# 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*). Nevertheless, it seems like aversive reinforcement is unique in its traits. One of its famous phenomena is that people are more inclined to choose the no punishment option than the matching reward option, named loss-aversion bias (*Kahneman & Tversky 1979)*.

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, which is harder to maintain remotely. Another concern we had to address was that subjects may 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, with respect to neuronal mechanisms, is whether reward and punishment have different neural systems in the brain, and a few hypotheses were suggested to address it.

Despite all the conceptual parallels (that was mentioned above) 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 (gaining more than expected) and negative reward prediction error (gaining less than expected), respectively (*Schultz 2007*), whereas in aversive learning mechanisms (such as punishment prediction error, which is the difference between expected and actual punishment) 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 (2020) show that the basolateral amygdala encodes a punishment prediction error. 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 and reward affect learning and behavior differently. Steel et al (2016) 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 (better performance) whereas the effect on the FTT task was negative (worse performance). Galea et al. (2015) found that punishment leads to faster motor adaptation whereas reward caused greater retention (Galea et al 2015)**.** Nevertheless, other studies did not find a significant difference in classical conditioning through reward and punishment (**Delgado et al**).

It is important to mention individual differences in of punishment. Animal research provide evidence of 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*).

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

In the study we test responses exclusively in an aversive context (punishment minimization condition), meaning that the only possible outcomes are punishment and the absence of punishment, hence, enabling only learning from punishment prediction errors. The subject's task was to avoid the stimulus that had the worse potential outcome by withdrawing their finger from the undesirable image and by this action choosing the other (supposedly better) one. The task simulates a situation of running away from a threat and choosing the least dangerous option, therefore all the attention is going to the punisher. Such excluding environment was established to capture the clear effects of different kinds of on learning and decision-making, commonalities, and differences between them (*Delgado et al 2011*).

The novelty of our methodology is that the experiment is running through a novel mobile platform design and is conducted outside of laboratory setting. This is a new approach for human studies that allow us to get highly dense daily data collection. This methodology has some limitations though, validating task requirements are followed, tracking subject's performance and technical variables that may interfere with the progress of the experiment. We used various measures to account for them 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 was 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.