HIGH SCHOOL EDITION
Journal of Student Research
Volume 11 Issue 4 (2022)

The Effectiveness of Counterconditioning and Extinction in Extinguishing Learning

Pradnya Rajalakshmi[1]

[1]Inventure Academy, India

ABSTRACT

Counterconditioning and extinction are two different forms of learning that inhibit the expression of a learned unwanted behavior by either eliminating the behavior (extinction) or replacing it altogether with a wanted behavior (counterconditioning). While counterconditioning has been shown to be more effective than extinction in some cases, both of these techniques can be vulnerable to the relapse of an unwanted behavior. Separate work has also shown that learning to inhibit unwanted behavior can be more effective if done across multiple different contexts rather than a single experimental context. A systematic comparison of extinction and counterconditioning across single and multiple contexts is, however, lacking. This research paper aims to identify the effectiveness of each of these learning techniques across a single context as well as multiple contexts to identify which is the most suitable for removing unwanted learned behavior. A two-by-two between-subjects design was used wherein four groups completed a learning task in either single or multiple contexts and with either extinction or counterconditioning. It was hypothesized that the counterconditioning group in multiple contexts should turn out to be more effective in extinguishing learning. The results did not support this hypothesis and instead suggest that these methods were all similarly effective at reducing unwanted behaviors.

Introduction

In recent years, extinction and counterconditioning have gained considerable attention in the cognitive science of learning and memory. Counterconditioning refers to replacing a behavior by associating it with a stimulus of the opposite valence. Extinction pertains to the complete removal of the stimulus, thus extinguishing the behavior entirely. It is imperative to acknowledge the strengths and limitations of extinction and counterconditioning in order to understand their various applications in learning and memory. Each technique is widely used to eliminate an aversive response or reaction. This type of learning, when introduced across multiple paradigms, would yield different results as compared to when introduced to solely the experimental paradigm. Counterconditioning and extinction have both been researched for therapeutic applications. Previous research has proved that counterconditioning could be more effective than extinction in removing an unwanted response. Raes and Raedt (2012) sought to determine whether counterconditioning can serve to reduce evaluative responses within fear conditioning. Fear conditioning refers to the pairing of the conditional stimulus with an aversive unconditioned stimulus that creates a response. Seventy participants were recruited, of which one-third were allocated to extinction, one-third for counterconditioning with the neutral stimulus (neutral tone), and one-third for counterconditioning with the positive stimulus (baby laugh). During the acquisition phase, the participants were exposed to distressing white noise as the unconditioned stimulus and a human face as the conditioned stimulus. The results concluded that both the counterconditioning groups with the neutral and positive stimuli were significantly more effective in reducing the unwanted response. One way in which the efficacy of counterconditioning versus extinction can further be analyzed is by testing each technique in either single or multiple contexts.

Introducing multiple contexts ensures that reinstatement of the aversive stimulus is eliminated in a range of more general settings than the single context used in the experimental paradigms. In general, extinguishing learned negative associations should be more effective in multiple contexts than in single contexts. It is, however, currently unknown whether the beneficial effects of counterconditioning and multiple contexts, when combined, interact to eliminate reinforcement, although previous research has provided enough evidence that that should be the case. Here, we sought to answer this question. We hypothesized that counterconditioning should be more effective than extinction in eliminating aversive learned responses. Specifically, counterconditioning in multiple contexts should extinguish learned aversive responses more than either of these individually, and also either extinction in multiple contexts or counterconditioning in a single context.

We first looked at previous research to analyze counter conditioning and extinction paired with single and multiple contexts. The famous little peter experiment established the comparison of counterconditioning with extinction and solidified the effectiveness of counterconditioning as a learning method. The experiment conducted by Jones in 1924 demonstrated the modification of behavior by replacing the unwanted response (fear) with an outcome of a positive appetitive stimulus. A rabbit (conditioned stimulus) was placed in a room far away from him while he was eating candy (appetitive unconditioned stimulus). After a certain period of time, the rabbit was moved closer to him until, eventually, he allowed the rabbit to nibble on his fingers without fear (conditioned response). The unconditioned stimulus was stronger than the conditioned stimulus in order to elicit fear removal. In this case, the delight peter got from eating the candy was significantly stronger than his fear of rabbits. This experiment provided evidence of the effectiveness of counterconditioning over extinction in reducing fear.

Despite this evidence demonstrating greater effectiveness for counterconditioning over extinction, some research on this topic has provided contradictory evidence. For example, studies have shown that counterconditioning and extinction are equally effective in reducing fear responses (G.T. Wilson, 1973). In this study, two experiments were conducted. The first experiment focused on eliminating a conditioned avoidance response by utilizing counterconditioning and response prevention techniques. The second experiment extinguished the conditioned avoidance response by maintaining the conditions for counterconditioning and response prevention for the same duration as the conditioned stimuli. The results showed that both methods were equally effective in decreasing conditioned fear.

The efficacy of single vs. multiple contexts was also researched previously. Shiban and colleagues (2013) focused on the exposure of multiple context extinction to reduce fear renewal. They compared this with extinction in a single context and noticed that renewal was significantly reduced in multiple contexts. These conclusions strongly support the statement that extinction in multiple contexts is far more effective than single context extinction. Gunther and colleagues (1998) analyzed the effects of extinction in multiple contexts by pairing the unconditioned and conditioned stimulus in one context, followed by the extinction of the conditioned stimulus in three other contexts. The experiment concluded that the subject conditioned in multiple contexts displayed a stronger response to the conditioned stimulus than subjects conditioned in a single context.

The body of work on this topic has raised several questions that the present work aims to answer. Under what circumstances is counterconditioning more effective than extinction at preventing relapse to an aversive stimulus? How effective are counterconditioning and extinction in single and multiple contexts? The present study found no difference between counterconditioning and extinction and no impact of the number of contexts on these. The results of the experiment concluded that single context counterconditioning was however effective in extinguishing learning. Although the results drawn from the experiment don't support the hypothesis, they raise questions for future discussions in extinction and counterconditioning.

Methods

Table 1: Single and multiple context extinction and counterconditioning design

	Context	Acquisition	Post-Acquisition	Test
Single Context Extinction	1	A+, F1-	F4+, F5-	
	2	F2+, F3-	A-, F6+	A?
	3		F7-	
	4		F8-	
	5			A?
Multiple Context Extinction	1	A+, F1-	F4+, F5-	A?
	2	F2+, F3-	A-, F6+, F7-, F8-	
	3		A-	
	4		A-	
	5			A?
Single Context Counterconditioning	1	A+, F1++	F4+, F5++	
	2	F2+, F3++	A++, F6+	A?
	3		F7++	
	4		F8++	
	5			A?
Multiple Context Counterconditioning	1	A+, F1++	F4+, F5++	A?
	2	F2+, F3++	A++, F6+, F7++, F8++	
	3		A++	
	4		A++	
	5			A?

Note: The contexts are the different pet stores; F1-F9 are the different pets (consisting of positive and aversive stimuli); + is the occurrence of the cue, and - is the nonoccurrence of the cue. ++ depicts the occurrence of counterconditioning.

Table 1 demonstrates the design for the single and multiple context extinction. In the first context for single context extinction, during the acquisition phase, participants in context 1 were exposed to the experimental cue's occurrence and the control cue's nonoccurrence. The post-acquisition phase consisted of the occurrence of two other controlled cues. Contexts 3 and 4 consisted of the nonoccurrence of the controlled cues. In multiple context extinction, the participants for context 1 were familiarized with the same experimental cue and control cues as participants in single context extinction. Additionally, the participants were presented with A- in contexts 2,3,4. The test was conducted during contexts 2 and 5 for single context extinction and 1 and 5 for multiple context extinction.

Experimental Design

The study was a two-by-two between-subject design that analyzed two different variables: single context or multiple contexts and extinction or counterconditioning. Four different groups were present: Single context counterconditioning, multiple context counterconditioning, single context extinction, and multiple context extinction. The contexts were the doors, and the stimuli were the animals. The experiment consisted of three phases that followed a standard extinction-learning procedure. Phase 1 (acquisition phase) consisted of 4 stimuli and a total of 48 trials (12 trials for each stimulus). In this phase, participants learned associations between the stimuli and outcomes. The second phase or the post-acquisition phase, consisted of 6 stimuli with 72 trials in total (12 trials of each stimulus). Here, participants learned new associations between stimuli and outcomes. In the last phase or the test phase, participants were shown a single stimulus and were asked to respond. Feedback on the correctness of their response was not provided. For the multiple context groups, the experimental cue was spread across the contexts; there were four trials for each context. The total added up to 128 trials. During the acquisition and post-acquisition trials, the prediction screen lasted for 3000 ms, wherein a context and the stimulus were shown. The outcome screen lasted for 1000 ms, where the feedback was displayed. The task took 15 minutes in total to complete.

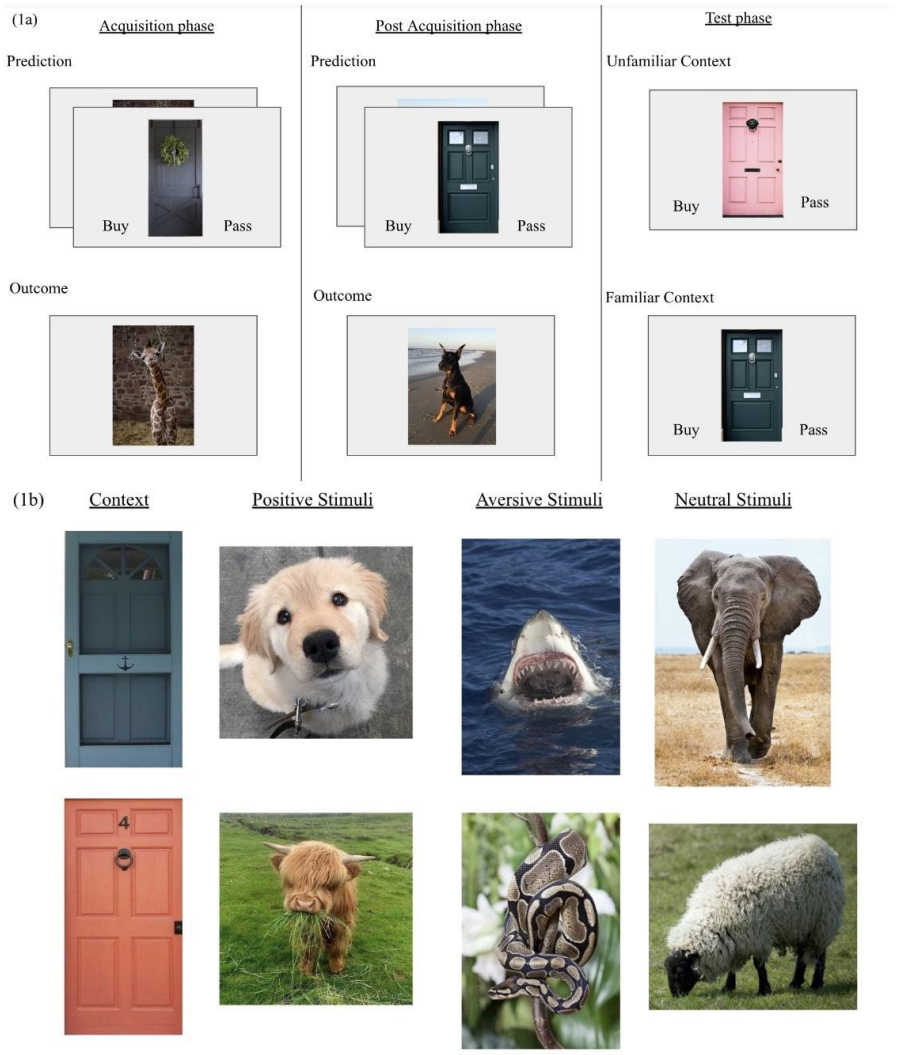

Figure 1: Design of the experiment. The experiment is split into three phases— The acquisition phase, Post Acquisition phase, and the test phase. In the Acquisition and post-acquisition phase, the prediction screen depicts the outcome hidden by a door (context). The participants are required to either select buy or pass, and the outcome screen displays whether the answer is correct. The final test phase displayed the prediction screen, and the outcome was not shown to participants. **(A)** outlines the experiment, and **(B)** depicts the contexts and stimuli. The positive stimuli were used for counterconditioning, the aversive stimuli for extinction, and the neutral stimuli for both.

Participants

There were 80 participants; 20 participated in each group. The subjects were recruited from prolific, an online platform that aids in launching the experiment to targeted groups of participants. Their ages varied between 18 and 35 years. Each participant was paid approximately $2. 50% of the participants were males, and 50% were females to ensure that the results were standardized across gender. The participants were given a set of instructions to read before starting the experiment. They were required to score 100% on a quiz testing their knowledge and understanding of these instructions prior to beginning the experiment.

Ethics

The procedure and design of this experiment was approved by a Scientific Review Committee (SRC) consisting of a high school educator, a medical professional, and a scientist who were all unaffiliated with the project. The SRC reviewed and provided approval for the experiment prior to experimentation and followed the guidelines of the Intel International Science and Engineering Fair (ISEF) rules.

Procedure

They were given a cover story that asked them to picture visiting a petting zoo, where they were told that they had to select the appropriate animals for the zoo. A colored door was shown to the participant, behind which the animal was present. The participant had to click on the 'F' key to pass and the 'K' key to buy the animal. There were five different contexts that were represented by the doors. For the acquisition and post-acquisition phase, the stimuli (animals) were shown as soon as the participant plugged in their response. The participants had to respond within 1500 ms.

Analysis

All data were analyzed using mixed-effects logistic regressions. First, we assessed the extent to which learning occurred by analyzing the proportion of bought and passed trials in the second half of the acquisition phase. We focused on the second half of the trials because we reasoned that choices at this point should represent learned preferences. Using *lmer*-style syntax, a mixed-effects logistic regression model was used to predict whether a buy decision was made. This model can be written as:

$$P(Buy\ Chosen) \sim ContextType * PostAcquisitionType + (1 \mid Participant)$$

Such that the predictor *ContextType* represented the context that a group completed the experiment in (coded as 0.5 if a participant was in the multiple context condition and -0.5 if a participant was in the single context condition) and the predictor *PostAcquisitionType* represented the type of post-acquisition that a group completed the experiment in (coded as 0.5 if a participant was in the counterconditioning condition and -0.5 if a participant was in the extinction condition). An intercept was included to capture the degree to which participants made decisions at a level that was different from chance, and this intercept was allowed to vary as a random effect for each participant. Fixed main effects of *ContextType* and *PostAcquisitionType* were included in the model alongside the fixed interaction effect between them. An identical mixed-effects logistic regression model was used to analyze the second half of the post-acquisition to assess the extent of extinguished behavior that resulted from each post-acquisition condition.

To address our primary question, which involved assessing the extent to which the extinguished behavior carried over into the test phase, we used two identical mixed-effects logistic regression models to predict buy choices in response to test stimuli in the test phase. In the first model, we included only test choices made in a familiar context, and in the second model, we included only test choices made in an unfamiliar context. In total, across all phases of the experiment, four regressions were conducted: one to analyze acquisition choices, one to analyze post-acquisition choices, and two to analyze test choices.

Results

These results clarify the effectiveness of counterconditioning and extinction in learning and memory across single and multiple contexts. We hypothesized that out of the four groups (single context counterconditioning, multiple context

counterconditioning, single context extinction, and multiple context extinction), counterconditioning in multiple contexts should turn out to be more effective in extinguishing learning. The experiment did not support the claims made in the hypothesis.

We first looked at the results of the acquisition phase in order to determine whether stimulus-outcome associations were learned (Figure 2; Left Panels). Participants in all groups were above chance performance as indicated by the intercept in our mixed-effects model of choice (β=0.11, 95% CI [0.01, 0.22], p = 0.039). Participants thus learned the associations between each stimulus and choice. We next assessed whether there were any differences between the groups. There was no main effect of context (β=0.11, 95% CI [-0.09, 0.31], p = 0.289), indicating that completing the acquisition phase in neither the single context condition nor the multiple context condition affected learning. There was also no main effect of cue type (β=-0.07, 95% CI [-0.27, 0.14], p = 0.527), indicating that there was no impact of neutral or positive cues on learning during the acquisition phase. Finally, there was also no interaction between cue type and context condition (β=-0.13, 95% CI [-0.54, 0.28], p = 0.527). This lack of interaction indicates that neither of these variables depends on the other to alter learning. Thus, overall, participants were able to learn the proper choices to make for each cue; however, there was no effect of our conditions on learning during the acquisition phase.

We then looked at the results of the post-acquisition phase to identify whether the stimulus-outcome associations were unlearnt (Figure 2; Right Panels). Participants in all groups were above chance, as indicated by the intercept in the mixed-effects model of choice (β=0.62, 95% Cl [0.52,0.72], p<0.001]. This showed that participants learned new associations between the stimulus and choice. We then assessed whether there were any differences between the groups. There is no main effect of context since the p-value is greater than 0.06, and the confidence interval is negative on one side and positive on the other. (β=-0.13, 95% Cl [-0.32,0.06], p=0.18). There was no main effect of cue type (β=-0, 95% Cl [-0.19,0.19], p=0.96), showing that learning was not affected by counterconditioning or extinction. The interaction between the context type and cue type (β=-0.37, 95% Cl [-0.75,0.01], p=0.06) showed that the single context counterconditioning group extinguished more, although we expected the multiple context counterconditioning group to do the same. Since the p-value is 0.06, this is a marginal effect.

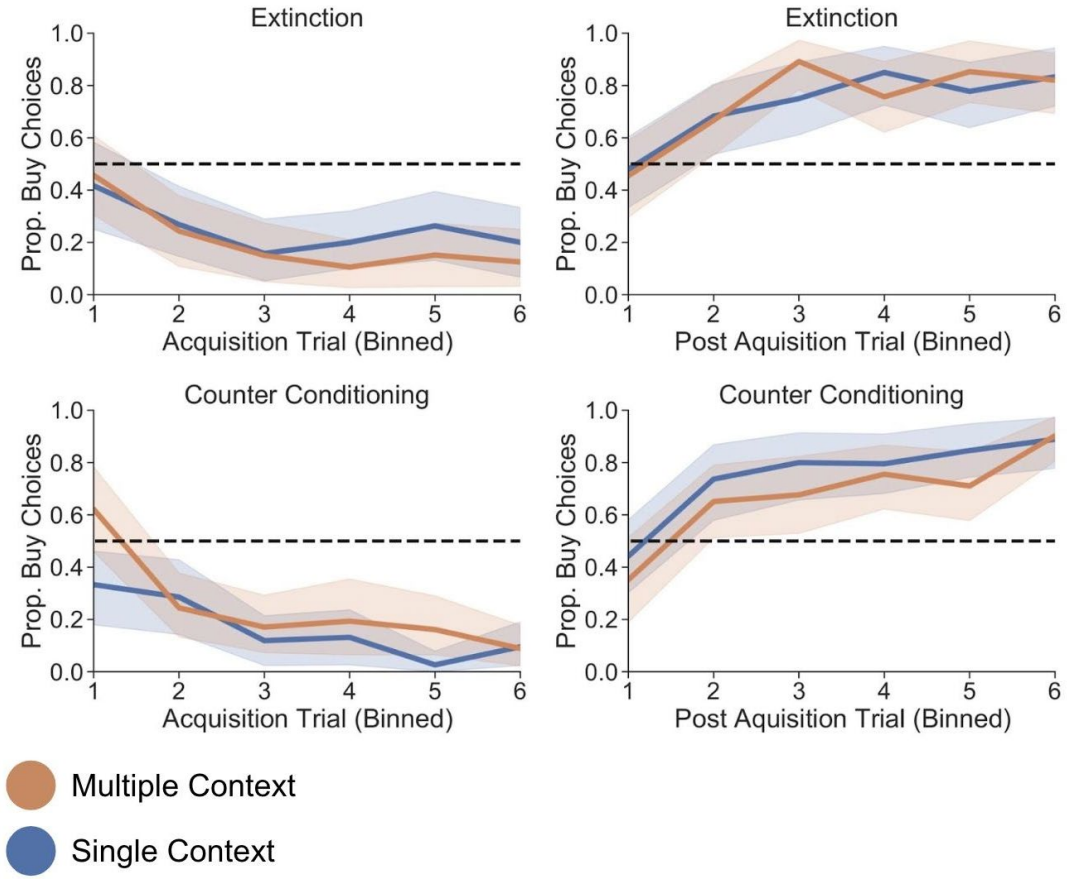

Figure 2: Acquisition and Post-Acquisition results. Depicts the results of the acquisition and post-acquisition phases in single and multiple contexts. Chance was at 0.5, which is represented by the black dotted line. The blue line represents the average proportion of Buy choices across the entire single context condition, and the orange line represents this information for the multiple contexts condition. Bands around lines represent 95% confidence intervals.

After establishing that participants were able to learn the extinguished response in the post-acquisition phase, we moved on to examine the test phase (Figure 3). The effects of context and cue type on familiar contexts were analyzed first. Participants in all groups chose the above chance, as shown by the intercept in the mixed-effects model of choice (β=5.56, 95% Cl [0.02,11.09], p=0.049). There was no significant effect of interaction between the context type and the cue type (β=0.004, 95% Cl [-3.83,3.84], p=0.99). There was no main effect of cue type (β=1.13, 95% Cl [-1.07,3.32], p=0.31) and no effect of context type (β=-1.44, 95% Cl [-3.34,0.47], p=0.14). Overall, in terms of the familiarity of the choices, there was no indication of multiple context extinction being more effective, which was the opposite of what we had predicted.

We then assessed the results of test choices made in unfamiliar contexts. Participants in all groups again chose the above chance, as shown by the intercept in the mixed-effects model of choice (β=2.03, 95% Cl [1.16,2.89], p<0.001). There was no main effect of context type (β=0.48, 95% Cl [-0.57,1.52], p=0.37). There was also no main effect of the cue type (β=-0.63, 95% Cl [-1.82,0.55], p=0.29). There was also no effect of the interaction between the two (β=-0.33, 95% Cl [-2.57,1.91], p=0.77). The results of the test choices showcased that there was no effect of the manipulation on the test choices.

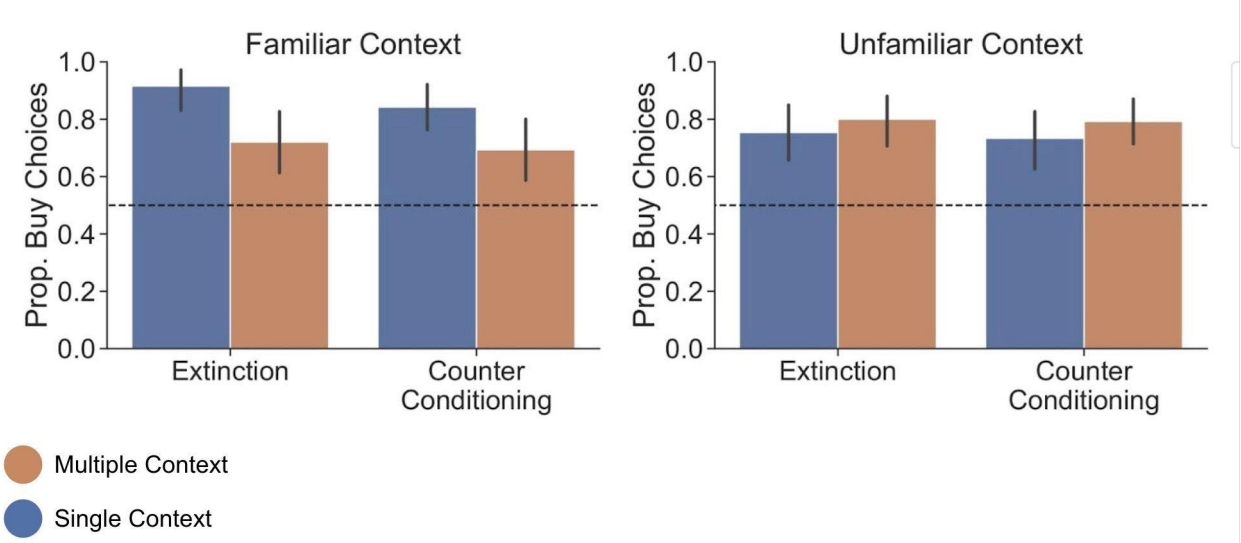

Figure 3: Familiar and unfamiliar contexts results. Shows the results of the familiar and unfamiliar contexts for extinction and counterconditioning in single and multiple contexts. Chance was at 0.5 which is represented by the black dotted line in each plot. The blue bar represents the group average proportion of Buy choices in a single context, and the orange bar represents this information for multiple contexts. Error bars represent 95% confidence intervals.

Discussion

Understanding the role of counterconditioning and extinction in learning and memory is extremely vital. The experiment conducted here analyzes the effects of the two in single and multiple contexts, which would prove useful when it comes to therapeutic applications. For example, these techniques could be used to eliminate the unwanted or aversive response to certain stimuli. We hypothesized that participants who completed counterconditioning across multiple contexts would extinguish learning more effectively than the other three groups. The results showed that all the groups extinguished their responses equally, indicating that none of these manipulations had a significant effect.

Previous research had concluded that counterconditioning was far more effective than extinction and using multiple contexts instead of the experimental paradigm would yield the best results. Based on this, we hypothesized that counterconditioning in multiple contexts would be the most effective learning method. However, the results of the experiment demonstrated that all groups extinguished the unwanted response equally.

One possible reason as to why the results did not support the hypothesis could be because of limitations in the experimental design. The number of participants in this study was relatively low, which decreases the reliability of the results. Increasing the number of participants would result in a wider range of data collected, therefore increasing the validity of the experiment. Extinction and counterconditioning in single and multiple contexts should be explored more extensively with a larger sample size as it may lead to promising results.

Although the experiment was designed such that external distractions were limited as much as possible, participants still completed the task online, which meant that there was limited control over the experimental conditions. The participants had to answer a few questions at the beginning of the experiment to ensure that they were paying 100% attention. They were not able to access the experiment unless all the questions were answered correctly. Distractions and other miscellaneous interruptions could be further avoided by conducting the experiment in an in-person, controlled setting.

Often, in studies experimenting with learning techniques, more salient unconditioned stimuli are utilized (shocks, money, and so on) in order to identify a more significant response; the experiment could be made more impactful by making use of more prominent stimuli. For instance, if the aversive stimuli were replaced with shocks, perhaps a stronger response could be identified.

Counterconditioning and extinction individually are extremely beneficial learning techniques. Previous studies have mentioned that both extinction and counterconditioning could be used in therapeutic applications. Now that all the combinations of extinction vs counterconditioning and single vs multiple have been established, single context and multiple context could be integrated in the learning techniques which could turn out to be extremely effective in practical applications. More research conducted on this would definitely lead to the advancement of learning and extinguishing behavior.

Ideally, counterconditioning in multiple contexts should be the most effective learning technique for extinguishing responses. Individually, other studies have demonstrated that both counterconditioning and the use of multiple contexts is far more effective at extinguishing unwanted behavior. The results of the present study conclude that although counterconditioning in multiple contexts can be used as a learning technique, it isn't more effective than the other three learning methods. The conditions under which these experiments are conducted need to be carefully controlled and monitored. This concept could be evaluated further in the future, leading to better applications for the learning techniques.

Acknowledgements

The author would like to acknowledge the Lumiere program as well as Jonathan Nicholas (mentor) for his support and guidance throughout the research and writing process. Without his attention to detail, knowledge, and enthusiasm, this paper couldn't have been written.

References

Dunsmoor, J. E., Niv, Y., Daw, N., & Phelps, E. A. (2015). Rethinking Extinction. *Neuron*, *88*(1), 47–63. https://doi.org/10.1016/j.neuron.2015.09.028

Keller, N. E., Hennings, A. C., & Dunsmoor, J. E. (2020). Behavioral and neural processes in counterconditioning: Past and future directions. *Behaviour research and therapy*, *125*, 103532. https://doi.org/10.1016/j.brat.2019.103532

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(4), 757–767. https://doi.org/10.1016/j.beth.2012.03.012

Keller, N. E., & Dunsmoor, J. E. (2019). The effects of aversive-to-appetitive counterconditioning on implicit and explicit fear memory. Learning & memory (Cold Spring Harbor, N.Y.), 27(1), 12–19. https://doi.org/10.1101/lm.050740.119

Shiban, Y., Pauli, P., & Mühlberger, A. (2013). Effect of multiple context exposure on renewal in spider phobia. Behaviour research and therapy, 51(2), 68–74. https://doi.org/10.1016/j.brat.2012.10.007

Dunsmoor, J. E., Ahs, F., Zielinski, D. J., & LaBar, K. S. (2014). Extinction in multiple virtual reality contexts diminishes fear reinstatement in humans. Neurobiology of learning and memory, 113, 157–164. https://doi.org/10.1016/j.nlm.2014.02.010

Bustamante, J., Uengoer, M., Thorwart, A., & Lachnit, H. (2016). Extinction in multiple contexts: Effects on the rate of extinction and the strength of response recovery. Learning & behavior, 44(3), 283–294. https://doi.org/10.3758/s13420-016-0212-7

Cover Jones, Mary (1924). "A Laboratory Study of Fear: The Case of Peter". *Pedagogical Seminary*. **31** (4): 308–315. doi:10.1080/08856559.1924.9944851