Published in final edited form as:

Behav Brain Res. 2017 January 15; 317: 122–131. doi:10.1016/j.bbr.2016.09.033.

Learning and generalization from reward and punishment in opioid addiction

Catherine E. Myers^{1,2}, Janice Rego³, Paul Haber⁴, Kirsten Morley⁴, Kevin D. Beck, PhD^{1,2}, Lee Hogarth^{5,6}, and Ahmed A. Moustafa^{1,7}

¹Department of Veterans Affairs, New Jersey Health Care System, East Orange, NJ, USA

²Department of Pharmacology, Physiology & Neuroscience, New Jersey Medical School, Rutgers, The State University of New Jersey, Newark, NJ, USA

³School of Social Sciences and Psychology, Western Sydney University, Sydney, NSW, Australia

⁴NHMRC Centre of Research Excellence in Mental Health and Substance Use, Discipline of Addiction Medicine, Central Clinical School, The University of Sydney, NSW, Australia

⁵School of Psychology, University of New South Wales, Sydney, NSW, Australia

⁶School of Psychology, University of Exeter, Exeter, UK

⁷Marcus Institute for Brain and Behaviour, Western Sydney University, Sydney, NSW, Australia

Abstract

This study adapts a widely-used acquired equivalence paradigm to investigate how opioid-addicted individuals learn from positive and negative feedback, and how they generalize this learning. The opioid-addicted group consisted of 33 participants with a history of heroin dependency currently in a methadone maintenance program; the control group consisted of 32 healthy participants without a history of drug addiction. All participants performed a novel variant of the acquired equivalence task, where they learned to map some stimuli to correct outcomes in order to obtain reward, and to map other stimuli to correct outcomes in order to avoid punishment; some stimuli were implicitly “equivalent” in the sense of being paired with the same outcome. On the initial training phase, both groups performed similarly on learning to obtain reward, but as memory load grew, the control group outperformed the addicted group on learning to avoid punishment. On a subsequent testing phase, the addicted and control groups performed similarly on retention trials involving previously-trained stimulus-outcome pairs, as well as on generalization trials to assess acquired equivalence. Since prior work with acquired equivalence tasks has associated stimulus-outcome learning with the nigrostriatal dopamine system, and generalization with the hippocampal region, the current results are consistent with basal ganglia dysfunction in the opioid-addicted patients. Further, a selective deficit in learning from punishment could contribute to processes by which addicted individuals continue to pursue drug use even at the cost of negative consequences such as loss of income and the opportunity to engage in other life activities.

Keywords

Opioid addiction; heroin; reward learning; punishment learning; generalization; acquired equivalence

1. Introduction

Addiction can be thought of as a disorder in which addicted individuals continue to seek out and use the addictive substance despite negative consequences, such as loss of income and loss of the opportunity to engage in other activities (1, 2). It is possible then that addiction is not merely about heightened reward-seeking, but rather an imbalance between the ability to adjust behavior to maximize reward while simultaneously avoiding punishment.

Of particular societal concern are the highly-addictive opioid drugs. Due to an increasing use of medically-prescribed opioid painkillers, accidental addiction to prescription opiate drugs is estimated to affect approximately 2 million people in the U.S. alone (3), and abuse of these drugs often leads to use of illegal opiates such as heroin, once access to the prescriptions drugs is withdrawn. Opioid addiction is extremely difficult to overcome, even when the addicted individual strongly desires to stop using the drug, and there is a high relapse rate for individuals who have undergone detoxification treatment (4). One option for individuals who have tried and failed to overcome opioid addiction is maintenance treatment, involving medically-supervised use of opioids such as methadone. However, a recent review of outcomes following maintenance therapy reported that, one month after discontinuation of treatment, rates of relapse to illicit opioid use exceed 50% (5). Given these figures, it is of great importance to understand reinforcement-based learning processes in opioid-addicted individuals, in order to develop more effective therapies to aid these individuals in overcoming their drug dependence.

1.1. Opiate addiction and reward processes

Numerous prior studies have examined how opioids affect reward pathways in the brain (6–8), including studies suggesting that opioid-addicted individuals are particularly impaired at temporal discounting tasks that require foregoing a small immediate reward in favor of a larger delayed reward (9, 10). However, few studies have attempted to examine the balance between reward-based and punishment-based learning in opioid-addicted individuals, or how this learning might generalize when familiar stimuli are presented in new ways.

In a prior study, Myers et al. (11) used a categorization task that interleaved trials on which participants learned to avoid negative consequences (point loss) with trials on which they learned to obtain positive consequences (point gain). There was also an ambiguous “no-feedback” outcome, which could signal either missed opportunity for reward or else successful avoidance of punishment. The category mappings were probabilistic, such that a particular stimulus belonged to one category on 80% of the trials on which it appeared and to the other class on 20% of trials. Thus, expectancies were sometimes violated even for a subject who had learned the category mappings well. The optimal strategy was therefore to choose the most-often correct category for each subject (“probability maximizing”), and to

continue to execute this strategy throughout the task, despite the occasional trials on which the response would fail to pay off. Results from this prior study showed that opiate-addicted individuals and never-addicted controls performed comparably in terms of total points accrued on the task. However, the addicted participants exhibited an increased tendency to shift response strategies after an unexpected loss (either a punishment or an omission of expected reward). In other words, the addicted group tended to “chase reward” by altering their response strategy whenever they experienced a negative outcome, while controls were more likely to “stick with” a stable response strategy that maximized long-term gain. However, another interpretation of the results is simply that opioid-addicted individuals were better at learning from punishment than controls, resulting in them being more able to adjust responding after a loss than controls. In the particular context of a probabilistic categorization task, this led to suboptimal behavior – but in a deterministic task, it might result in superior punishment-based learning.

The current task is designed to investigate this issue further, by considering a reward- and punishment-learning task where cue-outcome mappings are deterministic. If opioid-addicted individuals are simply better than controls at learning to adjust responding based on punishment, then they should similarly outperform controls on punishment-based trials in this deterministic task. On the other hand, if the results from the prior study were specifically due to the probabilistic nature of the task, then the opioid-addicted group might not outperform controls on a deterministic task. In addition, the prior study left open the question of whether learning would generalize in the same fashion in opioid-addicted and never-addicted groups, and if so, whether generalization was equivalent to stimuli that had previously been associated with reward versus those that had previously been associated with punishment.

To explore these questions, we adapted a widely-studied learning paradigm, acquired equivalence, in which prior training to treat two stimuli as equivalent increases generalization between them (12–14). In one, widely-used version of the paradigm, participants learn via trial and error to pair each of several antecedents with a consequent; some antecedents are implicitly equivalent in the sense that they should be paired with the same consequent. Subsequently, a subset of antecedents are paired with new consequents. Prior work has shown that healthy controls reliably generalize this new learning, tending to pair equivalent antecedents with the same consequents, even though these pairings were never explicitly trained (15).

1.2. Brain Substrates of Acquired Equivalence

A computer-based version of acquired equivalence has previously been used to demonstrate qualitative differences in learning vs. generalization in a number of psychiatric and neurological patient groups (for review, see 16). For example, the learning of stimulus-response pairs appears to depend on frontostriatal circuits, and is disrupted in individuals with frontostriatal dysfunction, such as patients with Parkinson’s disease tested on normal dopaminergic medication (15, 17), who show slow learning followed by successful generalization.

On the other hand, generalization appears to depend on medial temporal (hippocampal) function; thus, amnesic patients with bilateral hippocampal destruction (18) and nondemented elderly with hippocampal atrophy consistent with prodromal Alzheimer's disease (AD) (15) both spared learning followed by impaired generalization. Patients symptomatic for early (AD) show similar deficits on generalization, although they also show slower learning than age-matched controls, consistent with a more diffuse pattern of accumulating brain pathology in early AD (19). Disrupted generalization on acquired equivalence tasks is also seen in other psychopathologies that commonly involve hippocampal-region volume reductions, including schizophrenia (20, 21) and post-traumatic stress disorder (22).

Together, these studies suggest that the acquired equivalence paradigm provides a platform to dissociate frontostriatal-dependent associative learning from medial temporal-dependent generalization. Further evidence for this dissociation comes from functional neuroimaging (fMRI) studies in healthy young adults, which show a positive relation between caudate activity and performance on the initial feedback-based learning, while increasing activation in the hippocampus during training correlates with performance on subsequent generalization tests (23).

The importance of the hippocampal region for generalization in the acquired equivalence paradigm is consistent with theories suggesting that the hippocampal region helps establish stimulus representations that support later flexible use of the learned information (24, 25). On the other hand, the role of the basal ganglia in stimulus-response learning during the training phase is consistent with theories that emphasize the role of the nigrostriatal dopamine system in habit learning and action selection. For example, reinforcement learning (RL) models assume that trial-and-error learning results in the learner coming to choose actions that are expected to maximize reward and/or minimize punishment. Prediction error (PE), the difference between expected and experienced outcomes, is used to update the learner's expectations and guide action selection. A large body of single-unit neurophysiology implicates phasic dopamine signals in encoding PE during classical and instrumental conditioning (26–28), with dopamine targets in the anterior striatum associated with reward-based PE estimates (29), and anterior insula and dorsal striatum associated with punishment-based learning (30).

Addiction is, at its core, a biological process that involves drugs of abuse acting on the brain. Addictive substances appear to exert their effects upon a core set of brain structures that include the structures implicated in reinforcement learning and generalization. Thus, many drugs of abuse (such as cocaine and amphetamines) act directly on dopamine receptors and dopaminergic neurons in the VTA, while others (such as heroin and opioid medications) act indirectly via opiate receptors that modulate dopamine function (for review, see 31). The hippocampal region is also rich in dopamine receptors (32), and hippocampal functions, such as contextual processing, may play a key role in addiction and relapse, including an individual's response to drug-related cues and contexts (33). Given that the basal ganglia and hippocampus play central roles in drug effects and drug addiction, the acquired equivalence paradigm has been examined in the context of substance abuse disorders.

One prior study considered long-term cocaine users, and found that they were worse than non-addicted controls during the learning phase – specifically, on learning to map familiar antecedents to new consequents – but subsequently unimpaired on generalization (34). Conversely, abstinent individuals with a history of alcohol dependence showed the opposite pattern of spared learning but impaired generalization (35). These results suggest that cocaine addiction may be more associated with hippocampal dysfunction while alcohol dependence may be more associated with basal ganglia dysfunction. However, acquired equivalence has not yet been considered in the context of opioid addiction. Given that opioids affect brain structures implicated in reinforcement learning, and also brain structures implicated in generalization, it is possible that opioid-addicted patients would show abnormalities on the learning and/or generalization components of the acquired equivalence task.

1.3. Objectives of the Current Work

The current study used a novel task variant that embeds aspects of both acquired equivalence and reward-based and punishment-based training, as schematized in Table 1. Thus, during training Stages 1 and 2, some antecedents (A, C) are associated with one consequent (X1) while other antecedents (B, D) are associated with a different consequent (Y1); thus A is implicitly equivalent to C and B is implicitly equivalent to D. Two antecedents (A, C) are used for reward-based training, meaning that correct responses are rewarded with positive feedback and point gain while incorrect responses receive no feedback or point gain; the other antecedents (B, D) are used for punishment-based training, meaning that incorrect responses are punished with negative feedback and point loss while correct responses receive no feedback or point loss. Thus, the no-feedback outcome is ambiguous, as it could signal either missed reward (for a reward-based trial) or successfully avoided punishment (for a punishment-based trial). Next (training Stage 3), some antecedents are paired with new consequents (A with X2 and B with Y2). Finally, subjects are tested (without feedback) on all trained pairs as well as on new pairs, to assess prior learning as well as generalization for both reward- and punishment-based associations, within and across subjects.

The first question was whether, in this deterministic task, opioid-addicted participants would show facilitated punishment-based learning, relative to control participants; the second question was whether, having learned the initial antecedent-consequent pairs, the opioid-addicted patients would generalize well, like cocaine users and patients with frontostriatal dysfunction, or generalize poorly, like hippocampal-damaged patients and alcohol abusers – and whether this would apply equally to the generalization of reward-based and punishment-based learning.

1. Materials and Methods

1.1. Participants

The addicted group consisted of 33 opioid-dependent patients, recruited from the Royal Prince Alfred Hospital (RPAH) Drug Clinic Services and formed the addicted group. Inclusion criteria were heroin as principal drug of referral, current participation in methadone maintenance treatment, and educational level of at least 7 years. All patients

were undergoing methadone treatment on an outpatient basis; most patients were tested between 9:30AM and noon that is, within a few hours after receiving their daily methadone dose; thus patients were experiencing the acute actions of the drug. Daily methadone dosing was confirmed by clinic records. Prior heroin use was confirmed by urine drug screening at RPAH (by neurologists and medical registrars) before the commencement of methadone treatment. Subsequent use or non-use (of drugs other than methadone) was assessed via patient self-report during the testing session.

32 control participants, all of whom self-reported no current or past use/abuse of opioid drugs, were recruited from the community and Western Sydney University. Demographic information for each group is shown in Table 2. The groups did not differ significantly in age ($t(54.3)=1.67$, $p=.101$) or gender distribution (Fisher's Exact test, 2-sided, $p=.310$); there were no significant age differences between males and females, or between males and females in the two groups (ANOVA, all $p>.05$). However, the control group had significantly more years of education than the addicted group ($t(63)=8.91$, $p<.001$); in fact, education ranged from 10–15 years in the control group (i.e. all had at least some secondary education and many had some college) but ranged from 7 to 13 in the addicted group (with 19 participants reporting fewer than 10 years of education).

All but 2 patients in the addicted group reported cigarette use (ranging from 1 to 40 cigarettes per day). Most also reported mild alcohol use (e.g. social drinking) although 8 reported past problem drinking (>1 year) and 2 reported current heavy alcohol use. Many addicted participants reported current (within past month, $n=15$) or past ($n=8$) use of marijuana (e.g. 1–2 joints per day). Other self-reported secondary drug use in the addicted group included 3 individuals with current use of benzodiazepines and 7 with current use of methylamphetamines; 5 reported past (not current) use of cocaine and 6 reported past (not current) use of amphetamines.

Because secondary drug use was not an exclusion criterion in the addicted group, we also did not exclude control participants based on current or past use of drugs (excluding opioids). 9 control participants reported current (within past month) tobacco use, and an additional 5 reported past (not current) use. Most control participants reported mild alcohol use (e.g., social drinking), although none reported problem drinking or alcohol abuse; additionally two reported current marijuana use (1–2 joints per day), and 9 reported lifetime use. Additionally, seven reported past history of occasional use of (though not addiction to) drugs including amphetamines, cocaine, MDMA and/or LSD. No controls self-reported history of problem use of or addiction to any of these substances.

The study was approved by the Royal Prince Alfred Hospital Ethics Committee and the Ethics Committee at Western Sydney University, New South Wales, Australia. All participants provided written informed consent before the start of testing. Procedures conformed to guidelines established by the Declaration of Helsinki for the protection of human subjects. Participants were reimbursed AUD \$20 (about USD \$15) for their participation.

1.2. Behavioral Task

1.2.1. Materials—Antecedent stimuli were taken from a set of four drawings of cartoon characters (man, woman, boy, girl); consequent stimuli were taken from a set of four drawings of food items including two entrees (hamburger, hot dog) and two desserts (pumpkin pie, donut) as shown in Figure 1A. Images were modified using Adobe Photoshop from drawings provided as public domain clipart under Creative Commons license (CC0 1.0 Universal) at www.openclipart.org.

For each subject, the four characters were randomly assigned to be antecedents A, B, C, and D, and the two entrees were randomly assigned to be consequents X1 and Y1, and the two desserts randomly assigned to be consequents X2 and Y2. Among the antecedents, the boy and woman had blond hair while the girl and man had brown hair. Thus, each antecedent had three obvious, binary-valued features: age (adult vs. child), gender (male vs. female), and hair color (brown vs. blond). Each antecedent therefore shared exactly one feature with each other antecedent. The antecedents and consequents appeared about 1 in (approx. 2.5cm) high on the computer screen. Consequents could appear in either left-right order on the screen.

2.2.2. General Procedure—The computer-based task was a modified version of the acquired equivalence task described by Myers and colleagues (15). Software was presented on a Macintosh laptop, programmed in the SuperCard environment (Solutions Etcetera, Pollock Pines CA). Testing was conducted in a quiet, well-lit room, with the subject able to position his/her seating to achieve a comfortable viewing distance from the computer screen, typically 50–60 cm.

At the start of the training phase, the following instructions appeared on the screen: “*In part 1, you will see drawings of four people in a family. They look like this <drawings inserted>. Your job is to learn which food each person likes best. In the beginning, you will have to guess, but try to learn the correct answers. Sometimes, if you guess correctly, you’ll win points. Sometimes, if you guess wrong, you’ll lose points. Try to make as many points as you can. (We’ll start you off with 500 points now.)*” The experimenter read these instructions aloud to the participant and then clicked the computer mouse button to begin the training phase.

On each trial, an antecedent (character) appeared with two consequents (food items) as shown in Figure 2A, and the prompt: “*Which food does this person like better?*” The participant responded by clicking on one of the food items, which was then circled.

2.2.3. Training Phase—The training phase included three stages, each with increasing numbers of trial types, as illustrated in Table 1. Half of the trial types were reward-based, meaning that correct responses triggered feedback (“Correct!”; Figure 2B) and point gain (+10), while incorrect responses triggered no feedback or point gain (Figure 2C). The other half of trial types were punishment-based, meaning that incorrect responses triggered feedback (“Incorrect!”) and point loss (–10) (Figure 2D), while correct responses triggered no feedback or point loss. A running total of points appeared at the bottom right hand corner of the screen.

Since consequents could appear in either left-right ordering, there were 4 trial types in Stage 1, increasing to 8 in Stage 2 and 12 in Stage 3. Each stage consisted of a maximum of five blocks, each block consisting of one instance of each trial type, or terminated early if the participant reached criterion performance of a number of consecutive correct responses equal to one block of trials. Trial order within a block was randomized for each participant. The start of a new stage was not signaled to the participant. Because training stages contain different number of trials for each participant, a simple percent error measure is not appropriate, because two participants could have similar numerators but very different denominators; instead, following convention for this type of acquired equivalence task (e.g., 15, 17, 18, 19, 22, 34, 36, 37), we report total errors to criterion on each stage for reward-based and punishment-based antecedents, as well as number of trials to criterion on each stage.

We also report the percent of subject in each group who reached criterion on each type of trial (reward-based or punishment-based). Subjects who completed Stage 3 in fewer than 60 trials reached criterion on both trial types; however, a subject who did not finish Stage 3 early might still have reached criterion with one trial type (e.g. reward-based trials) if he or she made no errors on the last 6 trials of that type, even if he or she was still making errors on the other trial type.

2.2.4. Testing Phase—At the conclusion of Stage 3, the following instructions appeared: *“Good! Now, in part 2, you will need to remember what you learned so far. You will NOT be shown the correct answers. At the end of the experiment, the computer will tell you how many times you guessed right.”*

The testing phase consisted of 48 trials, including three blocks of 16 trial types, including all six pairs from the training phase (that is, retention trials, as they are the same as in prior training phase), plus the two new pairs shown in Table 1 (generalization trials), each with the consequents in each possible left-right ordering. Correct responding on generalization trials was defined as selecting the consequent consistent with acquired equivalence as illustrated in Table 1: i.e., $C \rightarrow X2$ (equivalent to A) and $D \rightarrow Y2$ (equivalent to B).

Trial events were the same as in the training phase, except that no corrective feedback or point gain/loss occurred. Trial order within a block was randomized for each participant. Because all subjects received the same number of testing trials, results for the test phase are reported as percent error, for retention and generalization trials with reward-based and punishment-based antecedents.

2. Results

3.1. Training Phase

Figure 3A shows number of trials increased across training stages, in each group. Mixed ANOVA with within-subjects factors of training stage, between-subjects factors of group (control vs. addicted), and covariates of age, sex and education, revealed a stage $\times$ group interaction ($F(2,120)=4.06$, $p=.020$) with no other effects or interactions (all $p>.100$). Specifically, the control group required fewer Stage 3 trials than the addicted group

($t(59.4)=3.08$, $p=.003$), with no significant differences on Stage 1 ($t(63)=0.40$, $p=.693$) or Stage 2 ($t(63)=1.29$, $p=.201$).

Figure 3B shows mean total errors at each training stage for each group, divided by trial type (reward-based vs. punishment-based). Mixed ANOVA with within-subjects factors of trial type (reward vs. punishment) and training stage, between-subjects factors of group (control vs. addicted), and covariates of age, sex and education, revealed a significant stage $\times$ group interaction ($F(2,120)=4.10$, $p=.019$); no other effects or interactions approached significance (all $p>.100$ except the within-subjects effect of stage, $F(2,120)=2.67$, $p=.073$).

To further examine the interaction, post-hoc t-tests were conducted to compare the groups in terms of errors on reward-based and punishment-based trials at each stage, with alpha corrected to $.05/6=.008$ to protect significance, and Welch's t used where the data violated assumptions of equal variance (Levene's test $p<.05$). In Stage 1, the two groups made equivalent errors on both reward-based ($t(56.1)=0.86$, $p=.413$) and punishment-based ($t(52.4)=2.26$, $p=.028$) trials; however on Stage 2, the addicted group made significantly more errors than the control group on punishment-based trials ($t(59.3)=2.80$, $p=.007$), although the groups performed equivalently on reward-based trials ($t(63)=1.47$, $p=.147$). Similarly, in Stage 3, the addicted group made significantly more errors on punishment-based trials ($t(34.9)=4.37$, $p<.001$) but the group difference on reward-based trials fell short of corrected significance ($t(63)=2.27$, $p=.026$).

As shown in Figure 3B, there were also within-subjects differences; specifically, although the control group made similar rates of errors to reward-based and punishment-based trials in Stage 1 (Bonferroni-corrected paired t-tests, $\alpha=.008$, $t(31)=2.45$, $p=.020$) and Stage 2 ($t(31)=1.40$, $p=.173$), they made significantly more errors to reward-based than punishment-based trials at Stage 3 ($t(31)=3.91$, $p<.001$). In contrast, the addicted group performed equivalently on reward-based and punishment-based trials at each training stage (all $t<2.0$, all $p>.050$).

3.2. Testing Phase

Figure 4 shows performance on old (retention) and new (generalization) pairs during the testing phase. First, to examine retention performance, mixed ANOVA was conducted, with within-subjects factor of trial type (reward-based vs. punishment-based) and between-subjects factor of group (control vs. opioid-addicted), and covariates of age, sex, and education. No main effects or interactions approached significance (all $p>.100$). These results did not change if training phase performance (i.e., stage 3 errors) on reward-based and punishment-based trials were added as covariates.

Second, to examine generalization, mixed ANOVA was conducted on performance to the new pairs. Again, there was no effect of group and no interactions (all $p>.100$). Results were similar when training phase performance (i.e., Stage 3 errors) on reward-based and punishment-based trials were added as covariates.

3. Discussion

The current study compared performance by a group of opioid-addicted patients on current methadone maintenance, and a group of never-addicted controls, on a novel acquired equivalence task that included separate reward-based and punishment-based learning components. Both groups were comparable at reward-based learning, even as memory load grew across training stages. On punishment-based learning, the control group outperformed the opioid-addicted group at later training stages that included more antecedent-consequent pairs to be learned. On the testing phase, both groups performed equivalently on retention trials involving previously-trained pairs, and also on hippocampal-dependent generalization trials, where subjects were challenged to demonstrate acquired equivalence.

We discuss each of these points further below.

4.1. Learning from Reward and Punishment

In a prior study using a probabilistic reward- and punishment-learning task, Myers et al. (11) found that opioid-addicted individuals were more prone to show lose-shift behavior after an unexpected negative outcome; one interpretation of that result might be that the addicted patients were simply better at learning to avoid punishment. One aim of the current study was to test that idea, by comparing opioid-addicted and control groups on a deterministic version of the reward- and punishment-learning task, where both rewarding and punishing outcomes are perfectly predictable. As illustrated in Figure 3B, the opposite effect was obtained: specifically, while the opioid-addicted and control groups were comparable on reward-based learning, the addicted group performed worse than controls on punishment-based learning. This in turn suggests that the increased lose-shift behavior of opioid-addicted participants, previously observed by Myers et al. (11), did not simply reflect superior learning to adjust responding after negative outcomes, but rather reflected impairment in the ability to “stick with” successful strategies to maximize performance in the face of unexpected negative outcomes during a probabilistic task.

A possible alternative explanation for the current findings is via reinforcement learning (RL) theories. As noted above, RL theories assume that learning is based on a prediction error signal that is used to update the learner's expectations and guide action selection; PE may be encoded by nigrostriatal dopamine (38); targets in the ventral striatum then use these signals to update predictions and guide learning (39). Piray et al. (40) examined patients with Parkinson's disease, a disorder that involves progressive cell loss in the ventral striatum, who had also developed impulse control disorders. When these patients were tested on their normal dopaminergic medication, the patients showed enhanced reward learning, but impaired punishment learning, relative to matched controls. Piray et al. attributed this impaired punishment learning to a dysfunctional critic module (in the patients' ventral striatum) which resulted in underestimation of adverse consequences, impairing the ability to modify responding to avoid those adverse consequences. A similar basis could underlie the poor punishment-based learning in the addicted patients of the current study. Functional neuroimaging studies could be useful here, to better understand whether individual differences in ventral striatum activity during punishment-based learning contribute to group differences on the task.

An important feature of the current task, as well as of the prior probabilistic reward- and punishment-learning task studied by Myers et al. (11), is the presence of an ambiguous, no-feedback outcome that can signal either missed opportunity for reward or successful avoidance of punishment. Thus, successful performance on punishment-based trials in the current task involves learning that the no-feedback outcome is reinforcing. It may be that the opioid-addicted individuals have difficulty learning this, particularly in the presence of intermixed reward-based trials where the same no-feedback outcome should be interpreted as a punisher. This possibility should be explored further in future work.

4.2. Generalization

A second aim of the current study was to examine the degree to which reward-based and punishment-based learning would generalize. Accordingly, the initial training was followed by a testing phase, which included retention trials to test previously-learned antecedent-consequent pairs, as well as generalization trials to test acquired equivalence, i.e. the extent to which learning about one antecedent would generalize to another antecedent that had been treated equivalently in the past. Figure 4A shows that both groups performed equivalently on retention trials involving both reward-based and punishment-based associations. Performance on punishment-based trials was somewhat better than performance on reward-based trials, although this did not reach statistical significance, which is generally consistent with the fact that subjects (particularly control participants) made fewer errors on punishment-based trials during training (Figure 3B).

There were no differences between groups on generalization trials (Figure 4B). Thus, the current results provide no support for the idea that opioid-addicted individuals generalize differentially than controls when information has been acquired via reward-based vs. punishment-based training. To our knowledge, this is the first study that has examined acquired equivalence in opioid-addicted individuals versus healthy controls. However, the absence of a group difference in generalization is similar to that previously observed in a prior study with long-term cocaine users (34), who were also impaired at learning but not at generalization. Given the body of prior work implicating the hippocampal region in generalization on acquired equivalence tasks (15, 18, 19, 23), the current findings do not provide evidence of impaired hippocampal function in the opioid-addicted patients.

However, it should be noted that total generalization errors were near chance levels (50% errors) in both groups. By contrast, prior studies with the computer-based acquired equivalence task have reported much better performance particularly in control groups, with control groups often averaging at or below 10–15% errors on generalization pairs (20, 21, 34), although one prior study using an Australian sample found error rates of about 31% in healthy elderly subjects (41). These prior studies did not include separate reward- and punishment-based trials; rather, on each trial the subject received explicit reward (feedback and point gain) for a correct response or explicit punishment (feedback and point loss) for an incorrect response. It is possible that the current task version, with separate reward-based and punishment-based trials, affects generalization. Further work is needed to better understand how different training regimes affect transfer performance on this type of associative learning task.

4.3. Limitations and Future Directions

One limitation of the current study, which considers opioid-addicted patients tested shortly after receiving their daily methadone dose, is that it cannot distinguish between the acute effects of current methadone maintenance treatment or the long-term effects of heroin (and methadone) use; it is also possible that group differences in behavior represent cognitive biases that were present prior to onset of opioid use and possibly conferred risk for exposure to or addiction to opioids. Clearly, this question is difficult if not impossible to disentangle in human studies, since otherwise healthy controls cannot be administered heroin or methadone for experimental purposes; conversely, it would be unethical and potentially medically dangerous to withhold methadone treatment from already-addicted individuals to examine behavioral effects. Even if it were possible, such a manipulation would introduce new confounds in the form of withdrawal symptoms. Given this reality, many other recent studies of cognition and brain function in patients undergoing methadone maintenance treatment have used the same design of comparing patients vs. healthy (never-addicted) controls (e.g., 11, 42, 43) – although a few studies have compared methadone-maintenance patients against individuals with opiate addiction not undergoing such treatment, in addition to comparing both patient groups against never-addicted controls (e.g., 44). A viable alternative to studies of already-addicted humans would be animal studies, where drug exposure can be randomized and controlled, to decouple the effects of long-term addiction vs. acute drug effects. Computational modeling can also be used to manipulate individual parameters under even tighter control than is possible in animal studies. Such results could provide useful complements to human studies, potentially providing converging evidence from many directions.

Another limitation of the current study is that the control group had significantly higher educational attainment than the patient group, with most control participants having completed at least a high school education, whereas many addicted participants did not. In fact, many other recent studies on drug addiction have also compared addicted patients to control groups with significantly more years of education (e.g., 11, 43, 45, 46, 47), and some studies have even suggested an explicit link between low educational attainment and drug addiction (e.g., 48, 49, 50). At least one prior large-scale longitudinal study has specifically assessed non-completion of grade 12 (secondary school) in Australia, and shown that polydrug- or alcohol-using adolescents were less likely to complete school than non-drug users (51). Of course correlation does not allow us to infer causation: low education could promote risk for drug abuse, early exposure to drugs could contribute to low educational attainment, and/or third variables (such as high cognitive ability or childhood socioeconomic status) could increase likelihood of both high educational attainment and lower rates of initiation of drug use and/or higher rates of cessation. Because of the group imbalance, education was included as a covariate in the current analyses, and there was no indication of an effect of education level on any performance measure. However, future studies could attempt to recruit and test a larger group of putatively healthy adults with fewer than 7 years of formal education, or a group of highly-educated individuals with opioid-addiction. In addition, educational attainment is often a marker for general intellectual ability, and future studies could include tests of general cognitive function, rather than simply collecting information about years of education.

Similarly, we relied on subject self-report regarding current and past use of addictive drugs (other than prior heroin use and current methadone use which was confirmed in the patient group). Given the high rates of polydrug use among heroin-addicted individuals, future studies could include urine screening for current drug use including but not limited to heroin, and the same screening could also be conducted in the control group.

Finally, we focused in this study on opioid-addicted individuals who specifically had a history of heroin addiction and who were currently receiving methadone maintenance. Given the large population of individuals addicted to other opioid drugs (e.g. prescription painkillers), it would be of considerable interest to examine the same task in subjects addicted to other opioid drugs, to determine the degree to which the current findings are specific to heroin and/or methadone, or apply more broadly to opioid drugs.

Some of the most important open questions involve the specific ways in which brain substrates of addiction contribute to behavior that supports and maintains opioid use, addiction, and relapse. In addition to pre-existing individual differences in brain structure and function that might confer risk of addiction in individuals who are exposed to opioid drugs, drug exposure itself affects the brain and brain plasticity, altering cognition in ways that may help sustain drug use. Beyond this, therapeutic interventions (including both psychotherapy and pharmacotherapy, as in methadone maintenance treatments) can themselves alter the brain. To this end, some of the most important future directions in addiction research are to better understand the link between brain and behavior. The current study attempts to shed some light by examining behavior of opioid-addicted individuals on tests of learning and generalization that have previously been associated with specific brain substrates known to be implicated in addiction and addiction vulnerability (i.e. ventrostriatal dopamine system for reward- and punishment-based learning, medial temporal system for generalization). However, future studies could use functional imaging (e.g. fMRI) to investigate whether the abnormal learning observed here in opioid addicts reflects abnormal activity in these brain substrates. In fact, a prior functional imaging study in cocaine-addicted individuals found altered activity in the nucleus accumbens during associative learning task performance (52). To our knowledge, similar studies have not yet been conducted with patients with heroin addiction, and so this remains an important avenue for future work.

Acknowledgments

This work was partially supported by Merit Review Award #101 000771 from the U.S. Department of Veterans Affairs Clinical Sciences Research and Development Service. The views expressed herein are those of the authors and do not necessarily reflect the views of the Department of Veterans Affairs, the United States government, or any institution with which the authors are affiliated.

References

1. Hyman SE. Addiction: A disease of learning and memory. *Am J Psychiatry*. 2005; 162(8):1414–1122. [PubMed: 16055762]
2. Alkon DL, Amaral DG, Bear MF, Black J, Carew TJ, Cohen NJ, et al. Learning and Memory. *Brain Res Rev*. 1991; 16:193–220. [PubMed: 1760656]

3. National Institute on Drug Abuse. [on February 11, 2015] Prescription and over-the-counter medications. 2014. Retrieved from <http://www.drugabuse.gov/publications/drugfacts/prescription-over-counter-medications>
4. Smyth BP, Barry J, Keenan E, Ducray K. Lapse and relapse following inpatient treatment of opiate dependence. *Ir Med J.* 2010; 103(6):176–179. [PubMed: 20669601]
5. Bentzley BS, Barth KS, Back SE, Book SW. Discontinuation of buprenorphine maintenance therapy: Perspectives and outcomes. *J Subst Abuse Treat.* 2015; 52:48–57. [PubMed: 25601365]
6. Wise RA. Opiate reward: Sites and substrates. *Neurosci Biobehav Rev.* 1989; 13(2–3):129–133. [PubMed: 2573023]
7. Bozarth MA. Opiate reinforcement processes: re-assembling multiple mechanisms. *Addiction.* 1994; 89(11):1425–1434. [PubMed: 7841852]
8. Di Chiara G, North RA. Neurobiology of opiate abuse. *Trends Pharmacol Sci.* 1992; 13(5):185–193. [PubMed: 1604711]
9. Cheng J-T, Lu Y, Han X, González-Vallejo C, Sui N. Temporal discounting in heroin-dependent patients: no sign effect, weaker magnitude effect, and the relationship with inhibitory control. *Experimental and Clinical Psychopharmacology.* 2012; 20(5):400–409. [PubMed: 22889040]
10. Kirby KN, Petry NM. Heroin and cocaine abusers have higher discount rates for delayed rewards than alcoholics or non-drug-using controls. *Addiction.* 2004; 99(4):461–471. [PubMed: 15049746]
11. Myers CE, Sheynin J, Baldson T, Luzardo A, Beck KD, Hogarth L, et al. Probabilistic reward- and punishment-based learning in opioid addiction: Experimental and computational data. *Behav Brain Res.* 2016; 296:240–248. PubMed PMID: 26381438 PubMed Central PMCID: PMC4734141. [PubMed: 26381438]
12. Bonardi C, Rey V, Richmond M, Hall G. Acquired equivalence of cues in pigeon autoshaping: Effects of training with common consequences and common antecedents. *Anim Learn Behav.* 1993; 21(4):369–376.
13. Grice G, Davis J. Effect of concurrent responses on the evocation and generalization of the conditioned eyeblink. *J Exp Psychol.* 1960; 59(6):391–395. [PubMed: 13829256]
14. Hall G, Ray E, Bonardi C. Acquired equivalence between cues trained with a common antecedent. *Journal of Experimental Psychology: Animal Behavior Processes.* 1993; 19(4):391–399. [PubMed: 8228835]
15. Myers CE, Shohamy D, Gluck M, Grossman S, Kluger A, Ferris S, et al. Dissociating hippocampal vs. basal ganglia contributions to learning and transfer. *J Cognit Neurosci.* 2003; 15(2):185–193. PubMed PMID: 12676056. [PubMed: 12676056]
16. Moustafa AA, Kéri S, Herzallah MM, Myers CE, Gluck MA. A neural model of hippocampal-striatal interactions in associative learning and transfer generalization in various neurological and psychiatric patients. *Brain Cogn.* 2010; 74(2):132–144. PubMed PMID: 20728258; PubMed Central PMCID: PMC2936107. [PubMed: 20728258]
17. Herzallah MM, Moustafa AA, Misk AJ, Al-Dweib LH, Abdelrazeq SA, Myers CE, et al. Depression impairs learning whereas anticholinergics impair transfer generalization in Parkinson patients tested on dopaminergic medications. *Cognitive and Behavioral Neurology.* 2010; 23(2): 98–105. PubMed PMID: 20535058. [PubMed: 20535058]
18. Myers CE, Hopkins R, DeLuca J, Moore N, Wolansky L, Sumner J, et al. Learning and generalization deficits in patients with memory impairments due to anterior communicating artery aneurysm rupture or hypoxic brain injury. *Neuropsychology.* 2008; 22(5):681–686. PubMed PMID: 18763887. [PubMed: 18763887]
19. Bódi N, Csibri E, Myers CE, Gluck MA, Kéri S. Associative learning, acquired equivalence, and flexible generalization of knowledge in mild Alzheimer disease. *Cognitive and Behavioral Neurology.* 2009; 22(2):89–94. PubMed PMID: 19506424. [PubMed: 19506424]
20. Kéri S, Nagy O, Kelemen O, Myers CE, Gluck M. Dissociation between medial temporal lobe and basal ganglia memory systems in schizophrenia. *Schizophr Res.* 2005; 77:321–328. PubMed PMID: 15893916. [PubMed: 15893916]
21. Shohamy D, Mihalakos P, Chin R, Thomas B, Wagner AD, Tamminga C. Learning and generalization in schizophrenia: Effects of disease and antipsychotic drug treatment. *Biol Psychiatry.* 2009; 67(10):926–932. [PubMed: 20034612]

22. Kostek JA, Beck KD, Gilbertson MW, Orr SP, Pang KCH, Servatius RJ, et al. Acquired equivalence in U. S. veterans with symptoms of post-traumatic stress: reexperiencing symptoms are associated with better generalization. *J Traum Stress*. 2014; 27(6):717–720. PubMed PMID: 25470729 PubMed Central PMCID: PMC4272630.
23. Shohamy D, Wagner AD. Integrating memories in the human brain: Hippocampal-midbrain encoding of overlapping events. *Neuron*. 2008; 60:378–389. [PubMed: 18957228]
24. Cohen, N.; Eichenbaum, H. *Memory, Amnesia and the Hippocampal System*. Cambridge, MA: MIT Press; 1993.
25. Gluck M, Myers C. Hippocampal mediation of stimulus representation: A computational theory. *Hippocampus*. 1993; 3:491–516. PubMed PMID: 8269040. [PubMed: 8269040]
26. Montague P, Dayan P, Sejnowski T. A framework for mesencephalic dopamine systems based on predictive Hebbian learning. *J Neurosci*. 1996; 16(5):1936–1947. [PubMed: 8774460]
27. Schultz W. Predictive reward signal of dopamine neurons. *J Neurophysiol*. 2002; 80:1–27.
28. Schultz W, Dayan P, Montague P. A neural substrate of prediction and reward. *Science*. 1997; 275:1593–1599. [PubMed: 9054347]
29. Mattfeld AT, Gluck MA, Stark CEL. Functional specialization within the striatum along both the dorsal-ventral and anterior/posterior axes during associative learning via reward and punishment. *Learn Mem*. 2011; 18:703–711. [PubMed: 22021252]
30. Palminteri S, Justo D, Jauffret C, Pavlicek B, Dauta A, Delmaire C, et al. Critical roles for anterior insula and dorsal striatum in punishment-based avoidance learning. *Neuron*. 2012; 76(5):998–1009. [PubMed: 23217747]
31. Kreek MJ, Levran O, Reed B, Schlussman SD, Zhou Y, Butelman ER. Opiate addiction and cocaine addiction: Underlying molecular neurobiology and genetics. *The Journal of Clinical Investigation*. 2012; 12(10):3387–3393.
32. Jay TM. Dopamine: A potential substrate for synaptic plasticity and memory mechanisms. *Prog Neurobiol*. 2003; 69:375–390. [PubMed: 12880632]
33. Gardner EL. Addiction and brain reward and antireward pathways. *Adv Psychosom Med*. 2011; 30:22–60. [PubMed: 21508625]
34. Vadhan N, Myers CE, Rubin E, Shohamy D, Foltin R, Gluck M. Stimulus-response learning in long-term cocaine users: Acquired equivalence and probabilistic category learning. *Drug Alcohol Depend*. 2008; 93(1–2):155–162. PubMed PMID: 24188172. [PubMed: 17976927]
35. Mátyáássy A, Kéri S, Myers CE, Levy-Gigi E, Gluck MA, Kelemen O. Impaired generalization of associative learning in patients with alcohol dependence after intermediate-term abstinence. *Alcohol Alcohol*. 2012; 47(5):533–537. PubMed PMID: 22582184. [PubMed: 22582184]
36. Herzallah MM, Moustafa AA, Natsheh JY, Danoun OA, Simon JR, Tayem YI, et al. Depression impairs learning, whereas the selective serotonin reuptake inhibitor, paroxetine, impairs generalization in patients with major depressive disorder. *J Affect Disord*. 2013; 151(2):484–492. PubMed PMID: 23953023; PubMed Central PMCID: PMC3797256. [PubMed: 23953023]
37. Meeter M, Shohamy D, Myers CE. Acquired equivalence changes stimulus representations. *Journal of Experimental Analysis of Behavior*. 2009; 91(1):127–141. PubMed PMID: 19230516; PubMed Central PMCID: PMC2614814.
38. Joel D, Niv Y, Ruppin E. Actor-critic models of the basal ganglia: New anatomical and computational perspectives. *Neural Networks*. 2002; 15(4–6):535–547. [PubMed: 12371510]
39. O'Doherty J, Dayan P, Schultz J, Deichmann R, Fristen K, Dolan R. Dissociable roles of ventral and dorsal striatum in instrumental conditioning. *Science*. 2004; 304(5669):452–455. [PubMed: 15087550]
40. Piray P, Zeighami Y, Bahrami F, Eissa AM, Hewedi DH, Moustafa AA. Impulse control disorders in Parkinson's disease are associated with dysfunction in stimulus valuation but not action valuation. *J Neurosci*. 2014; 34(23):7814–7824. [PubMed: 24899705]
41. Collie A, Myers C, Schnirman G, Wood S, Maruff P. Selectively impaired associative learning in older people with cognitive decline. *J Cognit Neurosci*. 2002; 14(3):484–492. PubMed PMID: 11970807. [PubMed: 11970807]
42. Zahari Z, Lee CS, Ibrahim MA, Musa N, Mohd Yasin MA, Lee YY, et al. Comparison of pain tolerance between opioid dependent patients on methadone maintenance therapy (MMT) and

- opioid naive individuals. *Journal of Pharmacy and Pharmaceutical Sciences*. 2016; 19(1):127–136. [PubMed: 27096697]
43. Sheynin J, Moustafa AA, Beck KD, Servatius RJ, Hogarth L, Haber P, et al. Exaggerated acquisition and resistance to extinction of avoidance behavior in treated heroin-dependent men. *J Clin Psychiatr*. 2016; 77:386–394. PubMed PMID: 27046310; PubMed Central PMCID: PMC4822714.
 44. Wang PW, Lin HC, Liu GC, Yang YC, Ko CH, Yen CF. Abnormal interhemispheric resting state functional connectivity of the insula in heroin users under methadone maintenance treatment. *Psychiatry Res*. 2016; 255:9–14. [PubMed: 27497215]
 45. Volkow ND, Wang G-J, Ma Y, Fowler JS, Wong C, Ding Y-S, et al. Activation of orbital and medial prefrontal cortex by methylphenidate in cocaine-addicted subjects but not in controls: Relevance to addiction. *J Neurosci*. 2005; 25(15):3932–3939. [PubMed: 15829645]
 46. Yip SW, DeVito EE, Kober H, Worhunsky PD, Carroll KM, Potenza MN. Anticipatory reward processing among cocaine-dependent individuals with and without concurrent methadone-maintenance treatment: Relationship to treatment response. *Drug Alcohol Depend*. 2016; 166:134–142. [PubMed: 27430401]
 47. Pennington DL, Durazzo TC, Schmidt TP, Abé C, Mon A, Meyerhoff DJ. Alcohol use disorder with and without stimulant use: Brain morphometry and its associations with cigarette smoking, cognition, and inhibitory control. *PLOS ONE*. 2015; 10(3):e0122505. [PubMed: 25803861]
 48. Weinberger AH, Pilver CE, Mazure CM, McKee SA. Stability of smoking status in the US population: A longitudinal investigation. *Addict Biol*. 2014; 109(9):1541–1553.
 49. Cai L, Cui W, You D, He J, Zhao K. Socioeconomic variations in nicotine dependence in rural southwest China. *BMC Public Health*. 2015; 15:1158. [PubMed: 26597724]
 50. van Oers JA, Bongers IM, van de Goor LA, Garretsen HE. Alcohol consumption, alcohol-related problems, problem drinking, and socioeconomic status. *Alcohol Alcohol*. 1999; 34(1):78–88. [PubMed: 10075406]
 51. Kelly AB, Evans-Whipp TJ, Smith R, Chan GC, Toumbourou JW, Patton GC, et al. A longitudinal study of the association of adolescent polydrug use, alcohol use and high school non-completers. *Addiction*. 2015; 110(4):627–635. [PubMed: 25510264]
 52. Volkow ND, Wang GJ, Telang F, Fowler JS, Logan J, Childress AR, et al. Cocaine cues and dopamine in dorsal striatum: Mechanism of craving in cocaine addiction. *J Neurosci*. 2006; 26(24):6583–6588. [PubMed: 16775146]

Figure 1.
The four cartoon characters that served as antecedents and the four food items that served as consequents.

Figure 2.

Example screen events on the computer-based task. (A) On each trial, the screen shows one antecedent and two consequents, and asks the subject to guess which food that person prefers. (B) On reward-based trials, correct responses are reinforced with feedback and point gain, while (C) incorrect responses trigger no feedback or points. (D) On punishment-based trials, incorrect responses trigger corrective feedback (“Incorrect – Lose 10 points!”) and point loss, while correct responses trigger no feedback or point loss.

Figure 3.

Performance on the training phase in terms of (A) mean trials at each training stage, and (B) mean total errors on reward-based and punishment-based trials, across the three stages.

Asterisks indicate significant differences between opioid-addicted and control groups ($p < .05$ in A, $p < .008$ in B).

Figure 4. Performance on the testing phase, including (A) retention of old (previously-trained) pairs and (B) generalization to new pairs, plotted separately for reward-based and punishment-based trials.

Table 1

Schematic of the acquired equivalence paradigm (15). On each trial, an antecedent stimulus (A, B, C, or D) is presented with a pair of consequents (X1 vs. Y1, or X2 vs. Y2), and the subject selects the correct consequent for that antecedent. Note that no feedback is provided during the testing phase. Antecedents A and C (+) are used for “reward-based” trials, meaning that correct responses are rewarded with positive feedback and point gain while incorrect responses trigger no feedback or point gain; antecedents B and D (–) are used for “punishment-based” trials, meaning that incorrect responses trigger negative feedback and point loss while correct responses trigger no feedback or point loss. The “correct” response (consistent with acquired equivalence) for each test trial is underlined.

Training Phase			Testing Phase
Stage 1	Stage 2	Stage 3	
A → X1 (+)	A → X1 (+) C → X1 (+)	A → X1 (+) C → X1 (+) A → X2 (+)	All trained pairs, plus: C → <u>X2</u> or Y2?
B → Y1 (–)	B → Y1 (–) D → Y1 (–)	B → Y1 (–) D → Y1 (–) B → Y2(–)	
			D → X2 or <u>Y2</u> ?

Table 2

Demographic information for the opioid-addicted and control groups; SD in parenthesis.

	Opioid-Addicted	Controls
N	33	32
Gender distribution	18 males / 15 females	22 males / 10 females
Mean age (in years)	39.3 (8.2)	35.0 (12.1)
Educational level (in years)	9.4 (1.6)	13.0 (1.6) *
Mean methadone dose (in mg)	75.5 (33.5)	N/A
Mean admission time to clinic (in years)	4.6 (3.6)	N/A
Mean duration of addiction to heroin (in years)	14.9 (3.7)	N/A

Asterisk indicates significant group difference, $p < .05$.