**Relapse Phenomena from Appetitive-to-Aversive and Aversive-to-Appetitive
Counterconditioning within a Human Predictive Learning Task**

by

R. Alexander Sauer

A thesis submitted to the Graduate Faculty of
Auburn University
in partial fulfillment of the
requirements for the Degree of
Master of Science

Auburn, Alabama
August 1st, 2015

Keywords: counterconditioning, predictive learning,
renewal, spontaneous recovery

Approved by

Martha Escobar, Associate Professor of Psychology
Chris Newland, Alumni Professor of Psychology
Sacha Pence, Assistant Professor of Psychology

Abstract

Counterconditioning (*CC*, Sherrington, 1947) involves training of a single cue with two different outcomes (i.e., cue-O1 then cue-O2) of opposite valence (e.g., appetitive and aversive). Thus, CC can be conducted in either the aversive-to-appetitive (cue-aversive O1 then cue-appetitive O2) or appetitive-to-aversive (cue-appetitive O1 then cue-aversive O2) direction. While CC effectively changes behavior elicited by the cue, this change is highly susceptible to relapse after a retention interval (i.e., spontaneous recovery) or a change of physical context (i.e., renewal). We compared relapse from appetitive-to-aversive and aversive-to-appetitive CC using a human predictive learning task. Participants made wagers of fictitious money based on the presentation of different cues and subsequently gained (appetitive outcome) or lost (aversive outcome) the value of their wager. In Experiment 1, we successfully demonstrated both appetitive-to-aversive and aversive-to-appetitive CC by pairing the target cue (*T*) with the loss (Loss-Gain condition) or gain (Gain-Loss condition) outcome during the original training phase and later pairing T with the gain (Loss-Gain condition) or loss (Gain-Loss condition) outcome during the CC phase. In Experiment 2, we observed spontaneous recovery from aversive-to-appetitive CC after both 7 day and 5min retention intervals, but recovery from appetitive-to-aversive CC was observed only after the 7day interval. In Experiment 3, we observed renewal only in the aversive-to-appetitive direction. After appetitive-to-aversive CC, responding at test was consistent with CC training regardless of the test context. Therefore, it appears that relapse is more likely from aversive-to-appetitive CC than appetitive-to-aversive CC. We discuss possible explanations to account for this difference.

Acknowledgements

Foremost, I want to express my sincerest gratitude to my advisor, Martha Escobar, for her guidance and support throughout this process. Because of her, I have grown as a researcher, a writer, and an academic. She is an amazingly knowledgeable, patient, and personable mentor, and I am so blessed to have the opportunity to work along-side her. I would also like to show my appreciation to my thesis committee members, Chris Newland and Sacha Pence, for their support of this project and their insightful questions, comments, and suggestions. Thank you all for everything you have done to make this project possible.

Additionally, I want to thank my laboratory colleges for all that they have done to help with this project. Tyler Toledo and Harrison Brown assisted in collecting data for these experiments and are phenomenal research assistants. Elizabeth Powell Dunaway, the senior graduate student in our laboratory, answered so many of my questions along the way and provided excellent feedback on my presentations. I am so grateful to have worked with these individuals.

Finally, I want to express my appreciation to my friends and family for their endless support throughout my life. My significant other, Rebecca Bresler, has been a daily source of encouragement during my journey as a graduate student. She is always willing to help me whenever possible, and I am so happy that she is a part of my life. My parents have continually pushed me to excel in academics, and I would not be in graduate school were it not for their dedication to my education early in life. They are a constant source of love, encouragement, and advice, and I am eternally grateful for all that they have done for me.

Table of Contents

List of Tables

List of Figures

Chapter 1

An organism must respond adaptively to its environment in order to survive. One way

this is accomplished is through classical conditioning, in which a conditioned stimulus (CS)

comes to elicit a conditioned response (CR) after multiple pairings with an unconditioned

stimulus (US; Pavlov, 1927; Rescorla & Wagner, 1972). This CR allows the organism to either

prepare for impending danger in the case of an aversive US or facilitate obtaining an appetitive

US. However, environmental conditions tend to change, requiring the organism to modify

existing CS-US relationships to accommodate new contingencies. For example, after a CS-US

relationship is formed the organism may experience the same CS paired with a different US.

This type of experience is known as *counterconditioning* (CC, *cf.* Sherrington, 1947).

In the laboratory, CC procedures occur across two consecutive training phases. During

the original learning phase, the CS is paired with a US (US1), resulting in a CR that reflects the

CS-US1 association. During the CC phase the CS is now paired with a different US (US2) rather

than US1, resulting in a new CR that reflects the CS-US2 association. Traditionally, US2 is of

an opposing valence to that of US1, such that behavior reflecting the CS-US1 association is

incompatible with behavior reflecting the CS-US2 association (see Dickinson & Pearce, 1977).

For example, many CC experiments in the animal literature use the presentation of food and the

onset of a mild foot shock as two opposing USs (e.g., Bouton & Peck, 1992; Nasser & McNally,

2012; Peck & Bouton, 1990; Powell, Escobar, & Kimble, 2013, Experiment 3). In such

experiments, foot shock is considered an aversive US, and the presentation of food is considered

an appetitive US. Depending on the valence of US2 as compared to that of US1, CC training can

occur in either the appetitive-to-aversive or aversive-to-appetitive direction. For appetitive-to-

aversive CC, US1 is appetitive and US2 is aversive (e.g., Bouton & Peck, 1992, Experiment 2; Nasser & McNally, 2012; Peck & Bouton, 1990, Experiments 2 and 3). For aversive-to-appetitive CC, US1 is aversive and US2 is appetitive (e.g., Bouton & Peck, 1992, Experiment 1; Peck & Bouton, 1990, Experiment 1; Powell et al., 2013, Experiment 3). If CC in either direction is successful, behavior elicited by the CS will change to reflect an anticipation of US2 rather than US1, resulting in a response that is more consistent with the current conditions of the organism's environment.

CC as an Interference Paradigm

CC is usually considered an *interference paradigm* (Bouton, 1993, 2004; Miller & Escobar, 2002) because, like other inference paradigms (e.g., latent inhibition, extinction), CC procedures result in a CS being associated with multiple outcomes across consecutive training phases. These multiple outcome associations are thought to interfere with one another because behavior consistent with one outcome (e.g., US1) is disrupted due to the association with the other outcome (e.g., US2). Interference can occur proactively when the CS-US1 association interferes with performance of behavior related to the CS-US2 association. Alternatively, interference can occur retroactively when the CS-US2 association interferes with performance of CS-US1 behavior. For most interference paradigms, both proactive and retroactive interference occur simultaneously. Proactive interference is seen primarily when the organism is slow to acquire the CS-US2 association after having acquired the CS-US1 association. However, after multiple pairings of the CS with US2, retroactive interference is evident because behavior begins to reflect the CS-US2 association.

The occurrence of interference in the case of CC seems to be related to the two outcomes having opposite motivational or affective valences. One possibility is that associations related to

appetitive and aversive USs activate different but mutually dependent motivational systems (e.g., Dickinson & Dearing, 1979; Dickinson & Pearce, 1977). According to this theory of opposing motivational systems, when either the appetitive or aversive motivational system becomes activated by the presentation of a US or CS, this activation simultaneously inhibits the other motivational system. Even from the first description of CC (Sherrington, 1947), the expectation is that brain areas associated with the appetitive and aversive responses will reciprocally inhibit each other. Indeed, neurons in the ventral tegmental area and the lateral habenula appear to provide physiological support for this view (Matsumoto & Hikosaka, 2009a, 2009b). In monkeys trained with distinct visual CSs that were differentially associated with apple juice (appetitive US) or an air puff (aversive US), presentations of the CS associated with the appetitive US inhibited neurons that fired during presentations of the CS associated with the aversive US, and vice-versa. Application of these findings to CC lead to the expectation that both aversive-to-appetitive and appetitive-to-aversive CC should result in proactive interference during acquisition of the CC association. Furthermore, retroactive interference should be observed after CC has been completed, because activation of one system will necessarily change behavior controlled by the other system.

Most applications of either the appetitive-to-aversive and aversive-to-appetitive CC procedures have yielded clear evidence of retroactive interference because, after completion of CC training, behavior towards the CS reflects the CS-US2 association rather than the CS-US1 association (e.g., Bouton & Peck, 1992; Engelhard, Leer, Lange, & Olatunji, 2014; Peck & Bouton, 1990; Powell et al., 2013; Thomas, Cutler, & Novak, 2012; Tunstall, Verendeev, & Kearns, 2012; Van Gucht, Baeyens, Vansteenwegen, Hermans, & Beckers, 2010). Nonetheless, proactive interference (i.e., retarded acquisition of the CC association) has been a less reliable

finding, and its observation may depend on the response system engaged during original learning. Aversive-to-appetitive procedures consistently show that original aversive learning interferes with subsequent acquisition of the appetitive CS-US2 association in the CC phase (Bouton & Peck, 1992; Bromage & Scavio, 1978; Kaye, Preston, Szabo, Druiff, & Mackintosh, 1987; Krank, 1985; Peck & Bouton, 1990; Scavio, 1974). In contrast, although a few studies have provided evidence of proactive interference in appetitive-to-aversive CC (Bouton & Peck, 1992; Dickinson, 1976, Experiment 1; Nasser & McNally, 2012, 2013), others have shown that original training with appetitive CS-US presentations actually *facilitates* aversive conditioning during the CC phase (DeVito & Fowler, 1982; Dickinson, 1976, Experiment 2; Scavio & Gormenzano, 1980). Furthermore, other reports suggest that prior appetitive conditioning has no noticeable effect on the rate of acquisition of subsequent aversive conditioning (Jackson, 1974; Peck & Bouton, 1990). The lack of consistency in observing proactive interference in appetitive-to-aversive preparations may indicate that the presumed aversive and appetitive motivational systems do not have such a direct effect on one another as the theory suggests.

Recent investigations of appetitive-to-aversive procedures have more reliably demonstrated proactive interference. Nasser and McNally (2012) used a procedure in which proactive interference had previously been demonstrated in rats (*cf.* Peck & Bouton, 1990). In their study, an auditory stimulus served as the CS, the presentation of food served as the appetitive US, and a mild foot shock served as the aversive US. Rats receiving appetitive conditioning during original learning were slower to acquire the fear response (freezing) during aversive conditioning in the CC phase than rats that did not receive appetitive conditioning. This proactive interference effect was fairly robust and resistant to several different manipulations,

4

such as reducing the CS-appetitive US contingency, changing the context between original learning and CC, and increasing or decreasing the magnitude of the appetitive US.

In a follow up study, Nasser and McNally (2013) investigated the neural correlates of appetitive-to-aversive CC by measuring the expression of phosphorylated mitogen activated protein kinase (pMAPK) as an indicator of plasticity in areas of the brain known to be associated with fear conditioning. Rats experiencing appetitive-to-aversive CC showed higher levels of pMAPK in the lateral amygdala, nucleus accumbens, and rostral agranular insular cortex compared to rats that only experienced appetitive or aversive conditioning. Given the relationship between the expression of pMAPK and neural changes in the brain, Nasser and McNally interpreted this increase in pMAPK as suggestive that rats in the CC group suffered from a larger prediction error about the outcome of the CS. That is, a greater change was required in these brain areas in order to accommodate for the unexpected presentation of shock after the CS, which could be related to the retarded acquisition of the CS-fear association in rats that received CC. While these pMAPK data do not necessarily demonstrate two interdependent and opposing motivational systems, they do show that a previous history of appetitive conditioning with a CS negatively alters subsequent fear learning about that CS, a notion that is in line with an associative opponent motivational account of the CC process (e.g., Dickinson & Dearing, 1979; Dickinson & Pearce, 1977).

Regardless of whether CC involves the mutual inhibition of different appetitive and aversive motivational systems, it is clear that both the appetitive-to-aversive and the aversive-to-appetitive procedures involve retroactive interference and at least some degree of proactive interference. The associative opponent account (e.g., Dickinson & Dearing, 1979; Dickinson & Pearce, 1977) and other associative accounts (e.g., Rescorla & Wagner, 1972) of interference

paradigms suggest the CS-US1 association is "unlearned" so that the CS-US2 association can be acquired. Although these explanations account for the change in behavior seen at the end of CC training, they cannot account for any return (or relapse) of behavior related to original training.

Relapse Phenomena from CC

CC (like other interference paradigms) is effective in changing behavior elicited by a CS; however, this procedure is also highly susceptible to relapse (e.g., Bouton, 1993, 2002, 2004; Powell et al., 2013). Relapse can be defined as a situation in which behavior consistent with the original association returns after completion of the CC phase. Such a return of behavior related to the original learning phase can occur under a variety of different conditions, giving rise to several different types of relapse phenomena. *Spontaneous recovery* (Pavlov, 1927) occurs when behavior consistent with the original association is again observed when time elapses between CC training and subsequent presentations of the CS in a testing phase (e.g., Bouton & Peck, 1992; Powell et al., 2013). *Renewal* (Bouton & Bolles, 1979) is a similar relapse phenomenon; however, it involves a physical context change from the original to the CC training phase rather than a temporal delay (e.g., Peck & Bouton, 1990; Thomas et al., 2012). 'Physical context' in these situations refers to the physical characteristics of the environment in which training takes place. While both spontaneous recovery and renewal depend on different variables (e.g., time and physical context), both of these relapse phenomena provide support to the view that the behavioral change resulting from CC (and other interference paradigms) is not permanent.

Many theories of associative learning (e.g., Dickinson & Dearing, 1979; Dickinson & Pearce, 1977; Rescorla & Wagner, 1972) cannot account for either spontaneous recovery or renewal, because they suggest that the CC association weakens or even erases the original association. However, if this were the case, then one would expect to see behavior reflecting the

CC association whenever the CS is presented rather than a return of the original behavior. An alternative to these earlier associative learning theories that much better accounts for relapse is provided by retrieval-based theories. For example, Bouton (1993) suggested that the CC association does not replace the original association; instead, each phase of training results in distinct associations to the same CS. Presumably, information about the CS, the outcome, and the physical and temporal contexts is stored as part of the memory for each training phase. CC training (or any other interfering training) makes the CS become an ambiguous predictor of two different outcomes; thus, at test, the organism must rely on contextual information to disambiguate the cue. If the context at test differs (either physically or temporally) from that of the recent CC training, the organism will fail to retrieve the CC association, resulting in the return of the originally-trained behavior. In this framework, spontaneous recovery can be attributed to a change in temporal context, whereas renewal can be attributed to a change in physical context.

Several experiments have investigated relapse phenomena in CC within the retrieval theory framework. Bouton and Peck (1992) investigated spontaneous recovery after both aversive-to-appetitive (Experiment 1) and appetitive-to-aversive (Experiment 2) CC. This procedure used shock as the aversive US, food as the appetitive US, and a tone as the CS. Subjects in the control condition received training with only one association, which matched the CC training of the CC group. All subjects were subsequently tested either 1 or 28 days after training was completed. CC was successful, and by the end of CC training, the magnitude of the CR was equivalent for both the CC and control conditions. More importantly, the original response was again observed in the CC group after the 28day interval but not after the 1 day interval. More recently, Powell et al. (2013, Experiment 3) investigated spontaneous recovery

after CC in the aversive-to-appetitive direction, using shock as the aversive US, sucrose pellets as the appetitive US, and a white noise as the CS. Compared to control subjects who received non-contingent CS-aversive US pairings during the initial training phase, subjects in the CC group showed more fear to the CS after a 3 day retention interval. Results from both Bouton and Peck (1992) and Powell et al. (2013) are in line with a retrieval theory account of spontaneous recovery from CC.

Studies investigating renewal from CC have also provided support for Bouton's (1993) retrieval theory. Peck and Bouton (1990) conducted a series of experiments to study the effects of changing the physical context on both aversive-to-appetitive (Experiment 1) and appetitive-to-aversive (Experiments 2 and 3) CC. Shock was the aversive US, food was the appetitive US, and a tone served as the CS. Three renewal paradigms were assessed: ABA, AAB, and ABC (where the first, second, and third letters reflect the context of original training, CC training, and testing, respectively). In both appetitive-to-aversive and aversive-to-appetitive preparations, renewal was evident in the ABA groups, although it was not observed in the AAB and ABC groups. Thomas et al. (2012) also compared renewal after extinction (i.e., CS-noUS training), a traditional aversive-to-appetitive CC, and a modified CC procedure. The extinction group received presentations of the CS alone during the second phase of training. The traditional CC group received CS-appetitive US pairings, which were not contingent on the subjects' behavior. Finally, the modified CC group received CS-appetitive US pairings, but they were required to lever press in order to obtain the US. ABA renewal was observed in the extinction and traditional CC group, but it was attenuated in the modified CC group. Although the authors suggested that their modified CC procedure offers a method of altering a CS-US association that is highly resistant to contextual changes, these results should be taken with caution. Animals in

the modified CC group could only receive reinforcement during Phase 2 upon lever pressing when the CS was present; therefore, this manipulation could be viewed as having simply established the CS as a discriminative stimulus for the availability of reinforcement upon lever pressing. From this perspective it is not surprising that rats in the modified CC group lever pressed more than other groups during the CS, because they were the only group to have the delivery of the appetitive US directly contingent upon lever pressing during the CS.

Research with Humans Investigating CC

The majority of studies investigating CC with human participants have used evaluative conditioning procedures (e.g., Baeyens, Eelen, Van Den Bergh, & Crombez, 1989; Engelhard et al., 2014; Gast & De Houwer, 2013; Kerkhof, Vansteenwegen, Baeyens, & Hermans, 2011; Raes & De Raedt, 2012; Van Gucht, Baeyens, Hermans, & Beckers, 2013; Van Gucht et al., 2010). In such procedures, participants are first asked to rate a list of different stimuli on a scale from highly unpleasant (negative values) to highly pleasurable (positive values). After completing this process the experimenter chooses stimuli that received extreme positive and negative ratings and subsequently pairs these stimuli with neutrally rated stimuli (values close to zero). Thus, a neutral stimulus can be used as a CS and the extreme positive and negative stimuli can be used as appetitive and aversive USs, respectively. After multiple pairings of a positive or negative stimulus with a neutral stimulus, the participant is again asked to rate an array of stimuli in order to assess to what degree the participant's evaluations of the neutral stimuli have changed to reflect the positive or negative stimulus with which they were paired.

The evaluative conditioning procedure can easily be used to study CC by pairing a single neutral stimulus with both a positive and a negative stimulus between subsequent phases of the procedure. It also allows for the use of a wide variety of stimuli, including pictures (Baeyens et

al., 1989; Engelhard et al., 2014; Raes & De Raedt, 2012), other visual stimuli (Van Gucht et al., 2013, 2010), and food items (e.g., cookies; Kerkhof et al., 2011). Evaluative CC appears to be more effective at altering US expectancy, and results in greater changes in participants' evaluation of the stimulus than extinction (Engelhard et al., 2014; Kerkhof et al., 2011; Raes & De Raedt, 2012; Van Gucht et al., 2013, 2010). Engelhard et al. (2014) and Van Gucht et al. (2013) suggested that the effectiveness of CC may have implications for treatment of conditions normally treated with extinction-based procedures, such as anxiety disorders (by changing the stimulus' emotional valence) and substance abuse (by reducing the appetitiveness of the abused substance).

Relapse from CC due to spontaneous recovery and renewal in the evaluative conditioning literature has been reported in just a handful of studies, and the results have been somewhat inconsistent. Kerkhof et al. (2011) investigated spontaneous recovery after both appetitive-to-aversive and aversive-to-appetitive CC using a within subjects design. Participants returned to the laboratory after a 7day retention interval at which point CS evaluations and affective priming were reassessed. Spontaneous recovery was not observed for the evaluation of stimuli in either valence direction, but there was some recovery for affective priming of stimuli receiving appetitive-to-aversive training (i.e., there was a recovery of positive ratings toward the CS). Using a similar 7-day retention interval and an appetitive-to-aversive CC procedure with chocolate as the appetitive US, Polysorbate 20 (a bitter tasting liquid) as the aversive US, and different colored trays as CSs, Van Gucht et al. (2010) observed spontaneous recovery of outcome expectancy ratings but no spontaneous recovery of ratings of chocolate cravings (spontaneous recovery of evaluative ratings was not assessed). A subsequent study showed that ratings of outcome expectancy, but not of chocolate cravings, exhibited ABA renewal (Van

Gucht et al., 2013). Thus, whether or not relapse from CC in humans can be observed is not clear, as it appears to be dependent on the variables measured.

CC has also been investigated with humans using a predictive learning task in at least one study (Rosas, Paredes-Olay, García-Gutiérrez, Espinosa, & Abad, 2010, Experiment 3). A predictive learning task involves the presentation a stimulus (i.e., a CS or cue) followed by some outcome. After multiple pairings of different cues with different outcomes, participants are asked to make predictions about the future occurrence of an outcome based on the presence of one of the cues. Rosas et al. (2010, Experiment 3) used CC in a predictive learning task by pairing two distinct visual cues (i.e., a star and a circle) with two different outcomes (i.e., destroying a plane and destroying a ship) in the original learning phase. During the CC phase, the outcome previously associated with one of these stimuli was changed to the alternative outcome (i.e., destroyed planes to destroyed ships or vice versa). Rosas et al. found that behavior elicited by the cue that underwent CC training changed to reflect the outcome presented during CC rather than the original learning phase. Because the cue is presented with two different outcomes between consecutive phases, this is technically a CC procedure; however, it differs greatly from previously described CC procedures in that it does not utilize outcomes of opposing emotional valence. Destroying planes and destroying ships are distinct events but neither would likely activate the appetitive or aversive motivational systems described by Dickinson and Pearce (1977).

Applications of CC

The most widely known applications of both aversive-to-appetitive and appetitive-to-aversive CC procedures involve clinical situations. For example, in the *systematic desensitization* procedure (*cf.* Wolpe, 1958), a previously fear-eliciting stimulus is changed to

signal focused relaxation over the course of several sessions with a therapist. This technique has been adapted to the treatment of other anxiety conditions, such as posttraumatic stress disorder (PTSD), by first classically conditioning a CS (e.g., pressing reasonably hard on the patients finger) to elicit strong emotional responses that are incompatible with fear or anxiety related emotional responses (Paunovic, 1999, 2002, 2003). After classical conditioning, patients experience CC training by exposing them to anxiety evoking stimuli, and when feelings of distress arise, the incompatible emotional responses are elicited via the previously trained CS. After multiple sessions, patients' responses to the anxiety-evoking stimuli decrease significantly. Disgust and anxiety towards spiders in individuals with spider phobias can also be attenuated with CC (de Jong, Vorage, & van den Hout, 2000). Additionally, CC techniques have proven effective in a wider range of medical issues, such as to supplement the use of anesthetics in providing pain relief during subcutaneous injections (Slifer, Eischen, & Busby, 2002), and to train biofeedback techniques that help children prevent cue-induced allergic reactions (Khan, 1977).

Most instances of applications of appetitive-to-aversive CC have been to treat substance abuse or addiction disorders (e.g., Childress, Ehrman, McLellan, & O'Brien, 1988; Frawley & Smith, 1990; Van Gucht et al., 2013). In such situations, cues associated with the abused substance are paired with strongly aversive stimuli (i.e., noxious tastes or odors and emetics) in order to decrease the occurrence of cue-induced cravings for the substance. For example, Frawley and Smith (1990) used chemical *aversion therapy* to improve outcomes for cocaine dependent patients by administering an emetic prior to the presentation of different paraphernalia (e.g., lines of cocaine, a mirror, razor blades) and the action of snorting a cocaine powder

substitute. A similar approach could also be used to treat food addictions by altering food cravings, as Van Gucht et al. (2013, 2010) did with normal populations.

A great limitation to the applicability of CC procedures is that the basic human literature investigating relapse phenomena associated with these procedures is relatively limited and inconclusive. In order for the application of CC procedures to be truly useful, the long-term maintenance of the behavior change produced by these methods must be better understood. For example, although animal research suggests that both appetitive-to-aversive and aversive-to-appetitive CC procedures will result in some degree of renewal and spontaneous recovery (Bouton & Peck, 1992; Peck & Bouton, 1990), it is not clear from the human literature whether or not these procedures are susceptible to relapse (Kerkhof et al., 2011; Van Gucht et al., 2013, 2010). Unfortunately, direct comparisons between relapse from appetitive-to-aversive and aversive-to-appetitive CC are not available, and such information could be valuable for therapists when devising treatment programs.

The Present Study

The present experiments were designed to investigate and compare the susceptibility of both appetitive-to-aversive and aversive-to-appetitive CC to renewal and spontaneous recovery under equivalent training conditions. A predictive learning task was chosen for these experiments because, as previously noted, such a procedure has been successful in yielding CC in human participants (Rosas et al., 2010); however, the present study used outcomes that have opposing valences rather than simply two distinct events. Predictive learning methods provide an advantage over evaluative conditioning methods because predictive learning methods rely on discrete behaviors rather than ratings of liking, disgust, or anticipation. Additionally, we chose a

predictive learning method because, to the best of my knowledge, predictive learning tasks have not been used previously to investigate relapse from CC procedures.

Our predictive learning task involved participants playing a computer game in which different cues (fictitious stock companies) resulted in aversive or appetitive outcomes (gaining or losing money, respectively). Expectations of these outcomes were assessed in terms of wagers of fictitious money that participants made in the presence of each stock, which were operationalized as behavioral (key press) responses. One of these cues (the target cue, hereon *T*) was counterconditioned over the course of two consecutive phases. Appetitive-to-aversive CC (Gain-Loss Condition) was conducted by pairing *T* with gains during the original learning phase and loss during the CC phase. Aversive-to-appetitive CC (Loss-Gain Condition) was conducted by pairing *T* with loss during the original learning phase and gains during the CC phase.

Experiment 1 established the task as an effective method of conducting a CC procedure in appetitive-to-aversive and aversive-to-appetitive directions. Based on the prior animal and human research reviewed above, we expected that, upon completion of the original learning phase, behavior towards *T* would reflect an anticipation of loss (Loss-Gain Condition) or gains (Gain-Loss Condition), and, upon completion of the CC phase, behavior towards *T* would reflect an anticipation of gains (Loss-Gain Condition) or loss (Gain-Loss Condition).

Experiment 2 investigated spontaneous recovery utilizing the same basic procedure, except that a 7day retention interval was imposed between the last trial of the CC phase and testing. A 7day retention interval was chosen because prior CC experiments with human participants have had success obtaining spontaneous recovery after such a delay (Kerkhof et al., 2011; Van Gucht et al., 2010). We also included a 5min retention interval as a control condition in order to assess the effects of the retention interval when all other factors (e.g., interrupting the

task and returning to the laboratory) were held constant. Based on the prior animal literature (Bouton & Peck, 1992; Powell et al., 2013) we expected to obtain spontaneous recovery at the 7day interval after both training procedures; however, given the inconsistencies in the human literature, we could not make any assumptions about potential differences that would be observed between these procedures.

In Experiment 3, our predictive learning task was used to investigate renewal after these CC procedures. We created different contexts by altering the background color and different screen elements presented on the computer. Other predictive learning tasks have had success in obtaining contextual effects by altering elements of the computer screen (Nelson, Lamoureux, & León, 2013; Rosas & Callejas-Aguilera, 2006), but it was unclear whether or not the current contextual manipulation would result in similar effects. However, if our manipulation resulted in contextual effects, then we expected to observe renewal in both appetitive-to-aversive and aversive-to-appetitive procedures based on the prior animal research (Peck & Bouton, 1990).

References

Baeyens, F., Eelen, P., Van Den Bergh, O., & Crombez, G. (1989). Acquired affective-
evaluative value: Conservative but not unchangeable. *Behaviour Research and Therapy*,
27, 279–287.

Bouton, M. E. (1993). Context, time, and memory retrieval in the interference paradigms of
Pavlovian learning. *Psychological Bulletin*, *114*, 80–99.

Bouton, M. E. (2002). Context, ambiguity, and unlearning: Sources of relapse after behavioral
extinction four mechanisms of relapse. *Biological Psychiatry*, *52*, 976–986.

Bouton, M. E. (2004). Context and behavioral processes in extinction. *Learning & Memory*, *11*,
485–494.

Bouton, M. E., & Bolles, R. C. (1979). Contextual control of the extinction of conditioned fear.
Learning and Motivation, *10*, 445–466.

Bouton, M. E., & Peck, C. A. (1992). Spontaneous recovery in cross-motivational transfer
(counterconditioning). *Animal Learning & Behavior*, *20*, 313–321.

Bromage, B. K., & Scavio, M. J. (1978). Effects of an aversive CS+ and CS− under deprivation
upon successive classical appetitive and aversive conditioning. *Animal Learning &
Behavior*, *6*, 57–65.

Childress, A., Ehrman, R., McLellan, A., & O'Brien, C. (1988). Update on behavioral treatments
for substance abuse. *NIDA Research Monograph Series*, *90*, 183–192.

De Jong, P. J., Vorage, I., & van den Hout, M. A. (2000). Counterconditioning in the treatment
of spider phobia: Effects on disgust, fear and valence. *Behaviour Research and Therapy*,
38, 1055–1069.

DeVito, P. L., & Fowler, H. (1982). Transfer of conditioned appetitive stimuli to conditioned

    aversive excitatory and inhibitory stimuli. *Learning and Motivation*, *13*, 135–154.

Dickinson, A. (1976). Appetitive-aversion interactions: Facilitation of aversive conditioning by

    prior appetitive training in the rat. *Animal Learning & Behavior*, *4*, 416–420.

Dickinson, A., & Dearing, M. F. (1979). Appetitive-aversive interactions and inhibitory

    processes. In A. Dickinson & R. a. Boakes (Eds.), *Mechanisms of learning and

    motivation: A memorial volume to Jerzy Konorksi* (pp. 203–231). Totawa, NJ: Erlbaum.

Dickinson, A., & Pearce, J. M. (1977). Inhibitory interactions between appetitive and aversive

    stimuli. *Psychological Bulletin*, *84*, 690–711.

Engelhard, I. M., Leer, A., Lange, E., & Olatunji, B. O. (2014). Shaking that icky feeling: Effects

    of extinction and counterconditioning on disgust-related evaluative learning. *Behavior

    Therapy*, *45*, 708–719.

Frawley, P. J., & Smith, J. W. (1990). Chemical aversion therapy in the treatment of cocaine

    dependence as part of a multimodal treatment program: Treatment outcome. *Journal of

    Substance Abuse Treatment*, *7*, 21–29.

Gast, A., & De Houwer, J. (2013). The influence of extinction and counterconditioning

    instructions on evaluative conditioning effects. *Learning and Motivation*, *44*, 312–325.

Jackson, D. E. (1974). CS-free food contingencies and subsequent acquisition of conditioned

    suppression: No transfer effect. *Bulletin of Pyschonomic Society*, *4*, 235–236.

Kaye, H., Preston, G. C., Szabo, L., Druiff, H., & Mackintosh, N. J. (1987). Context specificity

    of conditioning and latent inhibition: Evidence for a dissociation of latent inhibition and

    associative interference. *The Quarterly Journal of Experimental Psychology*, *39*, 127–

    145.

Kerkhof, I., Vansteenwegen, D., Baeyens, F., & Hermans, D. (2011). Counterconditioning: An effective technique for changing conditioned preferences. *Experimental Psychology*, *58*, 31–38.

Khan, A. U. (1977). Effectiveness of biofeedback and counter-conditioning in the treatment of bronchial asthma. *Journal of Psychosomatic Research*, *21*, 97–104.

Krank, M. D. (1985). Asymmetrical effects of Pavlovian excitatory and inhibitory aversive transfer on Pavlovian appetitive responding and acquisition. *Learning and Motivation*, *16*, 35–62.

Matsumoto, M., & Hikosaka, O. (2009a). Representation of negative motivational value in the primate lateral habenula. *Nature Neuroscience*, *12*, 77–84.

Matsumoto, M., & Hikosaka, O. (2009b). Two types of dopamine neuron distinctly convey positive and negative motivational signals. *Nature*, *459*, 837–841.

Miller, R. R., & Escobar, M. (2002). Associative interference between cues and between outcomes presented together and presented apart: An integration. *Behavioural Processes*, *57*, 163–185. http://doi.org/10.1016/S0376-6357(02)00012-8

Nasser, H. M., & McNally, G. P. (2012). Appetitive–aversive interactions in Pavlovian fear conditioning. *Behavioral Neuroscience*, *126*, 404–422.

Nasser, H. M., & McNally, G. P. (2013). Neural correlates of appetitive − aversive interactions in Pavlovian fear conditioning, *20*, 220–228.

Nelson, J. B., Lamoureux, J. A., & León, S. P. (2013). Extinction arouses attention to the context in a behavioral suppression method with humans. *Journal of Experimental Psychology. Animal Behavior Processes*, *39*, 99–105.

Paunovic, N. (1999). Exposure counterconditioning (EC) as a treatment for severe PTSD and

    depression with an illustrative case. *Journal of Behavior Therapy and Experimental*

    Psychiatry, *30*, 105–117.

Paunovic, N. (2002). Prolonged exposure counterconditioning (PEC) as a treatement for chronic

    post-traumatic stress disorder and major depression in an adult survivor of repeated child

    sexual and physical abuse. *Clinical Case Studies*, 148–169.

Paunovic, N. (2003). Prolonged exposure counterconditioning as a treatment for chronic

    posttraumatic stress disorder. *Journal of Anxiety Disorders*, *17*, 479–499.

Pavlov, I. P. (1927). *Conditioned reflexes*. London, England: Oxford University Press.

Peck, C. A., & Bouton, M. E. (1990). Context and performance in aversive-to-appetitive and

    appetitive-to-aversive transfer. *Learning and Motivation*, *21*, 1–31.

Powell, E. J., Escobar, M., & Kimble, W. (2013). Delaying interference training has equivalent

    effects in various Pavlovian interference paradigms. *Learning & Memory*, *20*, 241–244.

Raes, A. K., & De Raedt, R. (2012). The effect of counterconditioning on evaluative responses

    and harm expectancy in a fear conditioning paradigm. *Behavior Therapy*, *43*, 757–767.

Rescorla, R. A., & Wagner, A. R. (1972). A theory of Pavlovian conditioning: Variations in the

    effectiveness of reinforcement and nonreinforcement. In A. H. Black & W. F. Prokasy

    (Eds.), *Classical Conditioning II: Current Theory and Research* (pp. 64–99). New York

    City, New York: Appleton-Century-Crofts.

Rosas, J. M., & Callejas-Aguilera, J. E. (2006). Context switch effects on acquisition and

    extinction in human predictive learning. *Journal of Experimental Psychology. Learning,*

    Memory, and Cognition, *32*(3), 461–74.

Rosas, J. M., Paredes-Olay, M. C., García-Gutiérrez, A., Espinosa, J. J., & Abad, M. J. F. (2010). Outcome-specific transfer between predictive and instrumental learning is unaffected by extinction but reversed by counterconditioning in human participants. *Learning and Motivation*, *41*, 48–66.

Scavio, M. J. (1974). Classical-classical transfer: Effects of prior aversive condtioning upon appetitive conditioning in rabbits. *Journal of Comparative and Physiological Psychology*, *86*, 107–115.

Scavio, M. J., & Gormenzano, I. (1980). Classical-classicla transfer: Effects of prior appetitive conditioning upon aversive conditioning in rabbits. *Animal Learning & Behavior*, *8*, 218–224.

Sherrington, C. S. (1947). *The integrative action of the central nervous system*. Cambrige, UK: Cambridge University Press.

Slifer, K. J., Eischen, S. E., & Busby, S. (2002). Using counterconditioning to treat behavioural distress during subcutaneous injections in a paediatric rehabilitation patient. *Brain Injury*, *16*, 901–916.

Thomas, B. L., Cutler, M., & Novak, C. (2012). A modified counterconditioning procedure prevents the renewal of conditioned fear in rats. *Learning and Motivation*, *43*, 24–34.

Tunstall, B. J., Verendeev, A., & Kearns, D. N. (2012). A comparison of therapies for the treatment of drug cues: Counterconditioning vs. extinction in male rats. *Experimental and Clinical Psychopharmacology*, *20*, 447–53.

Van Gucht, D., Baeyens, F., Hermans, D., & Beckers, T. (2013). The inertia of conditioned craving. Does context modulate the effect of counterconditioning? *Appetite*, *65*, 51–57.

Van Gucht, D., Baeyens, F., Vansteenwegen, D., Hermans, D., & Beckers, T. (2010).

  Counterconditioning reduces cue-induced craving and actual cue-elicited consumption.

  Emotion, *10*, 688–695.

Wolpe, J. (1958). *Psychotherapy by Reciprocal Inhibition*. Standford, CA: Stanford University

  Press.

Chapter 2

In predictive learning preparations, a neutral cue is paired with the occurrence of an outcome (O1). After repeated cue-O1 pairings, the cue becomes a reliable predictor of O1 and presenting the cue results in behavior consistent with an expectation of O1. Nonetheless, in some circumstances, the cue may receive subsequent training with a different outcome, thus degrading the cue-O1 contingency (Miller & Escobar, 2002). For example, if after cue-O1 pairings the cue is presented alone (i.e., cue-noO), the cue-O1 association may undergo *extinction* (*cf.* Pavlov, 1927), and behavior consistent with expectation of O1 in the presence of the cue is reduced. Cue-O1 pairings can also be followed by pairings of the cue and a second outcome (O2). If O1 and O2 represent events of opposite valence, behavior consistent with the original association (cue-O1) comes to be replaced with behavior consistent with the interpolated association (cue-O2); that is, the cue undergoes *counterconditioning* (CC, *cf.* Sherrington, 1947).

Extinction and CC have been described as instances of *retroactive interference between outcomes* (Bouton, 1993; Miller & Escobar, 2002), in which retrieval of the original cue-O1 association is diminished due to training of the interpolated cue-noO (extinction) or cue-O2 (CC) association (for clarity, interference between outcomes will be summarized as cue-O1, cue-O2 training, where O2 may represent absence of the outcome [extinction] or a different outcome [CC]). This interference view is supported by the observation that behavior consistent with the original, cue-O1, association is likely to *relapse* under certain conditions. For example, if time elapses between interpolated (cue-O2) training and testing, behavior consistent with original (cue-O1) training may spontaneously recover (Bouton & Peck, 1992; Pavlov, 1927). Additionally, if interpolated (cue-O2) training occurs in a different physical context, testing

outside the context of interpolated training will result in renewal of the original (cue-O1) response (Bouton & Bolles, 1979; Peck & Bouton, 1990). Bouton (1993; also see Miller & Escobar, 2002) proposed that recovery from interference between outcomes occurs when the cue is encountered in an environment that is not consistent with the environment in which the interpolated (cue-O2) association was acquired. The assumption is that each association (cue-O1 and cue-O2) is acquired including information about the cue, the outcome, and the (physical and temporal) context. Original training makes the cue into a reliable predictor of O1, but interpolated training makes the cue an ambiguous predictor of both O1 and O2. Thus, the organism must rely on contextual information to disambiguate the cue and respond appropriately. If the context of retrieval differs (physically or temporally) from that of interpolated training, the organism will fail to retrieve the cue-O2 association, behavior will be controlled by the cue-O1 association, and relapse will be observed.

An important feature of CC is that it involves pairings of the cue with two distinct outcomes, O1 and O2, which are usually of opposite valence. That is, a cue that originally predicted an aversive O1 can be trained to be predictive of an appetitive O2 (aversive-to-appetitive CC), or a cue that originally predicted an appetitive O1 can be trained to be predictive of an aversive O2 (appetitive-to-aversive CC). In either arrangement, behavior comes to reflect the cue-O2 association as a result of the interference (CC) procedure. In theory, aversive-to-appetitive and appetitive-to-aversive CC should be equally effective and result in equivalent control of behavior (e.g., Dickinson & Dearing, 1979; Dickinson & Pearce, 1977); however, these two procedures appear to be functionally different to some degree. For example, the number of cue-shock and cue-food trials used in nonhuman animal studies are widely different, because the aversive outcome comes to control behavior much faster than the appetitive

outcome, regardless of the direction of training (aversive-to-appetitive or appetitive-to-aversive).

This potential difference in the biological relevance of the outcomes may in turn lead to

differential likelihood of relapse in aversive-to-appetitive and appetitive-to-aversive CC, which

may have important implications for practical applications of these paradigms (see the General

Discussion).

The majority of CC studies in human participants have used evaluative conditioning

procedures (Baeyens, Eelen, Van Den Bergh, & Crombez, 1989; Gast & De Houwer,

Experiment 2b, 2013; Kerkhof, Vansteenwegen, Baeyens, & Hermans, 2011; Raes & De Raedt,

2012; Van Gucht, Baeyens, Vansteenwegen, Hermans, & Beckers, 2010). In such procedures,

participants are initially asked to evaluate an array of stimuli on a scale ranging from highly

unpleasant (negative values) to highly pleasurable (positive values). Then, several of the stimuli

that received neutral ratings (values close to zero) are paired with stimuli that received extreme

negative or positive ratings (the outcomes), which leads to the previously neutral stimuli being

rated in a manner that reflects the rating of the stimuli with which they were paired. Evaluative

CC has been reported using a wide variety of stimuli, including pictures (Baeyens et al., 1989;

Raes & De Raedt, 2012), cookies (Kerkhof et al., 2011) and colored trays (Van Gucht, Baeyens,

Hermans, & Beckers, 2013; Van Gucht et al., 2010). (Odor stimuli seem to be insensitive to

evaluative CC treatments; see Stevenson, Boakes, & Wilson, 2000).

There is limited information about spontaneous recovery and renewal following

evaluative CC. In a within subjects design, Kerkhof et al. (2011) trained participants with both

appetitive-to-aversive and aversive-to-appetitive CC using pictures of different cookies as cues

and good or bad tasting cookies as outcomes. After a 7 day retention interval, they observed a

reduction in the magnitude of CC for both appetitive-to-aversive and aversive-to-appetitive CC;

however, affective priming (latency for categorizing cues as either 'positive' or 'negative')

exhibited recovery only in the appetitive-to-aversive condition. In a similar study, Van Gucht et

al. (2010) used colored trays as cues and chocolate and Polysorbate 20 (a bitter tasting liquid) as

outcomes to investigate spontaneous recovery of appetitive-to-aversive CC. After a 7 day

interval, they observed spontaneous recovery of ratings of outcome expectancy, but no

spontaneous recovery of ratings of chocolate cravings (spontaneous recovery of evaluative

ratings was not assessed). Using the same procedure, Van Gucht et al. (2013) assessed chocolate

cravings and outcome expectancy when original and interpolated training occurred in different

contexts, which were created by altering the lighting of the room in which the different training

phases occurred. As in their spontaneous recovery study, ratings of chocolate cravings were not

sensitive to renewal, but ratings of outcome expectancy were consistent with the training

(original or interpolated) that occurred in the context of assessment. Thus, these observations of

relapse from evaluative CC seem to be dependent upon what measure is used to assess the

relationship between the cue and its associated outcomes (relapse is observed when affective

priming or US expectancy, but not cravings, are assessed). Furthermore, no study has directly

compared the magnitude of relapse after aversive-to-appetitive and appetitive-to-aversive CC,

even when both procedures have been included in the same experiment (e.g., Kerkhof et al.,

2011).

Some human CC studies have used predictive learning tasks rather than evaluative

conditioning. In predictive learning tasks, expectation of the outcome to follow the cue is

assessed with a quantifiable measure. For example, Rosas, Paredes-Olay, García-Gutiérrez,

Espinosa, and Abad (2010, Experiment 3) used CC in a predictive learning task by pairing two

distinct visual cues (i.e., a star and a circle) with different outcomes (i.e., destroying a plane or

destroying a ship) in the original learning phase. During the CC phase, one of these two cues was suddenly paired with the alternative outcome (i.e., destroying planes to destroying ships or vice versa). After this training, participants' predictions of the outcome based on the cue changed to be consistent with the outcome presented during the CC phase rather than the original learning phase. This procedure differs greatly from previously described CC procedures in that it does not utilize outcomes of opposing emotional valence. Destroying planes and destroying ships are distinct events but neither would be considered appetitive or aversive within this task. To the best of our knowledge, no study has assessed relapse from CC in a human predictive learning task.

The present research investigated the development of CC in human participants, as well as its relapse due to changes in temporal and physical context using a predictive learning paradigm. For this purpose, we developed a task in which a cue is trained as predictor of one or more outcomes, and participants provided a measure of their expectation of outcome occurrence in the presence of the cue. Furthermore, we directly compared the susceptibility of aversive-to-appetitive and appetitive-to-aversive CC to relapse phenomena (i.e., spontaneous recovery and renewal). Human participants played a computer game in which different cues (fictitious stock companies) resulted in aversive or appetitive outcomes (gaining or losing money, respectively). Expectations of the outcome were assessed in terms of wagers of fictitious money made in the presence of each stock, which were operationalized as behavioral (keypress) responses. Experiment 1 established the effectiveness of the procedure to produce both appetitive-to-aversive and aversive-to-appetitive CC. Subsequent experiments assessed spontaneous recovery from CC after a retention interval (5min or 7day, Experiment 2) and renewal from CC (Experiment 3) in both CC preparations.

Experiment 1

The goal of Experiment 1 was to obtain both appetitive-to-aversive (gain-loss) and aversive-to-appetitive (loss-gain) CC using a human predictive learning procedure. Participants played a computer game in which they made wagers (fictitious money) about the outcome of three different cues (fictitious stocks). The presence of some cues resulted in a participant losing the value of their wager, while the presence of other cues resulted in a participant gaining the value of their wager. Cues G and L were followed by a consistent outcome (gain and loss, respectively) throughout the session. However, the outcome of target Cue T changed during the latter portion of the session to create two CC types, Gain-Loss and Loss-Gain. Gain outcomes were defined as appetitive, and loss outcomes were defined as aversive.

Methods

Participants

Nineteen Auburn University undergraduate students participated in this experiment either in partial fulfillment of a course requirement or to receive extra credit for a psychology course. The sample consisted of 6 males and 13 females, with a mean age of 18.78 years (range: 18-22). Participants were randomly assigned to one of two groups, Gain-Loss ($n = 9$) or Loss-Gain ($n = 10$), where 'Gain-Loss' and 'Loss-Gain' refer to the CC procedure imposed upon cue T in each group.

Procedure

Upon entering the laboratory and providing consent, participants were asked to pretend that they were applicants for a position at a financial institution, and, in order to earn the position, they would have to demonstrate their ability to anticipate increases and decreases in the price of stocks for a pharmaceutical company based on changes in the price of stock for different

high tech companies (see Appendix B for the complete cover story). During each trial of the task, participants were presented with the information that the stocks for one high tech company (cue) increased in price and were asked to invest (wager response) fictitious money in pharmaceutical company stocks based on this information. Wagers were entered via a response bar located on the right side of the screen. This response bar was initially set to $0 and could be increased using the keyboard arrows keys to a maximum of $10,000 in increments of $1,000. Upon selecting a wager, feedback was provided as to the outcome (gain or loss) of the trial. The winnings (or losses) for the trial were added to a "profits" counter, which kept track of the participant's earnings throughout the session. The magnitude of the wager for any given trial was assumed to reflect participants' anticipation of the outcome for that trial. The fictitious high tech companies were Gamma, Omega, and Theta, and they were counterbalanced within groups for identity as Cues T, G, and L.

Table 1 presents the critical aspects of Experiment 1's procedure. During *Phase 1 (original training)*, participants received five pairings of T and either the gain (Gain-Loss condition) or loss (Loss-Gain condition) outcome. During *Phase 2 (CC)*, the outcome was reversed, so that T was paired five times with either the loss (Gain-Loss condition) or gain (Loss-Gain condition) outcome. Five L-loss and G-gain pairings were randomly distributed within each phase of training. These distractor stimuli served to provide consistent experiences with both outcomes throughout the procedure.

A performance criterion was imposed to ensure that participants acquired the experimental contingencies. This criterion required that wagers during the last trial of G (which always predicted gains) was greater than wagers during the last trial of L (which always predicted losses). Data from three participants in the Gain-Loss condition and one participant in

the Loss-Gain CC condition were excluded from all analyses using this criterion. A significance level of $\alpha = .05$ was adopted for all analyses. Effect size ($\pm$ confidence interval, CI) for analyses of variance (ANOVAs) was calculated using the partial eta-squared (η_p^2) statistic.

Results and Discussion

Acquisition of the different outcomes occurred as expected with participants making higher wagers when cues were associated with gains and lower wagers when cues were associated with losses. CC of T was successful in both the appetitive-to-aversive (Gain-Loss CC type) and aversive-to-appetitive (Loss-Gain CC type) directions. The following analyses support these observations.

Responding to Cues G and L was compared to determine the effectiveness of the gain and loss outcomes with which they were consistently paired. A 2 (cue: G *vs.* L) x 2 (CC type: Gain-Loss *vs.* Loss-Gain) x 10 (trial) ANOVA revealed main effects of Cue, $F(1, 13) = 73.18$, $\eta_p^2 = .85$ (CI: .65, .90), and trial, $F(9, 117) = 2.78$, $\eta_p^2 = .18$ (CI: .03, .22) as well as a Cue x Trial interaction, $F(9, 117) = 9.57$, $\eta_p^2 = .42$ (CI: .27, .48). The main effect of CC type as well as all other interactions were not significant, all Fs < 1. Thus participants' wagers were larger for G than for L, and this differential responding was acquired over the course of the session. Responding to G was disrupted temporarily during its first presentation in the CC phase (see Figure 1). This disruption was likely due to the outcome of T changing at the start of the CC phase, and this disruption was not observed after participants were given feedback about the outcome of G during the CC phase.

CC of T was analyzed with a 2 (CC type) x 2 (phase: original training *vs.* CC) x 5 (trial) ANOVA which revealed a main effect of CC type, $F(1, 13) = 24.56$, $\eta_p^2 = .65$ (CI:. 31, .77), and a CC type x Phase x Trial interaction, $F(4, 52) = 23.82$, $\eta_p^2 = .65$ (CI: .48, .71). The main effect

of phase and trial and all other interactions were not significant, all Fs < 2.4 (see Figure 2). The occurrence of CC was confirmed by comparing the last trial of original training and the last trial of CC. CC occurred in both the Gain-Loss and Loss-Gain conditions, Fs(1, 13) = 7.01 and 9.85, η_p^2s = .35 (CI: .04, .57) and .43 (CI: .08, .63), respectively. That is, at completion of the CC phase, behavior elicited by the cue reflected the association trained during CC (T-O2) rather than the original association (T-O1).

Experiment 2

Experiment 1 established that our predictive learning task can yield both appetitive-to-aversive and aversive-to-appetitive CC in humans. Experiment 2 investigated spontaneous recovery from CC in our predictive learning task by assessing responding to T when a retention interval was introduced between CC and assessment of T. To the best of our knowledge, spontaneous recovery after appetitive-to-aversive and aversive-to-appetitive CC has not been assessed in human predictive learning. Nonetheless, spontaneous recovery from CC has been reported in evaluative conditioning procedures after a 7day interval (Kerkhof et al., 2011; Van Gucht et al., 2010). Therefore, we introduced a 7day retention interval between CC training and test, and expected to see a recovery of T-O1 responding in both CC directions after such a retention interval. A control condition in which participants were asked to leave the laboratory and return for testing 5min later was added to this study to assess the effects of the retention interval when all other factors (e.g., interrupting the task and returning to the laboratory) were held constant.

Methods

Participants

One hundred thirty one Auburn University undergraduate students participated in this experiment either in partial fulfillment of a course requirement or to receive extra credit for a psychology course. The sample consisted of 27 males and 104 females, with a mean age of 19.49 years (range: 18-29). Participants were randomly assigned to one of four groups: Gain-Loss/5min ($n = 35$), Loss-Gain/5min ($n = 32$), Gain-Loss/7day ($n = 32$), and Loss-Gain/7day ($n = 32$), where 'Gain-Loss' and 'Loss-Gain' refer to the CC procedure used in each group, and '/5min' and '/7day' refer to the interval imposed between CC and testing.

Procedure

Original and CC training phases occurred as described for Experiment 1. In order to investigate spontaneous recovery, half of the participants were asked to return 7 days after completion of CC training, and the other half of the participants were asked to return 5 min after completion of CC training. During the follow up session, participants were seated at the same computer they occupied during the initial session, read a brief reiteration of the cover story, and were given one final trial with cues T, G, and L. Wagers given to T during this test trial were compared to those given during the final T trial of the CC phase, and this comparison was expected to reveal the degree to which participants recalled either the T-O1 association or the T-O2 association (see Table 1).

The same performance criterion described in Experiment 1 was implemented in Experiment 2. Data from three participants in each of Groups Gain-Loss/5min and Loss-Gain/5min, two participants in Group Loss-Gain/7day, and six participants in Group Gain-Loss/7day were excluded from all analyses following this criterion.

<div style="text-align:center">

Results and Discussion

</div>

Experiment 2 successfully replicated the CC observed in Experiment 1. Furthermore, spontaneous recovery was observed for both appetitive-to-aversive and aversive-to-appetitive CC, but relapse was more likely in the aversive-to-appetitive than the appetitive-to-aversive direction. Aversive-to-appetitive CC was susceptible to spontaneous recovery after any temporal disruption, while spontaneous recovery from appetitive-to-aversive CC was only evident after the 7day retention interval. The following analyses support these observations.

The effectiveness of the G-gain and L-loss trials was confirmed by a 2 (cue: G *vs.* L) x 2 (CC type: Gain-Loss *vs.* Loss-Gain) x 2 (retention interval: 5min *vs.* 7day) x 10 (trial) ANOVA, revealing main effects of CC type and cue, Fs(1, 113) = 4.10 and 336.24, η_p^2s = .04 (CI: .00, .11) and .75 (CI: .68, .79), respectively, as well as a main effect of trial, F(9, 1017) = 14.23, η_p^2 = .11 (CI: .08, .14). There were also interactions of Cue x Retention interval, F(1, 113) = 4.61, η_p^2 = .04 (CI: .00, .11), CC type x Retention interval x Trial, Cue x Trial, and Cue x CC type x Trial, Fs(9, 1017) = 2.44, 70.42, and 3.32, η_p^2s = .02 (CI: .00, .03), .38 (CI: .34, .41), and .03 (CI: .01, .04), respectively. The main effect of retention interval and all remaining interactions were not significant, all Fs < 3.6. Participants consistently made higher wagers to G than L across trials for both CC types. As in Experiment 1, responding to G was briefly disrupted on the first trial where G was presented during the CC phase, but responding returned to normal by the second trial of G (see Figure 3). Prior to the imposition of the retention intervals, both the 5min and 7day conditions responded similarly to G and L.

Figure 4 presents the wagers made to T during each trial of original training (Phase 1) and CC (Phase 2) as well as the test trial for spontaneous recovery at the 5min and 7day retention intervals (SR Test). CC of T was assessed with a 2 (CC type) x 2 (Retention interval) x 2 (phase: original training *vs.* CC) x 5 (trial) ANOVA. We observed significant main effects of CC type

and phase, Fs$(1, 113)$ = 14.79 and 11.11, η_p^2s = .12 (CI: .04, .21) and .09 (CI: .02, .18),

respectively, as well as a main effect of trial, $F(4, 452)$ = 17.84, η_p^2 = .14 (CI: .09, .18). There

were also interactions of CC type x Phase, $F(1, 113)$ = 72.41, η_p^2 = .39 (CI: .27, .48), CC type x

Trial, Phase x Trial, and CC type x Phase x Trial, Fs$(4, 452)$ = 14.66, 8.41, and 113.04, η_p^2s =

.11 (CI: .07, .16), .07 (CI: .03, .10), and .50 (CI: .45, .54), respectively. The main effect of

retention interval and remaining interactions were not significant, all Fs < 1. Planned

comparisons between the last trial of the original training phase and the last trial of the CC phase

confirmed that CC was successful in both CC types. For Gain-Loss/5min and Gain-Loss/7day,

participants made smaller wagers to T during the last trial of the CC phase than during that of the

original training phase, Fs$(1, 113)$ = 41.87 and 27.30, η_p^2s = .27 (CI: .16, .37) and .19 (CI: .09,

.30), respectively. Conversely, for Loss-Gain/5min and Loss-Gain/7day, participants made

larger wagers to T during the last trial of the CC phase than that of the original training phase,

Fs$(1, 113)$ = 43.90 and 54.17, η_p^2s = .28 (CI: .17, .38) and .32 (CI: .21, .42), respectively.

Therefore, CC was successful for all four groups prior to imposing the retention interval.

Spontaneous recovery was assessed with a 2 (CC type) x 2 (retention interval: 5min $vs.$

7day) x 2 (trial: last trial of Phase 2 $vs.$ test trial) ANOVA. Eight participants (3 in Group Gain-

Loss/7day and 5 in Group Loss-Gain/7day) who met the performance criterion for inclusion in

the CC analysis did not return for the follow up session and therefore could not be included in

analysis of spontaneous recovery. This analysis revealed main effects of CC type and trial, Fs$(1,$

$105)$ = 40.04 and 8.25, η_p^2s = .28 (CI: .16, .38) and .07 (CI: .01, .16), respectively, and a CC type

x Trial interaction, $F(1, 105)$ = 47.50, η_p^2 = .31 (CI: .19, .41). The main effect of retention

interval and all remaining interactions were not significant, all Fs < 3.3. In order to assess

whether the T-O1 association recovered during the test trial, a planned comparison between the

last trial of the CC phase and the test trial was conducted for both CC types at the 5min and 7day retention interval. For the Gain-Loss CC type, responding to T increased during the test trial after the 7day interval, $F(1, 105) = 5.98$, $\eta_p^2 = .05$ (CI: .01, .14) but did not significantly change after the 5min interval, $F(1, 105) < 2.3$. This suggests that the T-gain association recovered only after the 7day interval. For the Loss-Gain CC type, responding to T significantly decreased during the test trial after both the 7day and 5min interval, $Fs(1, 105) = 16.42$ and 33.54, $\eta_p^2 s = .14$ (CI: .05, .24) and $.24$ (CI: .13, .35), respectively. It appears that the T-loss association (T-O1, Loss-Gain CC type) was prone to recovery after any temporal disruption upon completion of the CC phase (see Figure 4).

While these results suggest that spontaneous recovery occurred for both the Gain-Loss and Loss-Gain CC type, our failure to observe a significant main effect of retention interval presents a problem for this interpretation. It is possible that upon returning to the laboratory, participants' engagement in the task had decreased, and they may not have been attending to the stimuli as closely as they were during the initial session. Thus, the performance criterion was once again applied to responses to G and L made during the test session. Thirty-four participants failed to meet this second criterion (6 in Group Loss-Gain/5min, 9 in Group Loss-Gain/7day, 12 in Group Gain-Loss/5min, and 7 in Group Gain-Loss/7day. Spontaneous recovery was again assessed after excluding these participants, revealing a main effect of CC type, retention interval, and trial, $Fs(1, 71) = 40.11$, 8.82, and 6.99, $\eta_p^2 s = .36$ (CI: .21, .48), $.11$ (CI: .02, .23) and $.09$ (CI: .01, .20), respectively, as well as a CC type x Trial interaction, $F(1, 71) = 30.17$, $\eta_p^2 = .30$ (CI: .16, .42). All remaining interactions were not significant, all $Fs < 2.2$. Wagers for the last trial of T in the CC phase and the test trial of T differed after both the 7day and 5min interval in the Loss-Gain condition, $Fs(1, 71) = 12.85$ and 23.24, $\eta_p^2 s = .15$ (CI: .05, .28) and $.25$ (CI: .11,

.37), respectively, but only after the 7day interval in the Gain-Loss condition, $F(1, 71) = 6.19$, $\eta_p^2 = .08$ (CI: .01, .19) ($F(1, 71) < 0.04$ in the Gain-Loss/5 min group). These planned comparisons are consistent with those of prior analysis using only the first exclusion criterion. However, this analysis yielded a significant main effect of retention interval and, therefore, better supports our interpretation that spontaneous recovery was likely for both CC types with the Loss-Gain CC type being more susceptible to our temporal manipulations than the Gain-Loss CC type.

Experiment 2 demonstrated that retrieval of the T-O1 association increased in both appetitive-to-aversive and aversive-to-appetitive CC when time lapsed between CC and testing. These findings are consistent with previous studies that have investigated spontaneous recovery from CC procedures in human and nonhuman subjects (e.g., Van Gucht et al., 2010; Bouton & Peck, 1992). Interestingly, relapse from aversive-to-appetitive CC was more likely than relapse from appetitive-to-aversive CC, because even a 5 min interval was sufficient to obtain relapse from aversive-to-appetitive CC. Experiment 3 continued our assessment of relapse from CC with a comparison of the susceptibility of these two procedures to renewal due to a physical context change.

Experiment 3

As is the case for most forms of interference, responding after CC manipulations appears to be strongly modulated by contextual cues (Peck & Bouton, 1990). Such renewal has been observed in human participants tested with an evaluative conditioning task (Van Gucht et al., 2013), at least in the appetitive-to-aversive direction. In Experiments 3, two contexts (Contexts A and B) were created in the framework of our experimental task. Participants acquired the original association in Context A. CC training occurred either in the same context (Context A) or a different context (Context B) from that of original training. Then, responding was assessed

either in the context of original learning (Context A) or the context of CC training (Context B). This resulted in four conditions: AAA, ABA, AAB, and ABB, where the first letter represents the context of original training, the second letter represents the context of CC training, and the third letter represents the context of testing. Maximal renewal was expected in the ABA groups, in which testing was conducted in the context of original training of the original association.

Methods

Participants

One hundred three Auburn University undergraduate students participated in this experiment either as a research requirement or to receive extra credit for a psychology course. The sample consisted of 30 males and 73 females, with a mean age of 19.72 years (range: 18-31). Participants were randomly assigned to one of eight training groups, Gain-Loss/AAA ($n = 13$), Gain-Loss/AAB ($n = 13$), Gain-Loss/ABA ($n = 12$), Gain-Loss/ABB ($n = 13$); Loss-Gain/AAA ($n = 13$), Loss-Gain/AAB ($n = 13$), Loss-Gain/ABA ($n = 14$), and Loss-Gain/ABB ($n = 12$). In all group designations, 'Gain-Loss' and 'Loss-Gain' refer to the CC procedure used, and '/AAA,' '/AAB,' '/ABB,' and '/ABA' refer to the contextual manipulation being used.

Procedure

Original and CC training phases were conducted using the same procedure used in Experiments 1 and 2 with the following exceptions. Two different contexts were created based on different configurations of the screen (including location of the stimuli, response bar, and color of the background and text); each configuration was counterbalanced within groups to serve as Contexts A and B. Participants received no information regarding the existence or meaning of these contexts. All participants received original training in Context A. CC training occurred in either Context A (Groups AAA and AAB) or Context B (Groups ABA and ABB).

Testing consisted of an additional trial of Cue T, and it occurred in either Context A (Groups AAA and ABA) or Context B (Groups AAB and ABB). Each renewal group was subdivided to assess the Gain-Loss and Loss-Gain CC types, for a total of eight treatment groups.

In order to equate the association of the different contexts to the gain and loss outcomes during training, additional trials of L or G (consistently associated to the loss and gain outcome, respectively) were added to the original training and CC phase. There were five additional L-loss presentations when T predicted gains (Phase 1 in the Gain-Loss condition and Phase 2 in the Loss-Gain condition) and five additional G-gain presentations when T predicted loss (Phase 1 in the Loss-Gain condition and Phase 2 in the Gain-Loss condition). Therefore, regardless of treatment group, both the original training and CC phases contained 10 trials resulting in gain outcomes and 10 trials resulting in loss outcomes (see Table 1).

Eight participants (1 in each of Groups Gain-Loss/AAA, Gain-Loss/ABB, Loss-Gain/AAA, and Loss-Gain/AAB, and 2 in each of Groups Gain-Loss/AAB, and Loss-Gain/ABA) failed to meet the performance criterion described in Experiment 1, and their data were excluded from all analyses.

Results and Discussion

As in Experiments 1 and 2, CC was observed in the Gain-Loss and Loss-Gain conditions. Renewal was assessed after both appetitive-to-aversive and aversive-to-appetitive CC; however, renewal was only observed for the aversive-to-appetitive direction in the ABA renewal group. In the appetitive-to-aversive CC condition, the T-O1 association did not relapse after any of the contextual manipulations; rather, responding during the test trial was consistent with that of the CC phase. The following analyses support these observations.

The effectiveness of the G-gain and L-loss pairings was confirmed by a 2 (cue: G *vs.* L) x 2 (CC type) x 2 (CC context: same [as original training] *vs.* different [from original training]) x 15 (trial) ANOVA. This analysis revealed main effects of cue and CC type, Fs(1, 91) = 483.00 and 4.85, η_p^2s = .84 (CI: .79, .87) and .05 (CI: .00, .14), respectively, as well as a main effect of trial, $F(14, 1274) = 23.90$, $\eta_p^2 = .21$ (CI: .17, .23). There were also interactions of Cue x CC type, $F(1, 91) = 5.63$, $\eta_p^2 = .06$ (CI: .00, .15), CC type x Trial, Cue x Trial, Cue x CC type x Trial, and Cue x CC context x Trial, Fs(14, 1274) = 12.378, 74.61, 7.23, and 1.97, η_p^2s = .12 (CI: .08, .14), .45 (CI: .41, .47), .07 (CI: .04,.09), and .02 (CI: .00, .02), respectively. The main effect of CC context and all other interactions were not significant, all Fs < 1.5. Participants consistently made higher wagers to G than L throughout the session. As in Experiments 1 and 2, responding to G was disrupted on the first trial of G during the CC phase (6th trial of G for Gain-Loss CC type, 11th trial of G for Loss-Gain CC type), and this disruption was observed regardless of whether CC was conducted in the same or different context as original training. After feedback was provided for G of this trial, the disruption disappeared for all treatment groups (see Figure 5).

Figure 6 presents responding to T during each trial of the original (Phase 1) and CC (Phase 2) training, as well as the test trial to assess renewal (Renewal Test). CC of T was assessed with a 2 (CC type) x 2 (CC context) x 2 (phase) x 5 (trial) ANOVA, which revealed main effects of CC type, $F(1, 91) = 14.50$, $\eta_p^2 = .14$ (CI: .04, .25), and trial $F(4, 364) = 10.19$, $\eta_p^2 = .10$ (CI: .05, .14), as well as interactions of CC type x Phase, $F(1, 91) = 102.78$, $\eta_p^2 = .53$ (CI: .41, .61), CC type x Trial, Phase x Trial, CC type x Phase x Trial, CC context x Phase x Trial, and CC type x CC context x Phase x Trial, Fs (4, 364) = 24.17, 19.18 169.31, 3.11, and 2.77, η_p^2s = .21 (CI: .14, .26), .17 (CI: .11, .22), .65 (CI: .60, .68) , .03 (CI: .00, .06), and .03 (CI:

.00, .05), respectively. The main effect of CC context and all remaining interactions were not significant, Fs < 3.2. Planned comparisons between the final trial of both phases revealed that for the Loss-Gain CC type responding to T was higher on the last trial of CC than that of original training, Fs(1, 91) = 51.46 [Same CC context] and 55.97 [Different CC context], η_p^2s= .36 (CI: .23, .47) and .38 (CI: .25, .48), respectively. For the Gain-Loss CC type, responding to T was lower on the last trial of CC than that of original training, Fs(1, 91) = 87.79 [Same CC context] and 68.06 [Different CC context], η_p^2s = .49 (CI: .37, .58) and .43 (CI: .30, .53), respectively. Therefore, CC was successfully obtained in both CC types regardless of context in which CC occurred.

Renewal was assessed with a 2 (CC type) x 2 (CC context) x 2 (test context) x 2 (trial) ANOVA, revealing a main effect of CC type, $F(1, 87) = 88.53$, $\eta_p^2 = .50$ (CI: .38, .59), and interactions of CC type x CC context x Test context, CC type x Trial, CC type x CC context x Trial, CC context x Test context x Trial, and CC type x CC context x Test context x Trial, Fs (1, 87) = 4.28, 5.10, 8.90, 4.74, and 12.02, η_p^2s = .05 (CI: .00, .14), .06 (CI: .00, .15), .09 (CI: .02, .20), .05 (CI: .00, .14), and .12 (CI: .03, .23), respectively. All remaining main effects and interactions were not significant, all Fs < 3.8. Planned comparisons between the final CC trial and the test trial indicated that renewal of T-O1 responding was only observed in the Loss-Gain/ABA group, $F(1, 87) = 38.36$, $\eta_p^2 = .31$ (CI: .18, .42). For all other groups responding to T at test was consistent with the CC phase, Fs < 1.4 (see Figure 6).

The strength of the renewal was assessed with a 2 (CC type) x 2 (CC context) x 2 (test context) ANOVA of responding to T that occurred during the test trial. This analysis revealed a main effect of CC type, $F(1, 87) = 65.97$, $\eta_p^2 = .43$ (CI: .30, .53), and interactions of CC type x CC context, CC context x Test context, and CC type x CC context x Test context, Fs(1, 87) =

4.95, 7.11, and 12.95, η_p^2s = .05 (CI: .00, .15), .08 (CI: .01, .17), and .13 (CI: .04, .24),

respectively. The main effects of CC context and test context, and the CC type x Test context

interaction were not significant, all *F*s < 1.93. Four planned comparisons were conducted for

both CC types: AAA *vs.* ABA, AAA *vs.* AAB, ABA *vs.* ABB, and AAB *vs.* ABB. For the Loss-

Gain CC type, the AAA *vs.* ABA, AAA *vs.* AAB, and ABA *vs.* ABB comparisons were

significant, *F*s (1, 87) = 24.67, 7.52, and 12.66, η_p^2s = .22 (CI: .10, .34), .08 (CI: .01, .18), and

.13 (CI: .04, .24), respectively. The AAB *vs.* ABB comparison was non-significant, $F(1, 87) <$

1.8. Consistent with the lack of renewal observed in Gain-Loss condition, none of these

comparisons were statistically significant, *F*s < 0.8 (see Figure 6). These findings may indicate

that aversive-to-appetitive CC has a greater susceptibility to renewal than appetitive-to-aversive

CC.

General Discussion

The purpose of the present series of experiments was to assess and compare human

participants' susceptibility to relapse from appetitive-to-aversive and aversive-to-appetitive CC

procedures using a predictive learning task. In Experiment 1, we demonstrated that gaining and

losing fictitious money were effective appetitive and aversive outcomes, respectively, and our

predictive learning task successfully produced both appetitive-to-aversive and aversive-to-

appetitive CC of the target cue (T) within a single experimental session. Prior evaluative CC

procedures have also obtained CC in both valence directions (e.g., Kerkhof et al., 2011), but the

present studies provide the first report of both procedures using a predictive learning task. In

Experiment 2, we found evidence for spontaneous recovery from both appetitive-to-aversive and

aversive-to-appetitive CC after a 7day retention interval. Prior evaluative CC procedures have

observed spontaneous recovery in the appetitive-to-aversive direction (Kerkhof et al., 2011; Van

Gucht et al., 2010); however, we observed that aversive-to-appetitive CC was more susceptible to temporal changes between CC and testing than appetitive-to-aversive CC. This greater susceptibility is clear from the observation that a 5min retention interval resulted in recovery of T-O1 responding after aversive-to-appetitive CC, but not appetitive-to-aversive CC. Finally, in Experiment 3, we observed ABA renewal from CC but only in the aversive-to-appetitive direction. For appetitive-to-aversive CC, responding to T during the test phase was once again consistent with the outcome trained during the CC phase (i.e., the loss outcome) regardless of whether testing occurred in the context of CC or original training. Contrary to our results, Van Gucht et al. (2013) found evidence for ABA renewal from evaluative CC in the appetitive-to-aversive direction; however, this effect was not observed in all measures of the cue-outcome association (i.e., they found relapse of outcome expectancy but not self-reported appetitive cravings, and the latter may be more susceptible to the emotional valence of the outcome). Additionally, they did not conduct aversive-to-appetitive CC, so it cannot be determined whether there would be a difference based upon CC type in their type of preparation. Taken together, the results of Experiments 2 and 3 suggest that aversive-to-appetitive CC is more susceptible to relapse than appetitive-to-aversive CC.

Bouton's (1993) retrieval theory proposes that relapse of the originally learned outcome association (cue-O1) occurs due to the test conditions failing to activate memories of the interpolated association (cue-O2) after a temporal or physical change in context, but this theory does not specify potential differences in sensitivity to the effects of these contextual changes when O2 is appetitive or aversive. Our observations, while largely consistent with Bouton's retrieval theory, may suggest that contextual changes can have differential effects on relapse that depend upon the valence of the originally trained and interpolated associations. In Experiment 2,

as retrieval theory would predict, spontaneous recovery was observed in both CC directions after the 7day retention interval; however, our observation of spontaneous recovery after the 5min retention interval *only* in the aversive-to-appetitive direction is not expected within this theoretical framework. The temporal contexts of CC training and testing should decrease in similarity as the interval between them becomes larger, so it seems reasonable to assume that these temporal contexts would be more similar after a 5min than a 7day interval. Recovery of responding in the Loss-Gain condition even after this short interval (and when the CC training-testing temporal context similarity is presumably high) would suggest that aversive-to-appetitive CC is more susceptible to recovery than appetitive-to-aversive CC. Notably, brief interruptions (reading text extraneous to the task) between phases of a predictive learning task have been shown to have disruptive effects on interference procedures, in a manner relatively independent of the duration of the retention interval (Escobar, Arcediano, & Miller, 2003). However, whether the duration or presence of extraneous stimuli resulted in recovery of responding in the 5min condition of Experiment 2, these factors had greater impact after aversive-to-appetitive than appetitive-to-aversive CC condition, supporting our view that the former CC type is more susceptible to relapse.

The tendency for humans to be averse to loss outcomes could possibly account for the differences observed between aversive-to-appetitive and appetitive-to-aversive CC. Losses appear to be weighted more heavily than gains when making risky decisions (Tversky & Kahneman, 1981). Specifically, choices phrased in terms of gains, result in decisions to avoid risk, while choices phrased in terms of loss, result in decisions to take risky gambles to avoid certain loss. This propensity to avoid loss is quite robust and has been replicated across several different types of choice scenarios (e.g., De Martino, Kumaran, Seymour, & Dolan, 2006;

McElroy & Seta, 2003; Pinillos, Smith, Nair, Marchetto, & Mun, 2011; Smith & Levin, 1996).

While our stock market game was not designed as a risky decision-making task, there is certainly

an element of risk involved in making wagers to T during test trials given T's history with both

loss and gain outcomes. If our manipulations of retention interval and physical context caused

uncertainty about the outcome of T during the test phase, participants may have avoided loss by

making low wagers to T during the test phase regardless of whether the loss outcome was trained

during the original training phase (Loss-Gain) or the CC phase (Gain-Loss). From this

perspective, uncertainty about the outcome of T created by our [temporal and physical]

contextual manipulations, combined with a potential aversion to loss, could have masked relapse

in the appetitive-to-aversive direction and exacerbated relapse observed in the aversive-to-

appetitive direction.

Manipulating the ratio of loss to gain on trials presenting T could potentially reduce loss

aversion within this predictive learning task. In the current design, there is a 1:1 ratio between

wagers and outcome (e.g., a $1000 wager results in either a $1000 gain or a $1000 loss), but

Tversky and Kahneman (1981) demonstrated that losses and gains of the same quantity have

quite different subjective values. Therefore, the aversiveness of losing $1000 is likely to be of

larger subjective magnitude than the appetitiveness of gaining $1000. If we were to modify

these ratios such that the T-gain ratio was greater than the T-loss ratio (e.g., 1:2 for T-gain and

1:0.5 for T-loss), then we may be able to better equate the subjective value of our appetitive and

aversive outcomes. This would ideally eliminate the influence of loss aversion, and allow for a

better comparison of relapse between our appetitive-to-aversive and aversive-to-appetitive

preparations.

Differences in the biological significance of appetitive and aversive outcomes could offer another alternative for why relapse may be more likely in aversive-to-appetitive than appetitive-to-aversive CC. While appetitive outcomes (e.g., food and water) are necessary for survival in the long term, aversive outcomes (e.g., injury and sickness) are likely to be more biologically significant because they tend to have more immediate and detrimental survival consequences. Developing an adaptive response to the environment is one of the primary functions of Pavlovian conditioning (Domjan, 2005; Hollis, 1997), and 'relapse' is not necessarily maladaptive. Environmental situations are often ambiguous; therefore, an organism's ability to use both the temporal and physical features of the context to disambiguate the meaning of a cue may provide a highly adaptive way of selecting the most appropriate behavior given the current situation. This is especially true if the cue has previously been associated with an aversive outcome that may be dangerous or harmful. Thus, when there is ambiguity as to the outcome associated with a cue (e.g., when a change in context occurs after CC), higher likelihood of retrieval for the aversive than appetitive association is likely to be more adaptive, even if the cue was most recently followed by an appetitive outcome.

The biological significance of stimuli can have a large impact on interference training as well as cue competition within Pavlovian conditioning experiments (e.g., Denniston, Miller, & Matute, 1996; Domjan, 2005; Krause, Cusato, & Domjan, 2003; Miller & Matute, 1996; Oberling, Bristol, Matute, & Miller, 2000). Krause et al. (2003) found that after sexual conditioning cues having biological significance were more resistant to extinction than 'neutral' cues. Additionally, attempts to block and overshadow associations with biologically significant cues have been ineffective in reducing conditioned responses to these stimuli (Denniston et al., 1996; Miller & Matute, 1996; Oberling et al., 2000). While the effect of biological significance

on susceptibility to relapse phenomena has not been investigated, these previous studies illustrate the difficulty in modifying the associability of biologically significant stimuli. If aversive outcomes are more biologically significant than appetitive outcomes, then cues initially paired with aversive outcomes could be better protected from lasting behavioral changes resulting from CC than cues initially paired with appetitive outcomes. Our observed differences in relapse from appetitive-to-aversive and aversive-to-appetitive CC support this possibility.

While relapse may not be detrimental to survival, it can decrease the effectiveness of CC procedures when applied to modify problem behaviors. Aversive-to-appetitive CC is the mechanism underlying *systematic desensitization treatments* of fears and phobias (e.g., de Jong, Vorage, & van den Hout, 2000; Paunovic, 1999; Wolpe, 1958). Techniques derived from aversive-to-appetitive CC may also be applicable to a wider range of conditions, such as decreasing the pain of medical procedures without the need of analgesics (Slifer, Eischen, & Busby, 2002) and decreasing cue-induced allergic reactions (Khan, 1977). Techniques derived from appetitive-to-aversive CC have been employed in the treatment of substance abuse behaviors in order to alter the predictive value of different drug cues (Childress, Ehrman, McLellan, & O'Brien, 1988; Frawley & Smith, 1990). The present study suggests that these appetitive-to-aversive treatments may be more resistant to relapse than aversive-to-appetitive treatments. If future investigations corroborate this claim, it could have implications as to the utility of these procedures and better inform clinicians as they create treatment plans for problem behaviors.

References

Baeyens, F., Eelen, P., Van Den Bergh, O., & Crombez, G. (1989). Acquired affective-evaluative value: Conservative but not unchangeable. *Behaviour Research and Therapy*, *27*, 279–287.

Bouton, M. E. (1993). Context, time, and memory retrieval in the interference paradigms of Pavlovian learning. *Psychological Bulletin*, *114*, 80–99.

Bouton, M. E., & Bolles, R. C. (1979). Contextual control of the extinction of conditioned fear. *Learning and Motivation*, *10*, 445–466.

Bouton, M. E., & Peck, C. A. (1992). Spontaneous recovery in cross-motivational transfer (counterconditioning). *Animal Learning & Behavior*, *20*, 313–321.

Childress, A., Ehrman, R., McLellan, A., & O'Brien, C. (1988). Update on behavioral treatments for substance abuse. *NIDA Research Monograph Series*, *90*, 183–192.

De Jong, P. J., Vorage, I., & van den Hout, M. A. (2000). Counterconditioning in the treatment of spider phobia: Effects on disgust, fear and valence. *Behaviour Research and Therapy*, *38*, 1055–1069.

De Martino, B., Kumaran, D., Seymour, B., & Dolan, R. J. (2006). Frames, biases, and rational decision-making in the human brain. *Science*, *313*, 684–687.

Denniston, J. C., Miller, R. R., & Matute, H. (1996). Biological significance as a determinant of cue competition. *Psychological Science*, *7*, 325–331.

Dickinson, A., & Dearing, M. F. (1979). Appetitive-aversive interactions and inhibitory processes. In A. Dickinson & R. a. Boakes (Eds.), *Mechanisms of learning and motivation: A memorial volume to Jerzy Konorksi* (pp. 203–231). Totawa, NJ: Erlbaum.

Dickinson, A., & Pearce, J. M. (1977). Inhibitory interactions between appetitive and aversive

stimuli. *Psychological Bulletin*, *84*, 690–711.

Domjan, M. (2005). Pavlovian conditioning: A functional perspective. *Annual Review of

Psychology*, *56*, 179–206.

Escobar, M., Arcediano, F., & Miller, R. R. (2003). Latent inhibition in human adults without

masking. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *29*,

1028–1040.

Frawley, P. J., & Smith, J. W. (1990). Chemical aversion therapy in the treatment of cocaine

dependence as part of a multimodal treatment program: Treatment outcome. *Journal of

Substance Abuse Treatment*, *7*, 21–29.

Gast, A., & De Houwer, J. (2013). The influence of extinction and counterconditioning

instructions on evaluative conditioning effects. *Learning and Motivation*, *44*, 312–325.

Hollis, K. L. (1997). Contemporary research on Pavlovian conditioning: A "new" functional

analysis. *American Psychologist*, *52*, 956–965.

Kerkhof, I., Vansteenwegen, D., Baeyens, F., & Hermans, D. (2011). Counterconditioning: An

effective technique for changing conditioned preferences. *Experimental Psychology*, *58*,

31–38.

Khan, A. U. (1977). Effectiveness of biofeedback and counter-conditioning in the treatment of

bronchial asthma. *Journal of Psychosomatic Research*, *21*, 97–104.

Krause, M. a., Cusato, B., & Domjan, M. (2003). Extinction of conditioned sexual responses in

male Japanese quail (Coturnix japonica): Role of species-typical cues. *Journal of

Comparative Psychology*, *117*, 76–86.

McElroy, T., & Seta, J. J. (2003). Framing effects: An analytic–holistic perspective. *Journal of Experimental Social Psychology*, *39*, 610–617.

Miller, R. R., & Escobar, M. (2002). Associative interference between cues and between outcomes presented together and presented apart: An integration. *Behavioural Processes*, *57*, 163–185. http://doi.org/10.1016/S0376-6357(02)00012-8

Miller, R. R., & Matute, H. (1996). Biological significance in forward and backward blocking: Resolution of a discrepancy between animal conditioning and human causal judgment. *Journal of Experimental Psychology: General*, *125*, 370–386.

Oberling, P., Bristol, A. S., Matute, H., & Miller, R. R. (2000). Biological significance attenuates overshadowing, relative validity, and degraded contingency effects. *Animal Learning & Behavior*, *28*, 172–186.

Paunovic, N. (1999). Exposure counterconditioning (EC) as a treatment for severe PTSD and depression with an illustrative case. *Journal of Behavior Therapy and Experimental Psychiatry*, *30*, 105–117.

Pavlov, I. P. (1927). *Conditioned reflexes*. London, England: Oxford University Press.

Peck, C. A., & Bouton, M. E. (1990). Context and performance in aversive-to-appetitive and appetitive-to-aversive transfer. *Learning and Motivation*, *21*, 1–31.

Pinillos, A. N., Smith, N., Nair, G. S., Marchetto, P., & Mun, C. (2011). Philosophy ' s New Challenge : Experiments and Intentional Action. *Mind and Language*, *26*, 115–139.

Raes, A. K., & De Raedt, R. (2012). The effect of counterconditioning on evaluative responses and harm expectancy in a fear conditioning paradigm. *Behavior Therapy*, *43*, 757–767.

Rosas, J. M., Paredes-Olay, M. C., García-Gutiérrez, A., Espinosa, J. J., & Abad, M. J. F. (2010). Outcome-specific transfer between predictive and instrumental learning is unaffected by

extinction but reversed by counterconditioning in human participants. *Learning and Motivation*, *41*, 48–66.

Sherrington, C. S. (1947). *The integrative action of the central nervous system*. Cambrige, UK: Cambridge University Press.

Slifer, K. J., Eischen, S. E., & Busby, S. (2002). Using counterconditioning to treat behavioural distress during subcutaneous injections in a paediatric rehabilitation patient. *Brain Injury*, *16*, 901–916.

Smith, S. M., & Levin, I. P. (1996). Need for cognition and choice framing effects. *Journal of Behavioral Decision Making*, *9*, 283–290.

Stevenson, R. J., Boakes, R. a., & Wilson, J. P. (2000). Counter-conditioning following human odor–taste and color–taste learning. *Learning and Motivation*, *31*, 114–127.

Tversky, A., & Kahneman, D. (1981). The framing of decisions and the psychology of choice. *Science*, *211*, 453–458.

Van Gucht, D., Baeyens, F., Hermans, D., & Beckers, T. (2013). The inertia of conditioned craving. Does context modulate the effect of counterconditioning? *Appetite*, *65*, 51–57.

Van Gucht, D., Baeyens, F., Vansteenwegen, D., Hermans, D., & Beckers, T. (2010). Counterconditioning reduces cue-induced craving and actual cue-elicited consumption. *Emotion*, *10*, 688–695.

Wolpe, J. (1958). *Psychotherapy by Reciprocal Inhibition*. Standford, CA: Stanford University Press.

Apendix A

Table 1

Design of Experiments 1, 2, and 3

Group	Phase 1 (Original Training)	Phase 2 (CC training)	Retention Interval	Test
Experiment 1				
Gain-Loss	5 T+ / 5 L- / 5 G+	5 T- / 5 L- / 5 G+		
Loss-Gain	5 T- / 5 L- / 5 G+	5 T+ / 5 L- / 5 G+		
Experiment 2				
Gain-Loss/5min	5 T+ / 5 L- / 5 G+	5 T- / 5 L- / 5 G+	5min	T/ L / G
Loss-Gain/5min	5 T- / 5 L- / 5 G+	5 T+ / 5 L- / 5 G+	5min	T/ L / G
Gain-Loss/7day	5 T+ / 5 L- / 5 G+	5 T- / 5 L- / 5 G+	7day	T/ L / G
Loss-Gain/7day	5 T- / 5 L- / 5 G+	5 T+ / 5 L- / 5 G+	7day	T/ L / G
Experiment 3				
Gain-Loss				
AAA	[5 T+ / 10 L- / 5 G+]$_A$	[5 T- / 5 L- / 10 G+]$_A$		[T/ L / G]$_A$
AAB	[5 T+ / 10 L- / 5 G+]$_A$	[5 T- / 5 L- / 10 G+]$_A$		[T/ L / G]$_B$
ABA	[5 T+ / 10 L- / 5 G+]$_A$	[5 T- / 5 L- / 10 G+]$_B$		[T/ L / G]$_A$
ABB	[5 T+ / 10 L- / 5 G+]$_A$	[5 T- / 5 L- / 10 G+]$_B$		[T/ L / G]$_B$
Loss-Gain				
AAA	[5 T- / 5 L- / 10 G+]$_A$	[5 T+ / 10 L- / 5 G+]$_A$		[T/ L / G]$_A$
AAB	[5 T- / 5 L- / 10 G+]$_A$	[5 T+ / 10 L- / 5 G+]$_A$		[T/ L / G]$_B$
ABA	[5 T- / 5 L- / 10 G+]$_A$	[5 T+ / 10 L- / 5 G+]$_B$		[T/ L / G]$_A$
ABB	[5 T- / 5 L- / 10 G+]$_A$	[5 T+ / 10 L- / 5 G+]$_B$		[T/ L / G]$_B$

Note. T, L, and G were different fictitious stock companies. The number before each cue denote the number of trials in which the cue was presented. '+' represents *gain* outcomes, and '-' represents *loss* outcomes. Gain-Loss groups received T+ during Phase 1 and T- during Phase 2. Loss-Gain groups received T- during Phase 1 and T+ during Phase 2. Subscripts represent the context in which training took place (Experiments 3).
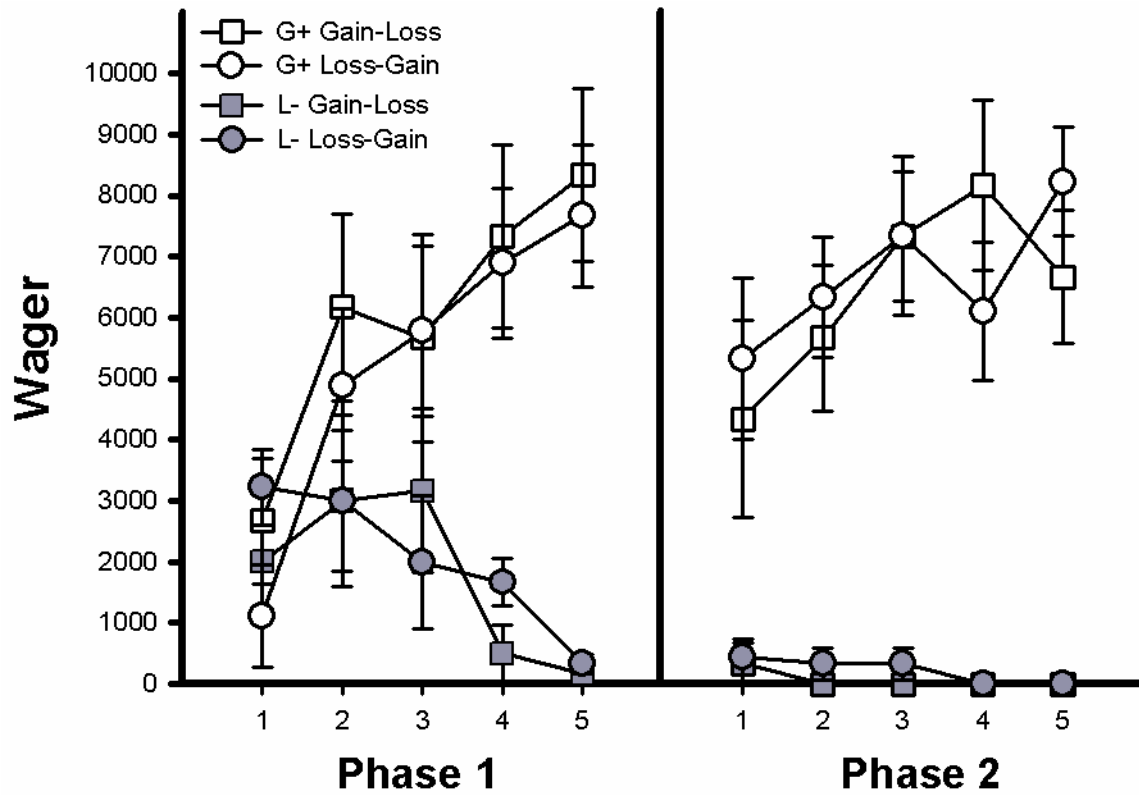
Figure 1. Results of Experiment 1 (Responding to G+ and L-). Phase 1 (original training) and Phase 2 (CC training) present mean wagers made to G+ and L- over trials in each phase of the procedure. For both the Gain-Loss and Loss-Gain conditions, G+ was always followed by the gain outcome, and L- was always followed by the loss outcome. Brackets represent standard error of the mean.
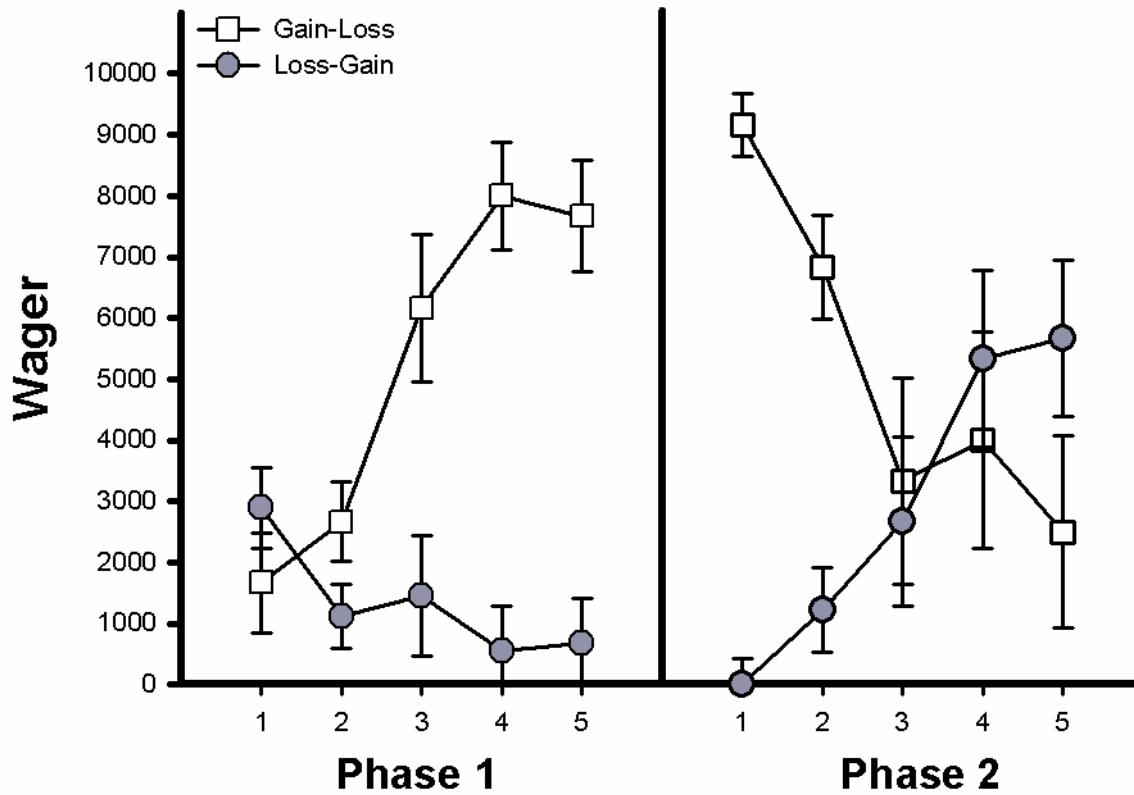
Figure 2. Results of Experiment 1 (CC of T). Phase 1 (original training) and Phase 2 (CC training) present mean wagers made to T over trials in each phase of the procedure. The Gain-Loss condition experienced T paired with the gain outcome in Phase 1 and the loss outcome in Phase 2. The Loss-Gain condition experienced T paired with the loss outcome in Phase 1 and the gain outcome in Phase 2. Brackets represent standard error of the mean.
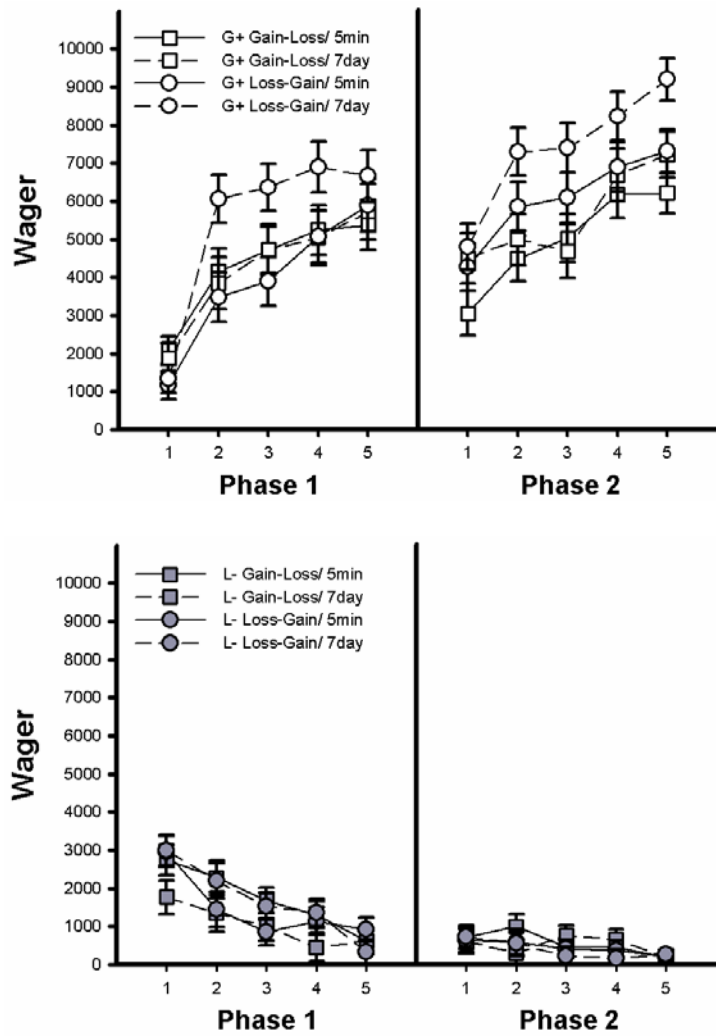
Figure 3. Results of Experiment 2 (Responding to G+ and L-). Phase 1 (original training) and Phase 2 (CC training) present mean wagers made to G+ (upper graph) and L- (lower graph) over trials in each phase of the procedure. For both the Gain-Loss and Loss-Gain conditions, G+ was always followed by the gain outcome, and L- was always followed by the loss outcome. '5min' indicates that the spontaneous recovery (SR) test occurred 5 minutes after completion of Phase 2, while '7day' indicates that the SR test occurred 7 days after completeion of Phase 2. Brackets represent standard error of the mean.
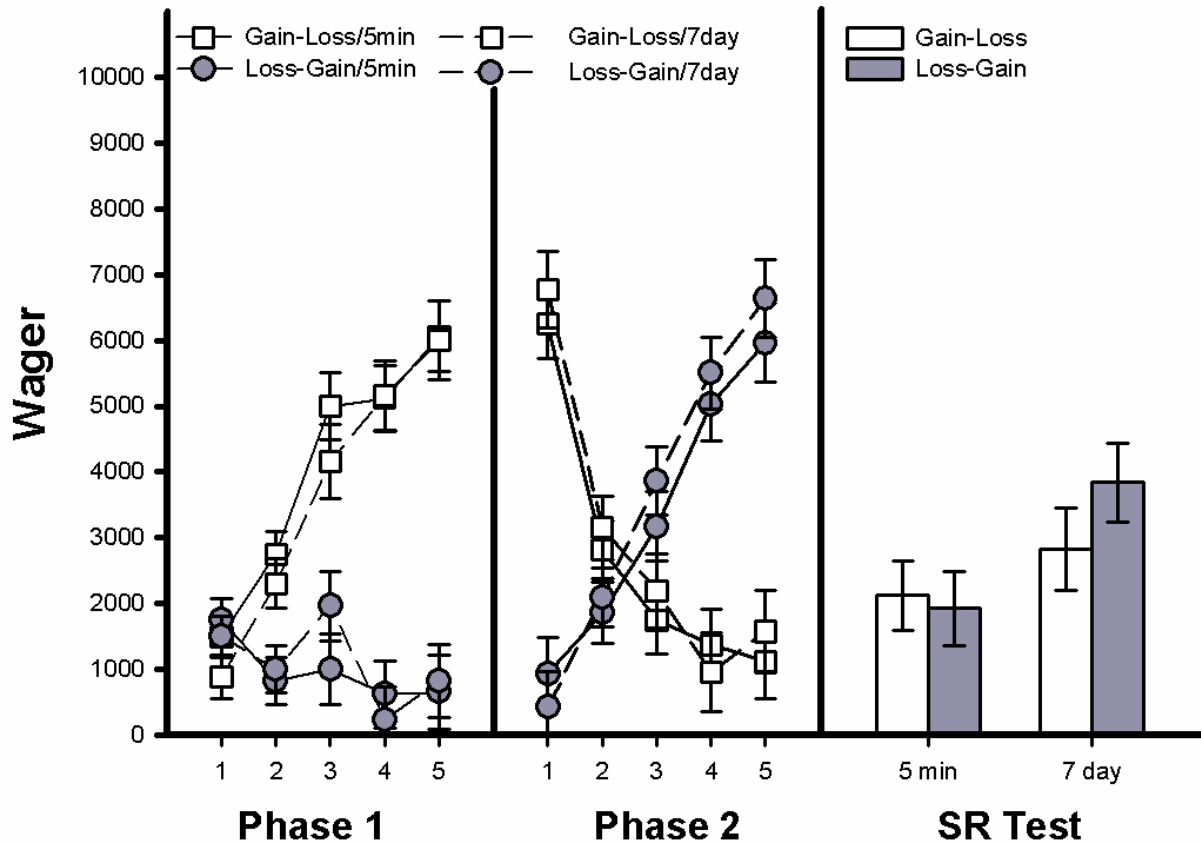
Figure 4. Results of Experiment 2 (CC of T and Spontaneous Recovery Test). Phase 1 (original training) and Phase 2 (CC training) present mean wagers made to T over trials in each phase of the procedure. The Gain-Loss condition experienced T paired with the gain outcome in Phase 1 and the loss outcome in Phase 2. The Loss-Gain condition experienced T paired with the loss outcome in Phase 1 and the gain outcome in Phase 2. '5min' indicates that the spontaneous recovery (SR) test occurred 5 min after completion of Phase 2, while '7day' indicates that the SR test occurred 7 days after completeion of Phase 2. SR Test presents mean wagers made to T over the test trial for the 5min and 7day retention intervals. Brackets represent standard error of the mean.
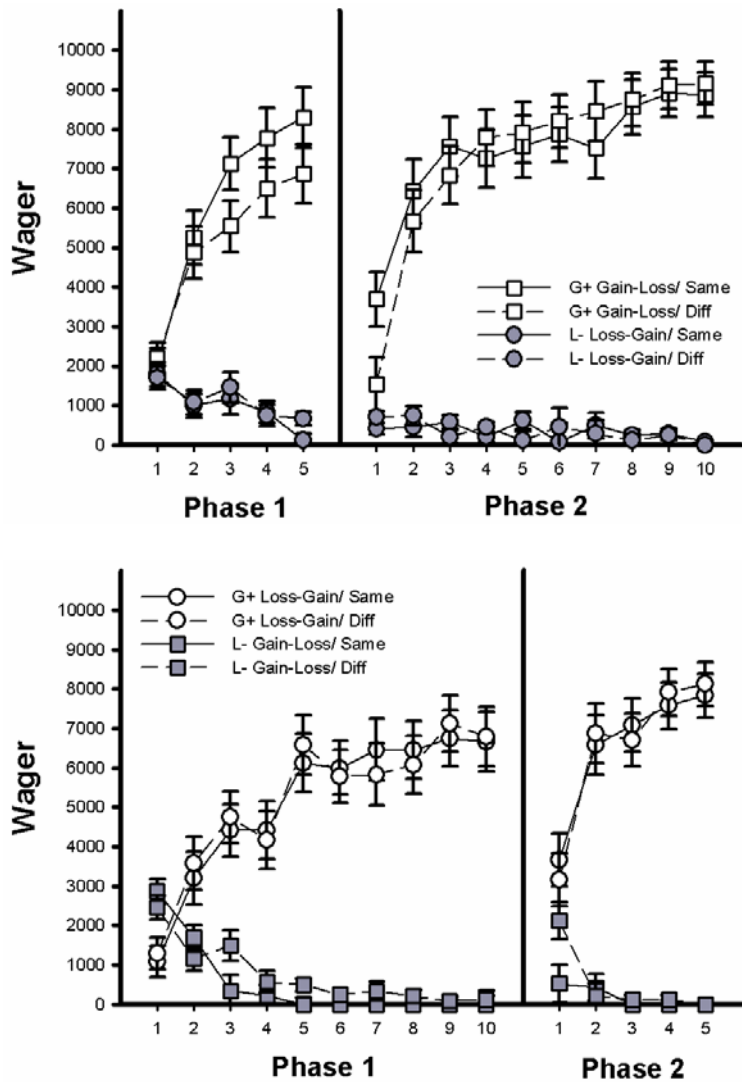
Figure 5. Results of Experiment 3 (Responding to G+ and L-). Phase 1 (original training) and

Phase 2 (CC training) present mean wagers made to G+ and L- over trials in each phase of the

procedure. The upper graph presents data from participants that received 5 G+ or L-

presentations in Phase 1 and 10 G+ or L- presentations in Phase 2. The lower graph presents data

from participants that received 10 G+ or L- presentations in Phase 1 and 5 G+ or L- presentations

in Phase 2. For both the Gain-Loss and Loss-Gain conditions, G+ was always followed by the

gain outcome, and L- was always followed by the loss outcome. 'Same' indicates that Phase 2

occurred in the same context as Phase 1, while 'Diff' indicates that Phase 2 occurred in a

different context from that of Phase 1. Brackets represent standard error of the mean.
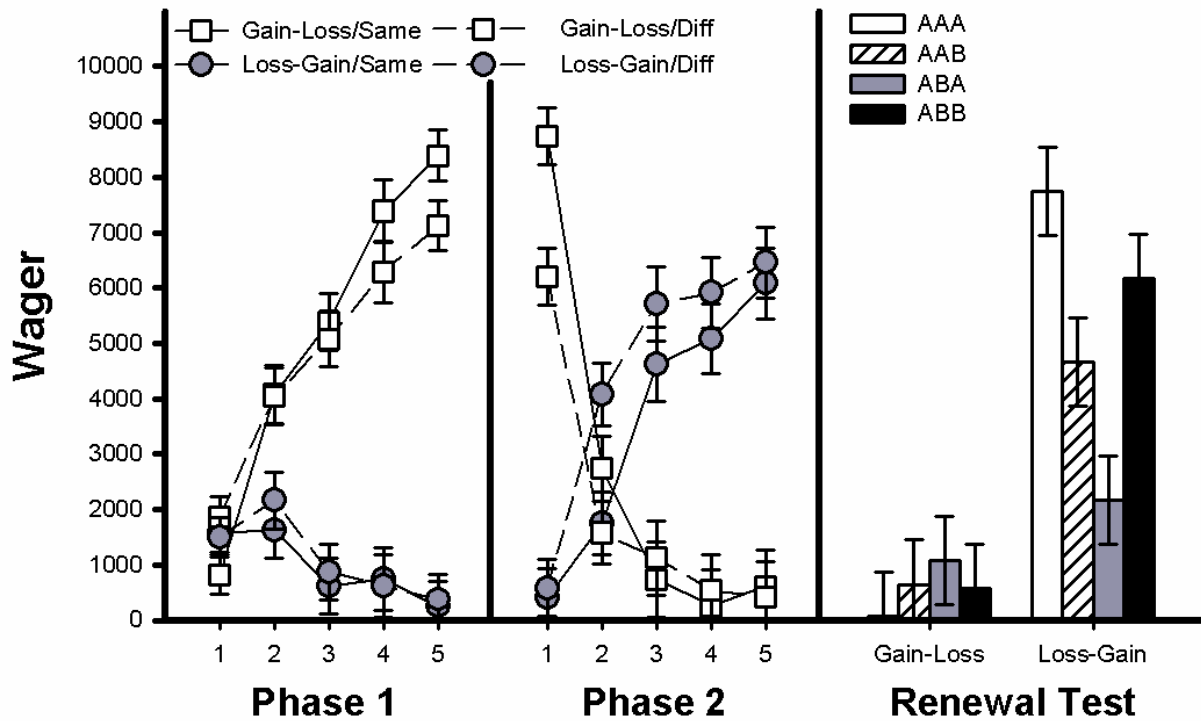
Figure 6. Results from Experiment 3 (CC of T and renewal test). Phase 1 (original training) and Phase 2 (CC training) present mean wagers to T over trials in each of the procedure. The Gain-Loss condition experienced T paired with the gain outcome in Phase 1 and the loss outcome in Phase 2. The Loss-Gain condition experienced T paired with the loss outcome in Phase 1 and the gain outcome in Phase 2. 'Same' indicates that Phase 2 occurred in the same context as Phase 1, while 'Diff' indicates that Phase 2 occurred in a different context from that of Phase 1. Renewal Test presents mean wagers to T over the test trial for the Gain-Loss and Loss-Gain conditions. Each condition name contains three letters that represent the context (A or B) in which Phase 1, Phase 2, and testing took place (see text for details). Brackets represent standard error of the mean.

Appendix B

Cover Story:

Participants will be asked to read the following cover story immediately prior to beginning the

task for Experiments 1, 2, and 3:

> *THANK YOU FOR PARTICIPATING IN THIS STUDY*
>
> *From this point, your total participation time will be of about 20 min. We would*
>
> *appreciate your completing the task to the best of your ability. The computer will guide*
>
> *you through the procedure at all times. At the end of the study, you will see a message*
>
> *indicating that you have completed the task. Please, wait for all other participants to*
>
> *finish the task before leaving the room.*
>
> *You can advance the instruction screens by pressing the right (→) arrow key and return*
>
> *to a previous instructions screen by pressing the left (←) arrow key.*
>
> *Imagine that you are applying for a high-paid job at a financial institution. The position*
>
> *has many applicants, and there will be many tasks that you must complete to be*
>
> *considered for the job. For example, you will need to demonstrate how well you can learn*
>
> *to anticipate the price of some stocks based on market changes.*
>
> *A lot of the company's clients have expressed a desire to invest in pharmaceuticals. Thus,*
>
> *you must demonstrate that you can use market cues to maximize your clients' profits; that*
>
> *is, gain a lot of money and lose little money when investing in stocks of pharmaceutical*
>
> *companies.*

One of the market analysts of the investment company suspects a relationship between the price of certain high-tech companies (the GAMMA, KAPPA, OMEGA, SIGMA, AND THETA companies) and the price of pharmaceutical company stocks. Using information on the price of the high-tech companies, you must determine whether or not to invest in pharmaceutical companies.

The company will let you purchase up to $10,000 in stock each day of the test period. You must be aware that this is a volatile market, in which gains and losses will equal the amount of your investment. That is, if you invest $5,000, you risk making a $5,000 gain or a $5,000 loss.

Each day, you will receive information about the performance of the high-tech stocks the company is watching, and you will be allowed to make your investment. Then, you will receive feedback on whether your investment resulted in a gain or a loss. You must be timely with your investments because the server can time out which will lock you out of purchasing stocks that day.

To make your investment, you must fill the bar located on the right side of the screen by pressing the up (↑) or down (↓) cursor (arrow) keys and then pressing ENTER to confirm your purchase. Remember, your goal is to make as much money as you can by investing more when you anticipate gains and less when you anticipate losses.

Press the SPACE BAR to continue