**Learning from punishments**

Ran Aaron Cohen (aharon.cohen2@mail.huji.ac.il)

Department of cognitive science, The Hebrew University of Jerusalem,

Mount scopus, Jerusalem, Israel

# Abstract

Studies suggest that people's sensitivity to punishment may change from day to day, and such changes could impact mood and mental disorders. Of particular importance may be the sensitivity to primary aversive stimulation such as pain or loud noises (as opposed to secondary, conditioned punishment like monetary loss) since this is the most basic and natural form of punishment in animals and humans. However, experimentally testing learning from primary aversive stimulation is difficult even in a single day lab experiment, let alone in an experiment that aims to assess and re-assess such learning over multiple days, which can be complicated by the problem of habituation to the experimental punishment. The goal of this project was to develop an experimental paradigm that can probe learning from primary punishment over multiple days. For this purpose, we developed a phone-based punishment learning task that administers punishment via headphones. The task allows quantifying both sensitivity and responsivity to punishment, with the former manifesting in the degree to which subjects tend to avoid experimental stimuli previously associated with punishment, and the latter in heartrate responses to punishment as measured through a wearable sensor. The experiment was conducted outside of the laboratory and was monitored remotely by the experimenter. 6 subjects performed an experimental session on their mobile phone twice a day, for 12 days. For comparison, another group of 6 subjects performed the experiment with secondary punishments, which consisted of monetary losses. On each session, subjects had to learn via trial-and-error the probabilities with which choices of different images (textured circles) led to punishment. Primary punishment consisted of a loud noise played through headphones put on during the sessions. To ensure that participants received the punishments properly, an attention listening task was administered randomly between trials. On each session subjects learned about new stimuli. To assess how well learned information was remembered, on some trials subjects were asked to choose between images they had already learned about on previous days. To evaluate whether the experiment successfully probed learning from punishment over the 12 days of the experiment, we examined the degree to which subjects successfully avoided images that were associated with larger probability of punishment. As an implicit and more direct measure of the emotional impact of the punishment, we also measured subjects’ heartrate responses to outcomes in the learning task. Subjects’ choices showed that they successfully avoided stimuli that were more likely to be associated with loud noise, even a few days after learning, which indicates successful and enduring conditioning. Moreover, almost no decline in choice accuracy was observed during the progress of experiment, which suggests low habituation to the experimental punishment. Analysis of subjects’ heartrate showed consistent responsivity to the loud white noise as compared to no punishment outcomes. The small sample size of this preliminary study did not allow conclusive comparison between the primary and secondary punishment groups. However, visual inspection suggested that choice sensitivity and heartrate responsivity to loud noise were at least as strong and enduring as sensitivity and responsivity to monetary loss.

These results confirm that the experimental procedure that we developed constitutes a valid and effective paradigm for investigating sensitivity and responsivity to punishment outside of the lab over multiple days. Furthermore, using this paradigm, future studies could probe the relationship between mood fluctuation and aversive learning mechanisms.

**Keywords:** Operant instrumental learning, Loud white Noise, Monetary loss.

### **Content**

* **Introduction** 
  + Punishment definition
  + Punishment and reward in operant instrumental learning – two sides of the same coin.
  + Differences between punishment and reward in literature
  + Punishment types - Positive & primary VS Negative & secondary punishment
  + Other studies
  + Our study
* **Materials & Methods** 
  + Participants
  + Materials/Stimuli
  + Procedure
  + The game
* **Results**
* **Analysis**
* **Discussion**
* **Conclusion**
* **references**

# Introduction

Punishment is considered a complementary to appetitive reinforcement (*Jean-Richard-Dit-Bressel et al 2018*). ***Conceptually***, they differ in the direction of change in the probability of choices. While Reward ***increases*** the likelihood of a behavior to occur, punishment ***decrease*s** it (*Azrin and Holtz 1966; Johnston, J. M. 1972*).

Eldar el al (2018) found that responsivity to reward prediction errors changes from day to day and these changes interact with mood fluctuations. This finding drove us to believe that a parallel mechanism may exists also in the responsivity to punishment.

## There are different types of reinforcers. One distinction would be between positive and negative punishments (*Franzoi, S. L. 2015*). A positive punisher is the appearance of an undesirable or aversive stimulus contingently after an operant response. Some examples in research of positive punishers will be the delivery of an air puff, electric shock, and loud noise. Conversely, a negative punisher is the removal of an appetitive stimulus such as food or money (*Jean-Richard-Dit-Bressel et al 2018*).

## Another distinction in the literature is of primary and secondary aversive stimulus (*Franzoi, S. L. 2015*). The primary punisher is one that the agent instantaneously perceives as aversive and unpleasant without being conditioned to it. Examples of this type are the delivery of an air puff, electric shock, and loud noise that create an immediate repulsion when encountered. Conversely, a secondary punisher is one that needs to be conditioned and learned, such as a parking ticket, monetary loss, and social punishments like an angry or disapproving face.

## Primary punishments are biologically aversive and naturally perceived in animals and human alike. Therefor it was interesting for us to probe the attributes of this basic form of reinforcer as opposed to a secondary punishment, like monetary loss. Especially important is to probe day to day sensitivity and responsivity to punishment. Sensitivity to punishment is the transformation from objective value into subjective utility while responsivity reflects how much attention is given to the dimension of punishment. We operationalized these two aspects as the degree to which subjects tried to avoid images associated with punishments and heartrate responses to the aversive stimulus, respectively.

## To carry out a longitudinal experiment on learning from primary aversive punishment we applied a novel mobile platform that can be used by subjects outside of the laboratory. This kind of design has proven reliable (Seow & Hauser 2021) and has many advantages but also a few challenges (Reips 2000). One of them the lack of experimental control, where experimenters should regulate all the experiment's aspects to avoid confounds. Another concern we had to address was that subjects will gradually habituate to the aversive punishment delivered.

## Subjects' data was uploaded and stored in a secured location every few hours. This data was reviewed by the experimenter regularly and a few reliability tests were made such as checking the reaction time was standard, side bias, performance, and tasks time schedules. To make sure that subjects wore the earphones during the learning game and that the hear through them the aversive noise (the sound volume was set at the beginning of the experiment and superimposed on the phone system), we added randomly between trials a task that proves they are listening and attentive to sounds delivered through the earphones.

*Delgado et al* conducted experiments to probe the effects of monetary loss as a secondary reinforcer and to examine the differences between a primary and secondary punisher in a fear conditioning paradigm. They found that the striatum has an important role in monetary loss punisher (secondary) as in mild shock punisher (primary). Interestingly, the amygdala was activated only in the mild shock condition. They concluded that learning from monetary losses may depend on reinforcement learning mechanisms whereas primary punishers rely more on biological mechanisms. Importantly, they did not find a significant difference between primary and secondary punishment in the acquisition of conditioned responses (*Delgado et al 2011*; *Delgado et al* 2006). In a different study, *Delgado et al* show punishment prediction error signals in the striatum both in primary and secondary punishers. (*Delgado et al* 2008).  
Interestingly, *Delgado et al* observed similaritiesbetween negative reward prediction error signal and monetary loss signal(*Delgado et al 2000*).

Like Electric Shock, Loud White Noise is a common unconditioned stimulus (US) used in punishment conditioning research (Sperl et al 2016). For experiments with many conditioning trials, Sperl et al asked which of the two US will cause a long and strong Conditioned Stimulus, that will satisfy an EEG and MRI test, and will be strong enough to avoid extinction. They designed a comparison study to test them and concluded that Loud White Noise had greater valence of unpleasantness, less extinction of Conditioned Response (CR), and a better recall of the CR after 24h (Sperl 2016). For this reason we used loud white noise as the primary punishment in our study.

Although punishment has an important role in learning processes and its extensive potential implications for psychiatric disorders (*Jean-Richard-Dit-Bressel et al. 2018; Wise & Dolan 2020*), we are still in ignorance with respect to some of its traits, neuronal and computational mechanisms, and its precise influence on human behavior (*Jean-Richard-Dit-Bressel et al. 2018; Wise & Dolan 2020*). Thus, the investigation of punishment is crucial and has a promising, fruitful prospect.

Creating new paradigms for probing punishment and contrasting it to reward is of great importance as it can help solving some open questions in the study of aversive learning.

Two theories offered to describe the **effect** of punishment on choices: The additive theory and the subtractive theory. The **additive** theory claims that punishment reinforce another stimulus that avoids the punishment, therefor reducing the occurrences of the punished stimulus, whereas the **subtractive** theory claims that punishment suppresses the appetitive attraction of a stimulus and thus causing a reduction in the probability of choosing that punished stimulus (Toshikazu et al, 2018). Furthermore, in the subtractive school there is a long dispute about the symmetrical nature of reward and punishment (*Rasmussen and Newland 2008*). *Rasmussen and Newland* concluded that punishment has a greater effect on choices than reward and therefor they are not symmetrical: a penny you lose perceived greater than a penny you earn. Interestingly, Palminteri et al show, that over time, in a punishing context, punishment shifts to a reward system of reinforcement. The agent starts associating the non-punishing response with reward and the punishing response with no-reward (Palminteri et al 2015). Another puzzling question is whether reward and punishment have different neural systems in the brain, and a few hypotheses were suggested to address it.

Despite all the parallels between reward and punishment, they differ **in some important characteristics** (*Jean-Richard-Dit-Bressel et al 2018*). The activation and nonactivation of dopaminergic neurons in the ventral striatum in the brain are classically attributed to reward mechanisms of positive reward prediction error and negative reward prediction error, respectively (*Schultz 2007*), whereas in aversive learning mechanisms the amygdala is more widely implicated (*Eldar et al. 2016a; Toshikazu et al. 2018; Jean-Richard-Dit-Bressel et al 2018*; *Costafreda et al 2008*). For example, Michely et al show that subregions in the basolateral amygdala encodes a punishment prediction error (Michely et al 2020). Furthermore, with respect to neurotransmitters, Dopamine is associated with reward prediction errors whereas Serotonin is implicated in punishment prediction errors (Cools et al 2008).   
Moreover, there is evidence that punishment has a different influence on learning and behavior than reward. Steel et al found that punishment had greater effect on learning in both sequencing skill task (SRTT) and motor skill task (FTT), although the effect on the SRTT task was positive whereas the effect on the FTT task was negative (Steel et al 2016). Galea et al. found that punishment produced faster learning for motor adaptation whereas reward caused greater retention (Galea et al 2015)**.** Moreover, according to the prospect theory (*Kahneman & Tversky 1979)* people are more inclined to choose the no-punishment option than the matching reward option, phenomenanamed as loss aversion bias. In animal research of punishment, Marchant et al found individual differences in the susceptibility to constant shock intensity in alcohol preferring Pack rats. The data showed a bimodal distribution in the response to punishment (*Marchant et al 2018*). Nevertheless, other studies did not find a significant difference in classical conditioning through reward and punishment (**Delgado et al**).

In the clinical psychiatric perspective, disfunctions in reward or punishment perception, result in different kinds of clinical disorders. For example, a dysfunctional reward system has been linked to depression and bipolar disorder (*Eldar 2016*). In contrast, a dysfunctional punishment system, was linked to anxiety (more precisely, cognitive anxiety; see *Wise & Dolan 2020*). For example, Aylward et al found that anxious participants learned faster from negative outcomes, i.e.- had higher learning rate (*Duits et al 2015; Aylward et al 2019*; although see *Wise & Dolan 2020*). On the other hand, a lower sensitivity to punishment is implicated in addiction and psychopathy (*Jean-Richard-Dit-Bressel et al 2018*).

In this study we test responses exclusively in an aversive context (punishment minimization condition). The only possible outcomes are punishment and the absence of punishment. Hence, there is only punishment prediction error. Importantly, the subject's choice is to avoid the stimulus that has the worse potential outcome by withdrawing their finger from the undesirable image and as a result, the other (supposedly better) one, is chosen. This design is implementing a scenario in which the agent can only lose, there is nothing to gain by avoiding the bad stimulus except avoiding punishments. It simulates a situation of running away from a threat and choosing the least dangerous choice, therefor all the attention is going to the punisher. As a result of this excluding environment, we wanted to capture the clear effect of punishment and its unique features and see more clearly how punishment affect learning and decision-making. Furthermore, we can test the effects of different kinds of punishments and compare between them (*Delgado et al 2011*).

**Importantly, this experiment has a novel methodology** to collect behavioral, physiological, and psychological data to probe learning mechanisms over long period of time. It is a longitudinal experiment conducted for 12 days. The experiment is running through a novel mobile platform design and conducted outside of laboratory setting. This is a new realistic approach for human studies that allow us to get highly dense and realistic data collection. There are challenges for this methodology, including the lack of experimental control, tracking subject's performance and technical variables that may interfere with the progress of the experiment. We used various measures to overcome these challenges which we elaborate on in the methods section.

The aim of this study is to validate that our design is working, and subjects do learn the values of the stimuli throughout the 12 days of the experiment, through punishments. Our target is to track the daily performance and heartrate responses to evaluate the learning process. These results will also give us a first, non-conclusive glimpse about the differences between a positive-primary punisher and a negative-secondary punisher, white noise and monetary loss, respectively. Finally, future study will compare aversive instrumental learning results with the data of a parallel experiment on reward instrumental learning conducted in the lab. In addition to the behavioral data, we monitored the heartrate of subjects while doing the tasks. With this, we hope to see physiological differences between punished and non-punished, to see the implicit emotional response to punishment. This experiment will hopefully give us good basic design for future studies that will get interesting conclusions about the process of learning from punishments and its implications.

## 

# Materials and Methods

## Participants:

19 healthy volunteers (mean-age: 25.47±3.53; range 20 – 33 years old; 12 female subjects) were recruited through social media advertisements and participated in a 12-day experiment conducted through a trial-and-error game installed on their phones. Subjects who did not have phones that are compatible with the experiment application were delivered a phone owned by the lab. All subjects underwent a screening process to exclude any motoric, auditory or vision disability. The screening process also excluded any past psychiatric disorder or the use of psychiatric medications and drugs. The experiment was approved by the Helsinki committee and subjects signed a consent form accordingly. All subjects were paid by the hour (40 shekels) plus the amount left in the experiment bank or a finishing bonus (for elaboration on the paying method see the appendix section).

**Exclusion:** Seven participants were excluded from the study. Three participants excluded due to technical problems, another three because of insufficient compliance to the schedule of tasks, and one that had repeated bad performance from the beginning of the experiment.

After exclusion, 12 volunteers (mean-age: 25.58±2.96; range 20 – 31 years old; 8 female subjects) completed the experiment.

## Materials/Stimuli

**The mobile platform:** for the longitudinal learning game we used an Android app that was developed for experiment purposes by the lab P.I. Dr Eran Eldar. It was programed using the Android Studio programing environment (Google, Mountain View, CA). The game in the app was made for reward and we adjusted it for a punishing environment. All data collected by the app was stored locally on the phone as SQLite databases and regularly uploaded to a cloud storage space, designated for it.

For generating the white noise and cutting sound duration (0.5 seconds) we used the sound editor software *Audacity* (<http://audacityteam.org/>) version 2.4.2.

For the "Wheel of fortune" task, we used and edited the example code from <https://github.com/zarocknz/javascript-winwheel> (Copyright (c) 2016 Douglas McKechie).

The Learning game contained 64 images (round fractals with styled backgrounds). Each stimulus had a fixed probability of delivering punishment (0, .33, .66, 1).

The games in the preliminary lab meeting consisted of 13 games and overall 388 trials. The games outside the lab consisted of 44 games (within 11 days) of overall 3088 trials. Altogether, each subject had 3476 trials. 2211 of them were training trials (with feedback) and 1265 were testing trials (without feedback).

On each trial, two images were presented to the subject, and the subject had to remove her finger from the image she thought will most probably deliver punishment.

Punishments were either the loss of a coin (worth approximately 0.2 shekels) for the monetary-loss group or, the delivery of a loud white noise (92-95 dB, duration: 0.5 seconds) for the noise group.

## Procedure

Subjects were assigned randomly into two groups of two types of punishments. Each participant went through a screening interview to suffice the experiment criteria and if successful the subject was invited to the preliminary lab meeting.

**Preliminary lab meeting:** First, subjects signed a consent form. Then the experimenter explained them the schedule of the task throughout the days of the experiment.

For the **monetary-loss group**, we conducted a preliminary task beforehand to create an experimental bank from which the subject can lose money. This task is somewhat similar to other monetary loss studies (such as *Delgado et al 2011*; Steel et al 2016; Steel et al 2020) although our task was unique. The task included spinning a Wheel of fortune (in a computer program) with different amounts of gains (400-1000 shekels). Unknown to the participants, the win was fixed on 600 shekels (eq to ~182 dollars) in order to equalize the amount of money for all subjects. This was their bank of money from which they can lose and therefore should do their best to avoid punishments. The aim of this task was to magnify the effect of money loss on subjects by creating a sense of endowment to make the subject value more the initial sum of money. We also wanted to create a sense of agency, to increase the engagement and interest of subjects (Taub et al. 2020). The reason for doing that is our fear that since the amount being reduced on each loss of a coin is meager (0.2 shekels, ~ 0.06 dollars) we might lose the loss aversion effect and the motivation to avoid punishments might be sparse too. Although we have many trials along the experiment, and therefor the potential amount of loss is great, still, participant might not look on the "big picture" and therefor disregard the loss on a single trial. The sense of endowment and agency over the money increases the aversiveness of the monetary loss, therefore magnifying the effect on the subject (*Delgado et al 2011*; *Tricomi et al., 2004; Zink et al., 2004;* *De Martino et al., 2009*).

**The loud white noise group** was supplied with earphones from the lab (Miracase MBTO106). To make sure that subjects are wearing the earphones throughout the games and listening to the delivery of punishments, we added "attention colors task" that appeared between trials. Randomly, every few trials appeared a screen with six rectangles of different colors. Then, one of the color names randomly asserted, sometimes to the left ear and sometimes to the right ear, and the subject needed to press the matching colored-rectangle. Also, in games where no-feedback trials administered, when a set of punishments was delivered, the task appeared randomly between noises.

After that, for both groups, the experimenter downloaded and installed the application of the experiment on the subject's phone (or, if it was not compatible, the phone was delivered by the lab – "Redmi Note 9 Pro").

**Volume Calibration:**

For the delivery of aversive audio stimuli in a web-based experiments, Seow & Hauser (2021) showed that they are reliable for inducing affective states similar to in-lab studies, with the right technical measures.

For the White Noise group, a sound calibration was made beforehand to set the system's volume to the range predetermined by the experimenter (92-95 dB). To check the sound volume we used a sound meter monitor "UT353 Mini Sound Level Meter". As in Sperl et al study, the range was between 92-95 dB for every subject in the noise group (mean dB=92.68±0.69; range 92 – 93.6 dB). The variability in the dB volume is due to the different phone systems and from the amount of intensity of the noise perceived by each subject. If the noise was unbearably intense for the subject, we lower the volume but maintained a minimum of 92 dBs. In addition to the white noise volume calibration, we also needed to calibrate the volume of the "attention colors task". We wanted to keep volume to the minimum necessary for the subject to hear the names of the colors asserted but not higher so that they could not hear it without earphones. For each subject, while wearing the earphones, we played the names of colors in the lowest volume and increased it until the subject said she hear them clearly, but not more than that.

After that, we instructed the subject how to put the wearable heartrate sensor on her body, a "Polar H10 " device monitor that measures heartbeat rates.

A person and person in swimsuits on a beach

Description automatically generated with low confidence

Subjects had to wear the sensor whenever they played the experiment game, and it included a rest state measure of five minutes before starting the games.

Next, we explained the tasks to the subjects in details, including the structure of the game.

**APP SCHEDULE:**

The schedule of the app starts in the morning and when the subject wakes up, she needs to press a "woke up" button in the app and report the quality of her sleep and the content of her dreams if she can recall. Also the subject needed to play two games of the task in the morning and two games in the evening and to fill three mood questionnaires throughout the day. Before going to sleep, a video recording task was delivered in which the subject needed to talk about her day within 20 seconds and following that, pressing a "went to sleep" button when she is ready to go to sleep. This routine was kept for 10 Consecutive days, following by a rest day and finally another experiment day meant as a summery test for all stimuli.

**THE GAME:**

The game itself is a trial-and-error learning game in which subjects need to choose between 2 stimuli in every trial. Each stimulus has its own probability to deliver punishment. Unknown to the subjects, the probabilities were set to be in a hierarchical structure with the probabilities of [0, .33, .66, 1] where 0 represents no chance of getting punished and 1 means that punishment will be delivered 100% of the time.

Notably, subjects started each trial by pressing on both the right-side stimulus and left-side stimulus and the way of choosing is by lifting the finger pressing the unwanted stimulus. as a result, the counter stimulus is chosen.

After they choose, subjects can see the outcome of their choice and learn the value of that stimulus. A punishment outcome was seen as a red arrow pointing down inside a circle with black background (and in the noise group was paired with a loud white noise) and no punishment outcome was seen as a blank black circle. It was emphasized to subjects, that the game is probabilistic and therefore a bad stimulus can sometimes not deliver punishment (although it is still the worst choice) and a not bad stimulus can sometimes deliver punishment (although it is still the best choice).

A picture containing shape

Description automatically generated

After enough times that a set of stimuli was repeated, a curtain covered the outcomes to conceal them from the subject. This means that the training phase is over for these set of stimuli, and the **testing** phase starts. This way we can see how much is learned throughout the trials of those stimuli and test for their recall. Nevertheless, the outcomes of choices were stored, and punishments are presented as a message to the subject after every 10 trials saying that she lost X number of coins in the last 10 trials, in the money loss group, and in the white noise group saying that she was punished X times in last 10 trials. In addition to the message, in the white noise group, a loud white noise was delivered the number of times that the hidden outcomes conveyed punishments. This way we kept the incentives for high performance without revealing the outcomes and therefore ending the learning phase.

## Analysis

# We analyzed the behavioral data using R studio, and for the physiological data (HR) we used MATLAB for processing.

First, we added to the analysis data from a parallel experiment on reward conducted in the lab. 22 subjects performed the same learning game in parallel schedule, but with rewards instead of punishment. This enabled us to compare both punishment groups with reward.  
We preprocessed data by excluding the lab session trials from analysis, it serves for us as training phase. To calculate the RT we subtracted the time the stimulus presented from the time the choice was made by the subject.

**Binomial test.** Explanation

# Results

**Table 1: Accuracy & Rt**

| **Group** | **Accuracy** | **RT (milliseconds)** |
| --- | --- | --- |
| *White noise* | 82.6% ± 9% | 1628.6 ± 63 |
| *Money loss* | 81.7% ± 4% | 1867 ± 202 |
| *Reward* | 83.1% ± 6% | 1707.7 ± 130 |

.

A picture containing text, sky, screenshot

Description automatically generated

# First, we examined the performance of subjects while learning the different images. Overall accuracy of 81% for both punishment groups was observed with an Rt of 1744ms. The white noise group was slightly more accurate in the feedback trials with 82% accuracy while the monetary loss group had 81.5% accuracy.

# Discussion

**This study is using a novel design** in which 12 subjects played a trial-and-error learning game on a mobile platform for 12 days outside of the laboratory at their natural environment.

**The results show** that subjects of both groups learned stimuli values throughout the 12 days of the experiment…

habituation…

motivation…

**Moreover,** from the heartrate data that we collected when subjects played the games, we can see….

**Habituation.**

**Motivation.**

**This study has a small sample size** and therefore will not enable us to make conclusions about the differences between groups. However…

**The between-subjects design** is different from some other human studies probing punishments. This design gives a cleaner effect to the US punisher, as we avoid the confounds that a within design might obtain (Charness et al 2012). Importantly, we used a different primary punisher (Loud White Noise) than *Delgado et al 2011* (Mild shock) that showed more efficacy in the Sperl et al study (2016).

Our target was to make the task **similar to the environment of punishment** and therefore the method of choosing was by avoidance. To do that we made participants withdraw from the stimulus they did not wanted and only by that the paired stimulus was chosen. This method choosing by withdraw was also used in Huys et al (2011).

**Limitations and confounds** of the study.

**Similar future studies should** explore more conclusively the differences between a primary-positive and a secondary-negative punishers and compare them to the mechanisms of reward reinforcers.

**Conclusions:**

# Acknowledgments:

# References:

* Azrin, N. H., & Holz, W. C. (1966). Punishment. In W. K. Honig (Ed.), Operant behavior: Areas of research and application (pp. 380–447). New York: Appleton-Century-Crofts.
* Jean-Richard-dit-Bressel, Philip & Killcross, Simon & Mcnally, Gavan. (2018). Behavioral and neurobiological mechanisms of punishment: implications for psychiatric disorders. Neuropsychopharmacology. 43. 1. 10.1038/s41386-018-0047-3.
* Johnston, J. M. (1972). Punishment of human behavior. American Psychologist, 27(11), 1033–1054.
* Trenholme, I. A., & Baron, A. (1975). Immediate and delayed punishment of human behavior by loss of reinforcement. Learning and Motivation, 6, 62–79.
* Toshikazu Kuroda, Carlos R. X. Cançado & Christopher A. Podlesnik (2018) Relative effects of reinforcement and punishment on human choice, European Journal of Behavior Analysis, 19:1, 125-148, DOI: 10.1080/15021149.2018.1465754
* Rasmussen, E. B., & Newland, M. C. (2008). Asymmetry of reinforcement and punishment in human choice. *Journal of the experimental analysis of behavior*, *89*(2), 157–167. <https://doi.org/10.1901/jeab.2008.89-157>
* Delgado Mauricio R, Li Jian, Schiller Daniela and Phelps Elizabeth A. 2008 The role of the striatum in aversive learning and aversive prediction errors. Phil. Trans. R. Soc. B3633787–3800. https://doi.org/10.1098/rstb.2008.0161
* Delgado, M. R., Jou, R. L., &amp; Phelps, E. A. (2011). Neural systems Underlying AVERSIVE conditioning in humans with primary and secondary reinforcers. Frontiers in Neuroscience, 5. doi:10.3389/fnins.2011.00071
* M. R. Delgado, C. D. Labouliere, E. A. Phelps, Fear of losing money? Aversive conditioning with secondary reinforcers, Social Cognitive and Affective Neuroscience, Volume 1, Issue 3, December 2006, Pages 250–259, https://doi.org/10.1093/scan/nsl025
* Nathan J. Marchant, Erin J. Campbell, Konstantin Kaganovsky. Punishment of alcohol-reinforced responding in alcohol preferring P rats reveals a bimodal population: Implications for models of compulsive drug seeking. Progress in Neuro-Psychopharmacology and Biological Psychiatry. Volume 87, Part A. 2018. Pages 68-77. ISSN 0278-5846. <https://doi.org/10.1016/j.pnpbp.2017.07.020>.
* Sperl MFJ, Panitz C, Hermann C, Mueller EM. A pragmatic comparison of noise burst and electric shock unconditioned stimuli for fear conditioning research with many trials. Psychophysiology. 2016 Sep;53(9):1352-65. doi: 10.1111/psyp.12677. Epub 2016 Jun 11. PMID: 27286734.
* Wolfram Schultz. Behavioral dopamine signals. Trends in Neurosciences. Volume 30, Issue 5,2007. Pages 203-210,ISSN 0166-2236. <https://doi.org/10.1016/j.tins.2007.03.007>.
* Cools, R., Robinson, O. & Sahakian, B. Acute Tryptophan Depletion in Healthy Volunteers Enhances Punishment Prediction but Does not Affect Reward Prediction. Neuropsychopharmacology 33, 2291–2299 (2008). <https://doi.org/10.1038/sj.npp.1301598>
* Steel, A., Silson, E., Stagg, C. et al. The impact of reward and punishment on skill learning depends on task demands. Sci Rep 6, 36056 (2016). <https://doi.org/10.1038/srep36056>
* Wise, T., Dolan, R.J. Associations between aversive learning processes and transdiagnostic psychiatric symptoms in a general population sample. Nat Commun 11, 4179 (2020). <https://doi.org/10.1038/s41467-020-17977-w>
* Aylward, J., Valton, V., Ahn, WY. *et al.* Altered learning under uncertainty in unmedicated mood and anxiety disorders. *Nat Hum Behav* **3,**1116–1123 (2019). <https://doi.org/10.1038/s41562-019-0628-0>
* Duits, P., Cath, D. C., Lissek, S., Hox, J. J., Hamm, A. O., Engelhard, I. M., Van Den Hout, M. A., & Baas, J. M. P. (2015). Updated meta‐analysis of classical fear conditioning in the anxiety disorders. Depression and Anxiety, 32(4), 239– 253. <https://doi.org/10.1002/da.22353>
* Galea, J., Mallia, E., Rothwell, J. et al. The dissociable effects of punishment and reward on motor learning. Nat Neurosci 18, 597–602 (2015). <https://doi.org/10.1038/nn.3956>
* Michely J, Rigoli F, Rutledge RB, Hauser TU, Dolan RJ. Distinct Processing of Aversive Experience in Amygdala Subregions. Biol Psychiatry Cogn Neurosci Neuroimaging. 2020 Mar;5(3):291-300. doi: 10.1016/j.bpsc.2019.07.008. Epub 2019 Aug 2. PMID: 31542358; PMCID: PMC7059109.
* Palminteri, S., Khamassi, M., Joffily, M. *et al.* Contextual modulation of value signals in reward and punishment learning. *Nat Commun* **6,**8096 (2015). <https://doi.org/10.1038/ncomms9096>
* Steel, A., Baker, C.I. & Stagg, C.J. Intention to learn modulates the impact of reward and punishment on sequence learning. *Sci Rep* **10,**8906 (2020). <https://doi.org/10.1038/s41598-020-65853-w>
* O'Doherty, J., Kringelbach, M., Rolls, E. *et al.* Abstract reward and punishment representations in the human orbitofrontal cortex. *Nat Neurosci* **4,**95–102 (2001). <https://doi.org/10.1038/82959>
* Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica*, 47, 263–292.
* S.G. Costafreda, M.J. Brammer, A.S. David, C.H. Fu, Predictors of amygdala activation during the processing of emotional stimuli: A meta-analysis of 385 PET and fMRI studies Brain Res Rev, 58 (2008), pp. 57-70
* Michelle Taub, Robert Sawyer, Andy Smith, Jonathan Rowe, Roger Azevedo, James Lester,The agency effect: The impact of student agency on learning, emotions, and problem-solving behaviors in a game-based learning environment,Computers & Education,Volume,147,2020,103781,ISSN,03601315,https://doi.org/10.1016/j.compedu.2019.103781
* Prefrontal Oscillations during Recall of Conditioned and Extinguished Fear in Humans, Erik M. Mueller, Christian Panitz, Christiane Hermann, Diego A. Pizzagalli, Journal of Neuroscience 21 May 2014, 34 (21) 7059-7066; DOI: 10.1523/JNEUROSCI.3427-13.2014
* Dynamic neural activity recorded from human amygdala during fear conditioning using magnetoencephalography, Sandra N. Moses, Jon M. Houck, Tim Martin, Faith M. Hanlon, Jennifer D. Ryan, Robert J. Thoma, Michael P. Weisend, Eric M. Jackson, Eero Pekkonen, Claudia D. Tesche, Brain Research Bulletin,Volume 71, Issue 5,2007,Pages 452-460,ISSN 0361-9230, <https://doi.org/10.1016/j.brainresbull.2006.08.016>.
* Dolan, R., Heinze, H., Hurlemann, R., & Hinrichs, H. (2006). Magnetoencephalography (MEG) determined temporal modulation of visual and auditory sensory processing in the context of classical conditioning to faces. NeuroImage, 32(2), 778–789. doi: 10.1016/j.neuroimage.2006.04.206
* Eldar, E., Hauser, T.U., Dayan, P. & Dolan, R.J. (2016a) Striatal structure and function predict individual biases in learning to avoid pain. Proc Natl Acad Sci U S A, 113, 4812–4817.
* Gary Charness, Uri Gneezy, Michael A. Kuhn,Experimental methods: Between-subject and within-subject design,Journal of Economic Behavior & Organization,Volume 81, Issue 1,2012,Pages 1-8,ISSN 0167-2681,https://doi.org/10.1016/j.jebo.2011.08.009.
* Franzoi, S. L. (2015). Psychology: a discovery experience. South-Western Cengage Learning.
* Huys QJM, Cools R, Go¨ lzer M, Friedel E, Heinz A, et al. (2011) Disentangling the Roles of Approach, Activation and Valence in Instrumental and Pavlovian Responding. PLoS Comput Biol 7(4): e1002028. doi:10.1371/journal.pcbi.1002028
* Seow, T. X. F., & Hauser, T. (2021). Reliability of web-based affective auditory stimulus presentation. In: bioRxiv.
* Ulf-Dietrich Reips, Chapter 4 - The Web Experiment Method: Advantages, Disadvantages, and Solutions,Editor(s): Michael H. Birnbaum,Psychological Experiments on the Internet,Academic Press,2000,Pages 89-117,ISBN 9780120999804,https://doi.org/10.1016/B978-012099980-4/50005-8.
* Johnson SL, Turner RJ, Iwata N. BIS/BAS levels and psychiatric disorder: an epidemiological study. Journal of Psychopathology and Behavioral Assessment. 2003;25(1):25–36. [Google Scholar]
* Carver CS, White TL. Behavioral inhibition, behavioral activation, and affective responses to impending reward and punishment: the BIS/BAS scales. Journal of Personality and Social Psychology. 1994;67:319–33.
* Wimmer G.E., Poldrack R.A. Reinforcement learning over time: spaced versus massed training establishes stronger value associations. bioRxiv. 2017; https://doi.org/10.1101/158964
* Iigaya K. Adaptive learning and decision-making under uncertainty by metaplastic synapses guided by a surprise detection system. eLife. 2016; 5: e18073
* Mcsweeney, F.K., Roll, J.M. Do animals satiate or habituate to repeatedly presented reinforcers?. *Psychonomic Bulletin & Review* **5,**428–442 (1998). <https://doi.org/10.3758/BF03208818>
* Howell, D. C. (2006). *Statistical Methods for Psychology* (6th ed.). Wadsworth Publishing.
* Touré-Tillery, M., & Fishbach, A. (2014). How to Measure Motivation: A Guide for the Experimental Social Psychologist. *Social and Personality Psychology Compass*, *8*(7), 328–341. https://doi.org/10.1111/spc3.12110

# Footnotes

\**Azrin and Holtz* uses the phrase ***immediate***. we used the word ***contingent*** because, though weaker, a punisher may also be delayed and still reduce the likelihood of the preceding behavior. Additionally if we explain the delay to the subject there is an increase in probability reduction (*Trenholme & Baron, 1975*)