (+SEM) probability of approaching the lever or food cup in

STs ($n = 7$) and GTs ($n = 11$), averaged in 3-trial blocks during R and NR periods on days 1 and 5 of discrimination training. The dashed horizontal lines near the top of each panel show mean probability of lever approach (in STs) or food cup approach (in GTs) recorded during the final day of PCA training (PCA 5), prior to discrimination training. GTs showed rapid discrimination between R and NR phases in the very first session **(A)**, and some improvement by the fifth session **(B)**. In contrast, STs showed no discrimination in session 1 **(C)**, but by the fifth session STs discriminated between R and NR periods **(D)**. Note that although in this and subsequent figures the R block is shown as occurring first, in fact, for half the animals the NR block occurred first, and the order was also counter-balanced across days. There was no effect of whether the R or NR blocks occurred first in the session.

Figure 5. In experiment 2, GTs ($n = 11$) learned to discriminate between R and NR periods faster than STs ($n = 7$). Behavioral responses (mean + SEM) during R versus NR periods, as well as the last day of PCA training (PCA 5), are shown for lever-directed behavior in STs **(A-C)** and food cup-directed behavior in GTs **(D-F).** The STs gradually learned to discriminate between R and NR periods with repeated training, and showed significant differences by session 5. The GTs showed rapid discrimination that was significant in all 5 sessions; ** $p < .01$, *** $p < .001$, post hoc t tests comparing R and NR blocks. **G-I**) Difference scores (R responses - NR responses) were significantly higher in GTs than STs; ++ $p < .01$, +++ $p < .001$, significant group effects; † $p < .05$, †† $p < .01$, post hoc t tests comparing STs and GTs.

Figure 6. In experiment 3, R and NR periods were signaled by spatially and temporally discrete cues (a cue light located at the front or back of the chamber illuminated for only 6s prior to the CS). The figure shows the within-session probability of GTs ($n = 18$) and STs ($n = 15$) approaching the lever or food cup (mean + SEM). The dashed horizontal lines near the top of each panel show mean probability of lever approach (for STs) and food cup approach (for GTs) during the last day of PCA training. GTs showed excellent discrimination in both session 1 **(A)** and session 5 **(B)**, whereas STs showed no discrimination in session 1 **(C)** and improvement by session 5 **(D)**.

Figure 7. In experiment 3, as in Experiment 2, GTs ($n = 18$) showed faster discrimination between R and NR periods than STs ($n = 15$). The figures show mean (+ SEM) lever-directed behavior in STs **(A-C)** and food cup-directed behavior in GTs **(D-F)** during the last PCA training day (PCA 5) and then the 5 discrimination sessions. STs initially showed no discrimination between R and NR periods, but significant differences emerged by the third session. In contrast, the GTs showed significant discrimination in all 5 sessions; * $p < .05$, ** $p < .01$, *** $p < .001$, post hoc *t* tests comparing R and NR blocks after significant interactions. **G-I)** For probability, contacts, and latency, differences between R and NR blocks were significantly greater in GTs than STs; +++ $p < .001$, significant group effects.

Figure 8. In experiments 2 and 3, we examined behavior during the first trial of R and NR blocks, when the light signals indicated whether or not the following CS would be reinforced. Responding during the very first trial of a new block was compared to the average the preceding 3 trials in the previous block, and difference scores were calculated (R minus NR, black symbols labeled R), and after the switch from rewarded to non-rewarded phases (NR minus R, white symbols labeled NR). In experiment 2 both STs **(A)** and GTs **(B)** learned to discriminate between the diffuse light cues in later sessions, showing an increase in contacts during the first trial of R phases and a decrease in contacts during the first trial of NR phases. In experiment 3, STs learned to discriminate between the discrete light cues by the last session **(C)**, but GTs did not **(D)**; * $p < .05$, significant interaction effects.

Figure 9. In experiment 4 STs ($n = 14$) and GTs ($n = 10$) were first trained to make an instrumental response (nose poke) for a food reward. **A)** STs and GTs initially acquired stable responding for food pellets at similar rates. Panel A shows active and inactive nose-poke (NP) responses during 8 days of acquisition training, broken into 3 phases with different reinforcement schedules and session lengths (FR1, fixed ratio 1; VI, variable interval). **B)** After training, when R and NR periods were introduced, both groups showed significant discrimination in all 10 sessions. However, the GTs showed higher rates of responding during R blocks than STs; *$p < .05$, main effect of group. **C)** Difference scores calculated from active nose-pokes (R minus NR) were also significantly higher in GTs than STs; *$p < .05$, main effect of group. However,

the absence of an interaction effect indicates the two groups learned the discrimination at the same rate. All results are mean ± SEM.