UNIVERSIDAD DE JAÉN

Escuela de Doctorado
Programa de Doctorado en Psicología

DOCTORAL DISSERTATION / TESIS DOCTORAL

●

[EXPERIENCING ASSOCIATIVE INTERFERENCE FACILITATES EXCITATORY PREDICTIVE LEARNING ABOUT A CONDITIONED INHIBITOR IN HUMANS]

[EXPERIMENTAR INTERFERENCIA ASOCIATIVA FACILITA EL APRENDIZAJE PREDICTIVO EXCITATORIO SOBRE UN INHIBIDOR CONDICIONADO EN HUMANOS]

PRESENTADA POR:
Gabriel González Tirado

DIRIGIDA POR:
Juan Manuel Rosas Santos
José Enrique Callejas Aguilera

JAÉN, 11 de DICIEMBRE de 2020

Autorización

Juan Manuel Rosas Santos, Catedrático de Psicología Básica de la Universidad de Jaén, y José Enrique Callejas Aguilera, Profesor Titular de Psicología Básica de la Universidad de Jaén,

Garantizan que:

La Tesis Doctoral titulada: "*Experiencing Associative Interference Facilitates Excitatory Predictive Learning About a Conditioned Inhibitor in Humans / Experimentar Interferencia Asociativa Facilita El Aprendizaje Predictivo Excitatorio Sobre un Inhibidor Condicionado en Humanos/"*, realizada por el doctorando Gabriel González Tirado, ha sido elaborada bajo nuestra dirección y reúne las condiciones de calidad, originalidad y rigor científico necesarias para que se proceda a su defensa pública de acuerdo con la legislación vigente.

Fdo.: Juan Manuel Rosas Santos

Fdo.: José Enrique Callejas Aguilera

Jaén, a 11 de diciembre de 2020

[“The only real mistake is the one from which we learn nothing”]

[*Henry Ford***]**

Agradecimientos

Son los agradecimientos, en mi caso, el texto que he escrito en último lugar. Seguramente por miedo a cometer el error de olvidarme a alguien o, quizás, porque es la parte más personal de este trabajo. Ruego perdón de antemano si alguien cree que debió estar aquí y no le nombro de manera especial. Seguro que está en lo cierto, pero escribir estas últimas líneas es más difícil de lo que imaginé en un principio.

Querría empezar por mis directores: José Enrique y Juanma. Cuando vine a la ciudad, llegué prácticamente con lo puesto. No conocía a nadie y ellos me abrieron las puertas de par en par. Además de la forma en la que me acogieron, tengo que agradecer esos pequeños detalles del día a día que te ayudan a sentirte como en casa. Gracias a ellos he aprendido infinidad de cosas (y sigo aprendiendo). Gracias José Enrique por tu cercanía. Siempre estás disponible para quién te necesita y das tu tiempo en ayudar a los demás. No sabes cuánto de importante es eso para quién lo necesita. Gracias Juanma. Tu valía personal es inconmensurable, así como tu capacidad de transmitir y enseñar. Sin la combinación perfecta de estos dos grandes profesionales, esta tesis doctoral no habría sido posible.

Por supuesto a mis compañeros Ale, José Andrés y Pedro. Qué grandes ratos que hemos pasado en el laboratorio, en la cafetería, en el despacho…vamos donde nos han dejado y donde hemos podido. Gracias Ale por tus clases magistrales de cómo usar la piscina y las cajas de Skinner. Gracias José Andrés por ser una pequeña gran Wikipedia de conocimiento científico y buen rollo. Gracias Pedro por la cantidad de cafés (y alguna cerveza) que hemos compartido juntos. Os merecéis todo lo bueno que os pase.

No podían faltar Rodolfo y los compañeros de México: Tere, Jaime, Mariel, Lalo…por enseñarme su pequeño lugar. Por compartir sus recursos con un desconocido y hacerme sentir como uno más. Siempre os estaré eternamente agradecido.

A la familia del 017: Rocío Donaire, Rocío Linares, María, Pablo, Teresa, Carmen, David y, por supuesto, los fichajes de última hora Sonia y Loida. Los ratos de despacho y escritura con vosotros eran más llevaderos. Compartir con vosotros esta experiencia ha sido un placer.

Por supuesto a la familia que no se elige. A mis padres, mi hermana y mi sobrino Darío. Por apoyarme en la distancia. Por ayudarme en la cercanía. Ellos me han enseñado el valor del trabajo y la constancia. A amar lo que uno ama y cuidarlo por encima de todo. Gracias por habérmelo dado todo sin pedir nada a cambio, solo por el amor incondicional que una familia muestra a sus seres queridos. No os pude elegir, pero si hubiera podido, lo habría hecho una y mil veces.

También a la familia que se elige: Yeyo, Mercedes, Antonio y Abraham y a las incorporaciones más recientes, Manu, Álvaro, Pedro, Peter, Vero y Lorena. Por todas las veces que me han preguntado por mi trabajo y me han apoyado. Porque alguna vez ellos vieron cosas en mí que ni uno mismo es capaz de ver. Gracias por nuestros faltadas en el pádel, las hamburguesas del Goiko, los domingos en "La Victoria" y las Alcázar en "La Manchega". Gracias por vuestro tiempo y ayuda. Os lo debo.

Por último, *resulta inevitable*, a Victoria. Mi compañera y mi amiga. Mi confidente y mi pareja. Mi quebradero de cabeza y también mi motivo de felicidad. Gracias por ser mi faro cuando estaba perdido. Gracias por creer en mi cuando ni yo mismo creía. Gracias por aguantarme cuando ni yo mismo lo hago. Gracias por acompañarme en este camino (y los que nos quedan). Tu también eres participe, a tu medida, de este trabajo. No existen palabras suficientes para agradecértelo. Aquí solo hay una pequeña muestra.

No quería dejar pasar a aquellas personas que, sin saberlo, han hecho que el desarrollo de una tesis doctoral sea más llevadero y que sus palabras, gestos y momentos hacen felices a aquellos que los disfrutan. Nati y Paco, Encarna (y sus croquetas), los chirigoteros…

Cualquier persona que ha realizado un trabajo semejante sabe que, de manera individual, cada uno planta su raíz, que las personas que te rodean son el agua y el fertilizante que la hace crecer y florecer y que, finalmente, estos pequeños párrafos son el fruto del trabajo.

Acknowledgements

Research was made possible by Grants PSI2014-52263-C2-1-P from the Spanish Ministry of Science and Competitiveness and PGC2018-097769-B-C22 from the Spanish Ministry of Science, Innovation, and Universities.

Index

Abstract

The goal of this doctoral thesis was to explore the effect of increasing prediction error on subsequent new learning. With this goal, we have developed two studies in human predictive learning evaluating to the effect of sudden increases in prediction error produced by different associative interference experiences on subsequent learning, assuming that increases in prediction error lead to general increases in attention that would facilitate subsequent learning.

This dissertation begins with a review of classical associative learning models and the role they have given to the error term in learning (Mackintosh, 1975; Pearce & Hall, 1980; Rescorla & Wagner, 1982; Sutton & Barto, 1981; Wagner, 1981), as well as some biological bases that support the explanation of the concept of prediction error at brain level. Next, the main procedures for generating associative interference and the implications they have on both contextual dependence of the information retrieval and other types of complex learning are detailed. The two empirical studies included in this doctoral thesis are devoted to explore the effects that the manipulation of the prediction error through different associative interference procedures have on inhibition reversal.

In general, both the theoretical review and empirical results suggest that the manipulation of the prediction error produces an increase in attention which, in turn, translates into an improvement in later learning, regardless of whether the increases in the prediction error involves extinction of a cue that already predicted the outcome, or pairing with the outcome a cue that was previously followed by the absence of it.

Preface

Organisms must adapt to the environment where they live to survive. Evolution leads to a selection of different abilities to predict and control their environment in different species, leading the species that way to obtain advantages that guarantee its survival. Alternatively, individuals learn about the events that occur in their habitat with multiple goals: Obtaining food, avoiding predators, choosing successful sex partners, etc. We could say that the mechanisms through which species adapt to their environment to survive are summarized in two: Natural selection and learning.

Natural selection occurs when a certain feature that appears randomly in one or more individuals of a species, suppose an advantage to survival. This advantage translates, later, in a greater reproduction probability. Therefore, those individuals that present this differential feature perpetuate it, converting it, after several generations, into a species characteristic. An example of the importance of natural selection is described by Endler (1978, 1980) in Poecilia reticulata Peters fish (Guppies). In this species, the color pattern has a double role: It is a mechanism to avoid predators, and it is the main attraction of the male to seduce the opposite sex. Individuals with a very striking color pattern will be more easily hunted by their predators; conversely, inconspicuous color patterns will make females choose other males (Endler, 1978, 1980). Survival, then, is in a perfect balance between the two extremes.

Alternatively, abrupt changes that occur in the environment may favor that individual features that did not suppose any adaptive advantage, acquire a high value in survival, so that individuals with that feature have a greater probability of reproducing and, therefore, that feature perpetuates as a species trait. A classic example of adaptation in the species is reported by Grant (1991) related to the size of the finches. The average size of the finches in Galapagos Islands increased due to a drought that took place there during the 1976-1978 period. Natural selection allowed larger finches to survive. A few generations later, the size of the finches was larger compared to their species' partners in other areas of the world where natural selection did not facilitate survival and reproduction of larger finches. As Grant (1991) reported, the size of the species may be decisive to resist long periods in the absence of water. Therefore, there may be a direct relationship among those features that facilitate survival in a specific environment and the natural features of the species that inhabit that environment. This relationship is bidirectional: Organisms show adaptive advantages to the adversities of a given environment that are, in turn, their own features because they served to overcome the dangers that characterized that environment in previous generations.

Natural selection as a mechanism of adaptation requires of multiple generations so that does not allow rapid adaptation to environmental features that are constantly changing. That task is awarded to learning mechanisms that are themselves favored by natural selection, and that allow the organism to quickly adapt to changing features of the environment. Rapid adaptation to sudden changes in the habitat requires learning mechanisms able of detecting both regularities and irregularities. For example, in some species like humans, people can learn from other members of the species to perform certain behaviors to obtain the same benefits, an ability that improves surviving, as allows learning to occur without the need of directly experiencing the specific situation. Knowledge and skills serve the individual and may even survive across generations (Galef & Laland, 2005; Johnston, 1982).

Although environments change over time, it is also true that they tend to remain relatively stable, presenting multiple regularities that may be used by the organisms to facilitate their survival. However, habitat features may also change across an individual lifespan because of many reasons. For example, habitats may change by human activities (resources exploitation, hunting, the spread of exotic species) or by climate change (Sih, Ferrari, & Harris, 2011). Thus, in the same way that regularities are established, they can cease to be useful if the conditions of the environment change. An organism with the ability to adapt to all those changes will be in an evolutionary advantage with respect to an organism without that ability.

Learning capacity and the species features combine to determine how learning proceed. For instance, LoLordo (1979) showed that, in pigeons, fear conditioning is easier to learn with auditory cues while visual cues are more useful when they are used to signal the presence of food (see Foree & Lolordo, 1973). This is known as stimulus specificity and reflects the ability of each species to respond to a specific stimulus dimension or characteristic better than others. In evolutionary terms, this selective capacity involves an adaptive advantage for surviving that explains the presence of the current feature in the species. Pigeons able to localize food while flying are more likely to survive, as they can localize food from the distance. Meanwhile, pigeons that use noise as a signal of danger are better prepared to detect predators. We could define learning as the ability to organize the environment through detection and prediction of the regularities in different situations. Those species able to detect and use environmental regularities more accurately and more quickly are more likely to survive and, therefore, they are more likely to reproduce (Iturbe, 2010). In a long selection process, this involves the perpetuation of the characteristic in a certain species. Related to this idea, when a learning ability has an explicit benefit and it is easy to obtain, natural selection occurs faster than when this benefit is not so clear or the ability is exceedingly difficult to achieve (Miller & Todd, 1991). Learning can lead to adaptive advantages in multiple aspects such as sexual behavior, reproductive, fear, aggressiveness and feeding among others (Domjan, 2005). In other words, to learn is to survive.

As a basic mechanism of survival and adaptation to the environment, the study of learning has received great attention since the beginning of time. Just take as a reference the Aristotle association laws where it is stated that two ideas will be associated with each other when they are presented together. Defining learning as the ability of detect regularities and establish predictions leads to understand the learning mechanism as a mechanism that corrects prediction error, understanding prediction error as the difference between what the organism expects and what it finally occurs in the environment. Learning is a nonstop mechanism. Learning will occur whenever prediction error is present. It will cease when the prediction error disappears. (e.g., Mackintosh, 1975; Pearce & Hall, 1980; Rescorla & Wagner, 1972). But it will be engaged again as soon as environment changes with the goal of eliminating the error (Miller & Escobar, 2002).

The studies conducted in this doctoral thesis depart from the idea that the mechanism of learning is a correction error mechanism by assuming that abrupt changes in prediction error may produce changes in the individual's attention to the environment as well, and that those changes in attention will facilitate subsequent learning. The general goal of the studies conducted in this thesis was to test whether experiencing prediction error facilitates subsequent learning in human predictive learning. The thesis is structured in five chapters. Chapter 1 will detail the state of the art, including information about the most relevant studies that relate attention and prediction error, as well as a brief description of the associative models of learning that relate both concepts. Specific goals and hypotheses will be described in Chapter 2. Chapter 3 includes two empirical studies that will try to shed some light on the goals and hypotheses presented in Chapter 2. In these two empirical studies the impact of experiencing associative interference on subsequent learning about a pre-exposed cue in human predictive learning is explored. These two empirical studies are presented in a publication format (the first one, has been already published, indeed). Chapter 4 includes a general discussion of both empirical studies and the main conclusions of the research conducted in this thesis. Finally, Chapter 5 consists of an extended summary in Spanish.

References

Domjan, M. (2005). Pavlovian conditioning: A functional perspective. *Annual Review of Psychology, 56*, 179-206. doi: 10.1146/annurev.psych.55.090902.141409

Endler, J. A. (1978). A predator's view of animal color patterns. In Max K. Hecht, William C. Steere & Bruce Wallace (Eds), *Evolutionary biology* (pp. 319-364). Springer, Boston, MA.

Endler, J. A. (1980). Natural selection on color patterns in Poecilia reticulata. *Evolution*, *34*(1), 76-91. doi: 10.1111/j.1558-5646.1980.tb04790.x

Foree, D. D., & LoLordo, V. M. (1973). Attention in the pigeon: differential effects of food-getting versus shock-avoidance procedures. *Journal of Comparative and Physiological Psychology, 85*(3), 551-558. doi: 10.1037/h0035300

Galef, B. G., & Laland, K. N. (2005). Social learning in animals: empirical studies and theoretical models. *Bioscience, 55*(6), 489-499. doi: 10.1641/0006-3568(2005)055[0489:SLIAES]2.0.CO;2

Grant, P. R. (1991). Natural selection and Darwin's finches. *Scientific American, 265*(4), 82-87. doi: 10.1038/scientificamerican1091-82

Iturbe, U. (2010). Adaptaciones y adaptación biológica, revisadas. *Revista de la sociedad española de biología evolutiva, SESBE.* Granada, España, Retrieved from: http://www.uaeh.edu.mx/investigacion/productos/4266/ADAPTACION.pdf

Johnston, T. D. (1982). Selective costs and benefits in the evolution of learning. In *Advances in the Study of Behavior* (Vol. 12, pp. 65-106). Academic Press. doi: 10.1016/S0065-3454(08)60046-7

LoLordo, V. M. (1979). Selective associations. In A. Dickinson & R. A. Boakes (Eds.), *Mechanisms of learning and motivation.* Hillsdale, N.J.: Erlbaum.

Mackintosh, N. J. (1975). A theory of attention: variations in the associability of stimuli with reinforcement. *Psychological review, 82*(4), 276-298. doi: 10.1037/h0076778

Miller, R. R., & Escobar, M. (2002). Associative interference between cues and between outcomes presented together and presented apart: An integration. *Behavioural Processes, 57*(2-3), 163-185. doi: 10.1016/S0376-6357(02)00012-8

Miller, G. F., & Todd, P. M. (1991). Exploring adaptive agency I: Theory and methods for simulating the evolution of learning. In D. S. Touretzky, J. L. Elman, T. J. Sejnowski, & G. E. Hinton (Eds.), *Proceedings of the 1990 connectionist models summer school* (pp. 65–80). San Mateo, CA: Morgan Kaufmann. doi: 10.1016/B978-1-4832-1448-1.50013-5

Pearce, J. M., & Hall, G. (1980). A model for Pavlovian learning: variations in the effectiveness of conditioned but not of unconditioned stimuli. *Psychological review, 87*(6), 532-552. doi: 10.1037/0033-295X.87.6.532

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 research and theory.* New York: Appleton-Century-Crofts.

Sih, A., Ferrari, M. C., & Harris, D. J. (2011). Evolution and behavioural responses to human-induced rapid environmental change. *Evolutionary applications, 4*(2), 367-387. doi: 10.1111/j.1752-4571.2010.00166.x

CHAPTER I: STATE OF THE ART

1. Prediction error as the learning engine

Learning allows the organism to predict the events that occur in the environment. Sometimes, those predictions are accurate. Other times, they are not. In those occasions in which the organism does not receive what it expects, so that there is a discrepancy between the organism's expectation and what happens in the environment it is said that there is a prediction error.

Traditional associative models assume that learning is driven by prediction error since the organism tries to reduce the discrepancy between expectations and events. Models differ in the explanatory mechanism that is assumed to be used by the organism to reduce prediction error, that is, in the learning mechanism, but they share the idea that it is prediction error what drives learning (e.g., Mackintosh, 1975; Pearce & Hall, 1980; Rescorla & Wagner, 1972; Sutton & Barto, 1981). Classical models of learning have expressed the prediction error with the following formula: $\lambda - \Sigma V$ (the delta rule). The first element of the equation, λ, is the maximum conditioning that the unconditioned stimulus (US) could award, that is, the asymptote of learning. Formally, λ value is assumed to be 0 when the US is absent, and 1 when the US is present in a determined learning trial. The second element of the equation, ΣV, represents the associative strength of all the conditioned stimuli (CS) that appear in a single trial associated to the given US presented in the learning situation. The discrepancy between λ and ΣV is what it is called prediction error.

Prediction error changes as learning develops. At the beginning of learning, the organism does not know the relationship between the CS and the US. As stated above, in each learning trial in which the outcome (the US) is presented, the value of λ is assumed be 1. The value of ΣV before the learning experience takes place would be 0, because the organism has not yet learned to expect the outcome in the presence of the cue. Prediction error in that situation is at its maximum level. The organism does not expect the outcome, and the outcome appears. As learning develops, ΣV is assumed to increase its value, gradually approaching the value of λ, and reducing the prediction error. When the organism learns to fully predict the presence of the outcome given the cue, the value of ΣV equates the value of λ, prediction error disappears, and learning stops.

The main difference among associative models lies in the explanatory mechanism proposed to explain how prediction error is reduced. Some theories, such as Rescorla and Wagner's model (1972), and Wagner SOP model (Wagner, 1981), assumed that the attention to different CSs is based on their physical features, and remain stable during learning, so that the focus of the attenuation of prediction error is in US relevance. Other models focus on the relevance of the attention the CS receives assuming that variations in prediction error affect CS attention. Among the later ones, are those proposed by Larrauri and Schmajuk, (2008), Mackintosh, (1975) and Pearce and Hall, (1980), among others. The Mackintosh model (1975) suggests that organisms pay more attention to CSs that reliably predict the occurrence of US; On the contrary, Pearce and Hall model (1980) suggests that the organism attends more to poor predictors of the CS, so that attention to the CS is reduced the better predictor of the US the CS becomes. From a different perspective, Larrauri and Schmajuk (2008) suggests there is a direct relationship between attention and novelty. More specifically, the more novel the event is, the greater the attention the organism pays to the different elements presented during the learning situation. Finally, special mention requires the hybrid model proposed by Le Pelley, Mitchell, Beesley, George, and Wills (2016) as it is able to reconcile the apparently opposite predictions from the Mackintosh (1975) and Pearce and Hall (1980) models (see also, Hall & Rodriguez, 2010, as another example of an attentional hybrid model). Next, we will briefly dwell on each of these models due to the theoretical implications they have for the development of the hypotheses that drive this doctoral thesis.

Rescorla and Wagner's model (1972) is one of the most influential associative learning models in the scientific literature, and its basic principles lay the foundations for many models that appeared years later (see Siegel & Allen, 1996; c.f., Miller, Barnet, & Grahame, 1995). This model assumes a direct implication of prediction error in learning. The model gives a main role to the unconditioned stimulus; even more, learning success depends on how surprising is the unconditioned stimulus that is presented with respect to what the organism is expecting. This model assumes that the associations established between CSs and the US are individually formed (Rescorla & Wagner, 1972). Prediction error is computed by taking in account all the stimuli that precede the US. At the beginning of training, surprise will be maximum since the difference between λ (predictive capacity of the US) and ΣV (organism expectation) is maximum as well. As training progresses, this difference gradually decreases until it disappears ($\lambda = \Sigma V$). Surprise is defined as the difference between what the organism receives, and what the organism expects to receive based in all the CS present in the situation. Thus, prediction error can be defined as the formal expression of surprise (as a psychological concept). At the beginning the learning, surprise is maximum, as there is no good predictor of the outcome (yet) in the learning situation, and prediction error is at its highest. As training progresses, one or more cues become good or poor predictors of the outcome depending on the specific experimental arrangement, prediction error diminishes, surprise diminishes as well, and so diminishes learning. Likewise, surprise increases if a cue that has already been established as a good predictor of the outcome, suddenly ceases to predict the outcome. The learning mechanism will be engaged until prediction error is again reduced to zero. As described, relationship between prediction error and learning is a direct one. In the next sections we will describe attentional models of learning in which the role of prediction error is indirect. Prediction error is assumed to modulate attention to the CS that, in turns, affects learning.

Mackintosh's attentional model of associative learning (1975) bases their explanations on the idea that the attention the CS receives, changes with the learning experience. This model proposes the CS perceived salience depends on whether the CS is a good or a poor prediction of the US, highlighting attention as a determining factor of learning. For his model, attention is part of the mediating and strengthening processes involved in establishing associations. It is assumed that the better predictor the CS is, the greater the attention the CS receive. Also, the opposite is true for poor predictors. Cues that do not predict the arrival of the US better than other cues presented in the situation tend to be ignored. That is, there is a direct relationship between the ability of the trained CS to predict the reinforcer and its associability: The greater the ability to predict the outcome, the greater its associability. The opposite occurs with the cues present that are comparatively identical or worse predictors of the outcome than the target cue: Their associability decreases. In the original Mackintosh's formulation [$\Delta V = \theta\ \alpha\ (\lambda - \Sigma V)$], θ is the learning rate parameter, and its value depends on the reinforcer used in a given trial (ranging from 0 to 1); α represents the associability of the cue and, again, $(\lambda - \Sigma V)$ is the prediction error. More precisely, if a given stimulus, X, is the best predictor in the learning situation, α_X will be high. Associative changes directly depend on the prediction error, as the value of the associability parameter depends on the discrepancy existing between λ and the relative associative strength of the target cue. More precisely, it is assumed that the greater the discrepancy, the less attention the cues receive since their predictive capacity is unknown. Consequently, the prediction error and attention are closely related in this model. This is known as the reverse hypothesis (Thomas, 1970). Likewise, subsequent learning also modifies this associability. Specifically, X consistent pairings with the outcome will increase the relative discrepancy between X and the rest of the stimuli present in the learning situation: V_T, $[(\lambda\text{-}V_X) < (\lambda\text{-}V_T)]$, so that α_X will be high and increasing. In those learning situations in which the rest of the stimuli presented in the learning situation are better predictors than X, $[(\lambda\text{-}V_X) > (\lambda\text{-}V_T)]$, X's associability ($\alpha_X$) will decrease (see Mackintosh, 1975). In summary, in Mackintosh's model there are two elements that determine the CS's associative strength, and both elements are interrelated: Prediction error ($\lambda - \Sigma V$) and α. Mackintosh (1975) points out that the greater the association between the conditioned stimulus and the reinforcer, the greater the attention it receives.

Pearce-Hall's (1980). Pearce and Hall (1980) take an approach that is contrary to the one taken by Mackintosh's model, by suggesting that the attention a CS receives is greater the worse predictor of the outcome the CS is. In this model, a CS loses associability the better predictor of the US becomes. Likewise, poor predictors of the outcome, gain associative capacity. In other words, bad predictors are expected to be more attended than good ones (Pearce & Hall, 1980). For example, Kaye and Pearce (1984) found that that the orientation response towards a light showed an inversely proportional relationship to its predictive capacity. They used rats that were trained under different reinforcement schedules (continuous versus partial), showing that for those animals that were under partial reinforcement, the orientation response towards the light remained higher than those that were trained under a continuous reinforcement program (Kaye & Pearce, 1984, Experiments 3 and 4). This result suggests that animals paid more attention to those cues that did not predict accurately the outcome. In the original model formulation [$\Delta V = \alpha$ (λ

$- \Sigma V)]$, α represents the associability of the trained cue and it is determined by the evolution of the prediction error, as it is shown in the following formula: $\alpha_n = | \lambda - \Sigma V |_{n-1}$; λ is the intensity of the US and ΣV is the summatory of the associative strength of all the cues present in the situation. This associability will grow when an unexpected outcome occurs and, on the contrary, it will tend to decrease when an expected outcome occurs. More specifically, the associability for stimulus X in trial n will be determined by the absolute value of the discrepancy between what the organism receives and what the organism expects $|\lambda - V_X|$ before the occurrence of X, ($\alpha_{Xn} = |\lambda - V_X|_{n-1}$). Thus, the associability of this stimulus will increase when it is followed by an unexpected EI (when the value of $|\lambda - V_X|$ is high), and it will decrease when it is followed by an expected EI (when the value of $|\lambda - V_X|$ is low). These attentional changes are shown in the α value since α is the element that indicates the absolute value of the predictive error in each trial (because $\alpha = |\lambda - \Sigma V|$) (Schulz & Dickinson, 2000). This model predicts that the associability of the trained cue will increase depending on how unexpected the US is. Therefore, for expected stimuli, the $|\lambda - V|$ will be low and, therefore, less attended. Alternatively, for unexpected unconditioned stimuli, the $|\lambda - V|$ will be high and, consequently, more attended. As in Mackintosh's (1975) model, there is a close relationship between prediction error and the associability of the CS and, thus, between prediction error and learning. This close relationship is indirect, as prediction error does not directly affect the learning rate, but the associability of the stimuli that is inversely related to their predictive power.

Larrauri and Schmajuk (2008). More recently, other models have suggested the existence of a relationship between attention and novelty. Larrauri and Schmajuk (2008) suggested that attention will be high at the beginning of the task, that is, when the information is completely novel. Furthermore, animals are assumed to respond to novelty with greater attention to the general environment, and not only to the CS involved in the novelty situation. Novelty is assumed to be altered by increases in the prediction error, as for example occurs when an associative interference procedure is introduced (Schmajuk, Lam, & Gray, 1996).

Le Pelley, Mitchell, Beesley, George, and Wills (2016) hybrid attentional model. These authors presented a hybrid model that allows for reconciling the opposite predictions of Mackintosh and Pearce and Hall models. This hybrid model assumes that there are two different attentional mechanisms that enter into play under different circumstances: Attentional exploration an Attentional exploitation. Attentional exploration mechanism will be engaged in those situations in which the organism needs to search for accurate predictors of the outcome. Attentional exploitation mechanism would be activated once the good predictors have been identified. Thus, increases in prediction error would lead the attentional exploration mechanism to enter into play, while decreases in prediction error would lead to activation of the attentional exploitation mechanism.

Finally, there are models in the literature that incorporate additional variables such as the relevance of the temporal factors involved in learning (Balsam, Drew, & Yang, 2002), the relevance of comparator processes at testing (e.g., Denniston, Savastano, & Miller, 2001; Miller & Matzel, 1988; Stout & Miller, 2007), the role of context (Bouton, 1993, 1997; Rosas, Callejas-Aguilera, Ramos-Álvarez, & Abad, 2006), or the role of different brain areas on variations and changes in the prediction error (Schultz, 2016; Schultz & Dickinson, 2000). At any rate, the relationship between the prediction error, the delta rule, and the learning process seems quite clear from most associative perspectives. The different learning models that are based on the prediction error agree that the appearance of the outcome must be surprising for learning to take place (Schultz & Dickinson, 2000). That surprise is operationalized, one way or the other, as the prediction error (λ - ΣV). This theoretical approach to prediction error has two main implications: a) learning appears as a way to correct prediction error, and b) prediction error also has a direct impact on the attention the organism pays to the signal or to the environment (and thus it does also have an indirect impact on learning by modulating the attention received by potential CSs).

1.1) Biological basis of learning mechanisms based on prediction error

Aside the multiple behavioral tests associative learning models have received, their principles have been also tested in the biological field in which the idea of associative strength as a CS-US link is translated in joint pre- and post-synaptic activity (Schultz & Dickinson, 2000). To our knowledge, Widrow and Hoff (1960) were the first authors in proposing a relationship between prediction error and neural substrates. This idea was tested years later after the advances in neuroimaging techniques. In the brain, there are multiple units that work simultaneously: Neurotransmitters, brain areas and systems. The role of each of these elements in terms of error prediction is different and, in some cases, their complete implication is still to be determined. In parallel, and related to prediction error, certain neural bases are also involved in attention. Although the study of biological correlates of behavioral effects is beyond the scope of this doctoral dissertation, it seems appropriate to give a general overview of the biological substrates under both, prediction error and attentional mechanisms.

Neurotransmitters and prediction error. The chosen task seems to be crucial to establish which area or neurotransmitter is activated and when. For example, using a go-no go procedure in monkeys, it was observed that dopaminergic neurons in ventral striatum are activated during reinforcement, coding the prediction error when it takes place (Ljungberg, Apicella, & Schultz, 1992; Schultz, Apicella, & Ljungberg, 1993). Thus, it seems that the activity of dopaminergic neurons is involved in prediction of which reinforcer is expected, and in when the reinforcer is expected to appear. Also, monkeys' dopaminergic neurons seem to show great precision on detecting novel or intense conditioned stimuli in a motor task (Ljungberg et al., 1992). However, monkey dopaminergic neurons seem to show low effectiveness in working with conditioned aversive stimuli and with unpredicted rewards (Mirenowicz & Schultz, 1994). Accordingly, Mirenowicz and Schultz (1996) suggested that dopamine response to aversive stimuli seems to be a motivational response. According to formal assumptions of attentional theories, omitting reinforcements should promote changes in neural bases that point in the same direction that unexpected reinforcers do, as they are based in the same theoretical principle, increases of prediction error. However, this is not observed in dopaminergic activity. The introduction of a delay time interval in the arrival of the reinforcer produces a dopamine activity depression when the reinforcer is expected. In contrast, an extra activation with the arrival of the unexpected reinforcer takes place (Hollerman & Schultz 1998).

Neurotransmitters and attention. Norepinephrine neurons focus their temporary activity when the meaning or occurrence of stimuli changes. Using monkeys in a space vigilance task, locus coeruleus norepinephrine neurons were observed to activate selectively on those cues that were attended (Aston-Jones, Rajkowski, Kubiak, & Alexinsky, 1994). In addition, norepinephrine neurons can acquire quickly new responses during a reversal discrimination procedure. Norepinephrine neurons are also responsible for changing the attention to new predictors in this type of learning tasks (Aston-Jones, Rajkowski, &

Kubiak, 1997). Furthermore, norepinephrine neurons seem to play a role in those events that are unexpected and can discriminate between neutral events and those that have a motivating component (Aston-Jones et al., 1994). Thus, it seems that norepinephrine neurons play an activating function of attention (Schultz & Dickinson, 2000).

Brain areas and prediction error. Nucleus basalis activity is related to the presence of good predictors and primary reinforcers. Richardson and DeLong (1986, 1990) have showed that the neurons of the nucleus basalis were activated by visual cues that accurately predict the outcome (see also, Mitchell, Richardson, Baker & DeLong, 1987; Richardson & DeLong, 1986). These results suggest that the activity of the nucleus basalis is determined by the precision of the organism to predict the reinforcement occurrence. When the organism can accurately predict the outcome, activity of the nucleus basalis is high. Hence, the relationship between the nucleus basalis and the prediction error seems to be inversely proportional: The lower the prediction error is, the greater the activation of the nucleus basalis.

Climbing fibers are also involved in learning (Marr, 1969). Marr found that prediction error can guide the development of learning at the brain through the projection that the climbing fibers made from the inferior olive to the Purkinje neurons in the cerebellar cortex. Furthermore, in studies with monkeys, the climbing fibers activation is greater when animals execute tasks in which they must adapt their behavior according to visual feedback (Gilbert & Thach, 1977; Ojakangas & Ebner, 1992). Their activation varies according to the error magnitude to achieve the visual targets (Kitazawa, Kimura, & Yin, 1998) and, finally, they are responsible to carry out errors between eye and target positions in visual tasks where the animals had to track visual target cues which were interrupted by errors that made tracking difficult (Kettner et al., 1997). In this sense, their way of pointing out errors is similar to the way in which the Delta rule (Kawato & Gomi, 1992) does. However, there are some results that also suggest that climbing fibers are activated by aversive stimuli that cannot be predicted (Kim, Krupa, & Thompson, 1998; Thompson & Gluck 1991), what suggests a way of working in aversive conditioning tasks that approaches Rescorla and Wagner's rule (Thompson & Gluck 1991). In similar lines, the middle layer of the superior colliculus fires when a predicted visual stimulus occurs (Mays & Sparks 1980). These neurons code the difference between the current and future eye position. Also, his activity increases when targets are extinguished. These features suggest that the prediction error could be better described by the delta rule between the visual cue and its current state (Schultz & Dickinson, 2000).

Finally, the frontal cortex is differentially activated in specific areas when prediction error changes. The anterior cingulate cortex is involved in the error signals processing during the execution of motor tasks. For example, in humans, in tasks in which well-trained motor cues are suddenly changed, there is an increase in blood flow (Berns, Cohen, & Mintun, 1997). Likewise, these changes could be reflected in the comparisons established between current and future behavior (Carter et al., 1998). In the dorsolateral prefrontal cortex of monkeys and using a modified differential reinforcement of long latencies task, neurons fire either when an error is made or when the experimenter removes the expected reward (Niki & Watanabe 1979). Thus, it seems that the prefrontal cortex is involved in both, coding performance and prediction errors, even though these errors occur in different neural populations. Related to the orbitofrontal cortex, blood flow increases when an error occurs in the time and space prediction in visual attention tasks with humans (Nobre, Coull, Frith, & Mesulam, 1999). These neurons are also activated in monkeys when reinforcement is unexpected and when the introduction of novelty changes the expectation of reinforcement (Tremblay & Schultz, 1999). Finally, frontal eye neurons show differences between current and future eye positions in spatial processing tasks in monkeys. These differences are like performed by the superior colliculus (Umeno & Goldberg 1997).

Regardless of the specific function of the areas, systems and neurons, they all seem to act as an organized system that uses the prediction error as the organizing element of the neuronal function. In the same way that the prediction error serves as an activator of attention at a behavioral level, the prediction error appears to play a regulatory role in neuronal processing (Schultz & Dickinson, 2000).

2. Prediction error after associative interference

We have described the role played by prediction error in the main associative models of learning (Mackintosh, 1975; Pearce & Hall, 1980; Rescorla & Wagner, 1972). We have also presented an overview of the neural bases that seem to underlie prediction error. To activate prediction error is necessary that environment changes, disconfirming what the organism expects. This sudden increases in prediction error may occur either because a CS is followed by an unexpected US at the beginning of learning, but also when a CS that was usually followed by a US is either no longer followed by it (extinction) or followed by a different US (extinction and interference). These changes may occur naturally (the environment where the organism develops its daily activity suddenly changes) or they can be introduced through artificial procedures in the laboratory. Procedures that lead to a sudden increase in prediction error are usually named as associative interference procedures.

2.1 What is an associative interference?

Learning about regularities requires that the elements to be learned remain stable over time. Interference appears when there is an alteration in the previously established associations. Polack, Jozefowiez, and Miller (2017, p. 129) described associative interference as a "*decrement in performance on a target learning task resulting from a specific memory conflict between memory of the target training and nontarget training events that have some similarity to target training, but also some unique components*". Polack et al. (2017) point out that interference may appear due to multiple memory conflicts in the natural environment. However, researchers usually have focused their studies on interference in only one of the instances to simplify the analysis. Along similar lines, Bouton (1993), while exploring what the interference paradigms have in common, points out that "*the common feature of the paradigms is that the significance of a conditioned stimulus (CS), or its association with some other event or outcome (typically an unconditioned stimulus, or US), changes between phases of the experiment*" (Bouton, 1993, p. 80).

Associative interference may be defined as a procedure, as the behavioral result of applying such as procedure to an organism, and as the internal mechanism that produces such a change in the organism behavior once the interference procedure is applied. Specifically, associative interference is any procedure that changes the information the organism receives over time, leading to both, a difficulty in recovering the information learned in the first place (retroactive interference) or the second information learned (proactive interference). Finally, associative interference requires sequential and consistent training. Taking in account these premises, the number of procedures or parametric variations to produce interference is quite large, including blocking (Vadillo, Castro, Matute, & Wasserman, 2008) and other cue competition phenomena. However, the basic general principles underlying behavioral effects of proactive and retroactive interference are assumed to be similar (or the same) regardless of the specific procedure used. Accordingly, we will focus our description in the most used in research such as extinction, counterconditioning, reversal discrimination, and latent/differential inhibition, as this latter procedure will be the one used in the experimental sections of this dissertation.

Extinction. Extinction consists of a two-phase training. In the first phase, the target cue is trained with an outcome (Cue-Outcome). In a second phase, this association is broken, so that the target cue is no longer followed by the outcome (Cue-NoOutcome) (Pavlov, 1927). Extinction, in addition, has been the most used procedure to study multiple associative interference phenomena such as spontaneous recovery, disinhibition, reinstatement or renewal (Rescorla, 2000).

Counterconditioning. In a standard counterconditioning procedure, the target cue is paired with an outcome with a given motivational during the first phase. During the second phase, the same target cue is trained with an outcome of an opposite motivational value (e.g., Pearce & Dickinson, 1975). Counterconditioning is a common procedure in fear conditioning experiments (see De Jong, Vorage & Van den Hout, 2000; Holmes, Leung, & Westbrook, 2016). The order to produce counterconditioning can be varied (from appetitive to aversive conditioning or vice versa), although altering the presentation sequence seems to be decisive in the later implications that this learning has (e.g., Bouton & Peck, 1992; Krank, 1985).

Reversal discrimination. This procedure combines extinction and new excitatory conditioning. Let's take an example of classical conditioning. During Phase 1, two cues are trained with the presence or the absence of the outcome (or with two different outcomes): Cue 1 will be followed by the outcome (Cue 1-Outcome) while Cue 2 will not be followed by the outcome (Cue 2-No Outcome). In Phase 2, the relationship that both cues maintain with the outcome is reversed; Cue 1 will be extinguished (Cue1-NoOutcome) while Cue2 will be conditioned with the outcome (Cue2-Outcome). This type of interference has been implemented in multiple tasks such as avoidance learning (Spear et al., 1980), appetitive conditioning with pigeons (Thomas, Moye, & Kimose, 1984) and conditioned supression in rats (Bouton & Brooks, 1993).

Latent/differentia inhibition. We could describe latent inhibition procedure as the mere exposure of a cue before starting conditioning with it. As a behavioral phenomenon, latent inhibition shows as a retardation of conditioned responding when the exposed cue is later paired with an outcome, compared to a cue that was never exposed (Lubow & Moore, 1959). Though the explanation of latent inhibition is still under discussion (Lubow, 1989), it has been suggested that the lack of changes in the environment during the exposure phase leads the organism to ignore the exposed cue (Lubow, 1989; Lubow & Moore, 1959). It is important to distinguish between a latent inhibitor and a conditioned inhibitor, especially because simple exposure of the CS when other CS are presented followed by the outcome may end conducting to what has been called differential inhibition (O'Boyle & Bouton, 1996; Rescorla & Lolordo, 1965). While the latent inhibitor does not predict anything about the consequence (at least, not in a directly observable way), a conditioned inhibitor signals the outcome absence. Latent inhibitors tend be ignored as they do not predict neither the presence, nor the absence of the outcome; on the contrary, in conditioned inhibition training, the cue becomes a good predictor of the absence of the outcome. Given that the same procedure may result either in a latent or in a differential inhibitor, testing conditioned inhibition needs complementary test procedures. The traditional test of latent inhibition is a retardation test, in which latent inhibitors are expected to condition slower than neutral stimuli. But this test is also passed by a conditioned inhibitor. Accordingly, to detect whether a cue is a latent inhibitor or a conditioned inhibitor, two tests are usually used, the retardation and summation tests (Rescorla, 1969). Retardation test makes possible to observe whether a cue behaviorally presents a delay in its acquisition, although this delay alone does not allow to delimit its nature. Retardation test consists of the comparison of target training cue versus a novel cue when both are followed by the outcome. If the target cue training proceeds slower than novel cue we can affirm that our procedure produces some type of inhibition, but to detect whether such inhibition is latent (Lubow & Moore, 1959) or conditioned (Williams, Overmier, & LoLordo, 1992) we will need and extra confirmation: the summation test. Summation test is based on the idea that a conditioned inhibitor would subtract excitatory capacity from a cue that is already an excitor. Therefore, the simultaneous presentation of an excitatory cue plus the potential conditioned inhibitor should produce a decrease in the conditioned response compared to that produced after the presentation of the excitatory cue plus a novel one (Pavlov, 1927; Rescorla, 1969). When a cue passes the retardation test but does not pass the summation one it is categorized as a latent inhibitor; if the cue passes both tests, it is considered a conditioned inhibitor.

2.2 Associative interference and prediction error. Implications.

Interference treatment, as its name suggests, introduces a change in the regularities that were already established through learning. Detecting these changes is also part of the learning mechanism and it is an adaptive advantage. However, this detection change may have implications that go beyond the ability to perceive that something has switched. Bouton (1997) suggested that the interference treatment generates an ambiguity in the information that causes it to become dependent on the context in which it is acquired. This ambiguity can be generated through any treatment to generate associative interference and it is assumed that it would only affect retrieval of the information involved in the interference treatment. More recently, Rosas et al. (2006) suggested that ambiguity does not only affect retrieval of the information involved in the interference treatment, but all the information learned acquired shortly after the interference treatment, regardless of the whether the information was involved in the interference treatment or not. Thus, ambiguity generated by the associative interference treatment produces a change in what is being learned, as the organism begins to pay attention to the context where the information is acquired, and that context becomes a relevant factor for retrieval of the information. An example of how such ambiguity plays a role on retrieval of the information can be clearly seen in the extinction makes acquisition context specific (EMACS) effect (Rosas & Callejas-Aguilera, 2006) which has been tested in humans (e.g., Ogállar et al., 2019; Rosas & Callejas-Aguilera, 2006; for a recent results) and animals (e.g., Bernal-Gamboa, Rosas, & Nieto, 2018). The extinction of a previously trained cue in phase 1 produces, in the excitatory training of a different cue during phase 2, its conversion into dependent on the context where it is trained (Rosas & Callejas-Aguilera, 2006).

The potential mechanism that explains the impact of ambiguity on context-dependence of learning has been under discussion for the last three decades. Bouton (1993) suggested that it was inhibitory or second-learned information what become context-specific. The reason because second-learned information could become context specific was given by Bouton (1997). He suggested that ambiguity leads the organism to pay attention to the context where learning is taking place trying to break the ambiguity of the situation. Paying attention to the context is assumed to make inhibitory information, or second information learned (e.g., Nelson, 2002), context-dependent, while retrieval of first learned information does not depend on the context where such retrieval is tested, as such information is not part of the conflicting information that is assumed to become context-dependent (Bouton, 1997). This idea cannot explain why simple acquisition becomes context-dependent after the extinction treatment (the EMACS effect, Bernal-Gamboa et al., 2018; Ogállar et al., 2019; Rosas & Callejas-Aguilera, 2006). Rosas et al. (2006) suggested a role for ambiguity that allowed to explain renewal and EMACS effects with the same mechanism. These authors suggested that ambiguity leads the organism to pay attention to all the information that is present during the task, and not only to the ambiguous one, so that all the information in the situation is coded as context dependent (see Ogállar, Ramos-Álvarez, Alcalá, Moreno-Fernández, & Rosas, 2017, for a recent review).

In addition to associative interference, the introduction of a novel stimulus can produce an increase in attention. Specific, Schmajuk and Larrauri (2006) suggested that novelty is an essential factor on the attention that animals pay to a learning situation. According to these authors, novelty is defined in a broad way; that is, novelty understood as total inexperience with the information that is being presented. In this way, the mere introduction of an unknown cue would produce an increase in novelty that would, in turn, increase attention. Therefore, the greater the novelty, the greater the attention the organisms pay to the situation and this could improve subsequent learning (Alcalá, González, Aristizabal, Callejas-Aguilera, & Rosas, 2018). Schmajuk et al. (1996) suggest that the effects produced by phenomena such as extinction overtraining (North & Stimmel, 1960) or saving effects in extinction-to-acquisition experiments (Frey & Ross, 1968) can be explained thanks to the relationship between novelty and the CS. For this reason, novelty seems responsible for, at least under certain parameters, directing attention according to the learning task. Also, as organisms learn about the new learning situation, novelty, and attention decrease (Schmajuk et al., 1996).

Taking together the concepts of associative interference and prediction error, it seems reasonable to suggest that the prediction error is the mechanism responsible for activating the attentional resources. In the case of Bouton (1993, 1994, 1997) and Rosas et al. (2006) studies, the introduction of associative interference seems responsible to convert the learned information into context-dependent. Additionally, the introduction of new information also produces an increase in attention (Schmajuk & Larrauri, 2006). Therefore, it seems reasonable to suggest that any change in prediction error, regardless of whether it is due to interference or novelty, produces an increase in attention. Hence, attention, prediction error and associative interference are related. Even more. The implications that the prediction error has both in the nature of the tasks and in their consequences, makes us think that the change in the prediction error affects attention in a generalized way.

Recent studies suggest that the increased attention that associative interference produces may indeed lead to a general improvement in subsequent learning; not only as far as the context is concerned, but to the whole learning situation (e.g., Alcalá, Callejas-Aguilera, Lamoureux, & Rosas, 2019; Alcalá, Callejas-Aguilera, Nelson, & Rosas, 2020; Alcalá et al., 2018). The idea that underlies the hypothesis is that the increase in prediction error that follows an associative interference treatment would would produce a general increase in attention so that the organism would start paying attention to all the information present in a learning situation in comparison with the same learning situation without the interference treatment. If that were the case, a general improvement in learning after an interference treatment would be expected. This hypothesis has been tested in temporal conditioning (Alcalá et al, 2019), and in spatial learning (Alcalá et al., 2020), using Wistar rats as subjects in both studies.

In the temporal discrimination study, three experiments were conducted to evaluate the facilitation of subsequent learning after the introduction of an associative interference experience, in this case, a discrimination reversal training (Alcalá et al., 2019). Animals were trained in Skinner boxes with two different CS (a tone and a light), one was paired with food, and the other was not. The associative interference consisted of the reversal of the Cue-Outcome relationship in effect during the first phase. Rats in group R (reversal) received the associative interference treatment while group A (Acquisition) was kept under the original discrimination training. In the test phase, half of the animals in each group were trained in a situation in which food was delivered under a fixed interval program 60" (R-FT and A-FT, respectively) while the other half were trained under a variable time program 60" (R-VT and A-VT, respectively). Experiment 1 found that there were not differences between groups R and A did not differ in performance under a variable time food delivering schedule, as expected, because there is nothing in the learning situation that sign when the US is presented and no temporal discrimination was observed. However, group R learned faster than group A the temporal discrimination when food was delivered under a fixed time interval; both groups, R and A, end showing a temporal discrimination performance, but this temporal discrimination was acquired faster in the group that had the associative interference experience. Experiments 2 and 3 confirmed these results extending them to 30-sec interval suggesting the experience of interference that is produced by discrimination reversal facilitates subsequent learning about a temporal discrimination in rats (Alcalá et al., 2019).

In the spatial learning study, Alcalá et al. (2020) conducted two experiments with the goal of evaluating whether the experience of interference would facilitate subsequent learning about a new spatial location of the platform in a Morris water maze. Associative interference was produced by moving the position of the platform in the Reversal group

(R) across two different training phases. In the first phase, the platform was presented in a position equidistant and near to two visual cues. In the second phase, the location of the platform was placed equidistant but far away from these same cues. Animals in the control conditions received the same treatment as R group received in the first training phase, so that no interference was introduced. During the test, the platform was placed in a new position with respect to a different pair of cues from the ones used during training Results showed that animals in group R found the new platform faster than controls, suggesting that previous experience of interference facilitated subsequent learning within the spatial domain. This result was replicated in Experiment 2 in which alternative explanations for the results of Experiment 1 were controlled.

There are a few studies in the animal learning literature that are also compatible with this idea of improvement in learning after interference experience. Shanab and Cotton (1970) found that experiencing extinction in a runway facilitated subsequent acquisition of a T-maze discrimination and vice versa. De la Casa, Mena, Ruiz-Salas, Quintero, and Papini (2018) suggests that the effects of SNC could go further than a simple increase in the exploratory behavior. In this experiment, authors used a consummatory Successive Negative Contrast (cSNC), trained animals with a dissolution of 32% sucrose before giving them access to a dissolution of 4% sucrose. Control group received access to the dissolution of 4% sucrose from the start. Downshifting the value of the reward produced the standard drop in consumption in the experimental group below consumption of the control animals. After a session of downshift, half of the animals in each group were exposed to a conditioned stimulus (CS), while the other half received equivalent exposure to the context. The retardation of conditioning that follows nonreinforced exposure to the CS was attenuated in the animals that had the cSNC experience, suggesting that experiencing cSNC somewhat preserved the associability of a pre-exposed CS. In summary, improvement of learning after experiencing associative interference appears in animals in different tasks: Spatial learning in a water maze (Alcalá et al., 2020) and in a T-maze (Shanab & Cotton,1970), appetitive temporal discrimination (Alcalá et al., 2019) and after successive negative contrast procedures (De la Casa et al., 2018).

The effect of experiencing associative interference facilitating later learning has been also shown in one study in humans. Nelson, Fabiano and Lamoreux (2018) conducted an experimental series with humans to test the effect experiencing extinction on subsequent learning about a temporal discrimination. They trained color sensors (cues) that signaled the arrival of an enemy spacecraft (outcome) in a video game. In the task, participants should respond to the proper sensors to prevent the arrival of the enemy spacecraft. The authors established two groups: After Extinction (AE) and No Extinction (NE). Both groups received presentations of a sensor followed by the outcome and, in phase two, the AE group extinguished the same previously trained cue while the NE group remained the same. Finally, in the test phase, 8 unsignaled outcomes appeared under a 20 second fixed-interval schedule, so that, temporal learning about unsignaled outcomes was evaluated. Therefore, if extinction produces an increase in attention, discriminating temporal learning about unsignaled outcomes should occur faster in the AE group than in the NE group. Results in Experiment 1 showed that AfterExtinction group learned faster about unsignaled outcomes during the test phase versus the No Extinction group.

Studies with humans on memory tasks seem to yield results that may be compatible with the idea of interference facilitating subsequent learning. Potts and Shanks (2014) suggested that prediction error may be a useful mechanism to promote learning. These authors conducted four experiments in which participants had to learn definitions about unusual words in English. In Experiment 1, all participants, during the study phase, had to learn all the words from a list of 60 which appeared once using the three possible teaching methods: Read, Generate and Choice. Read, consisted of presenting the word to learn plus its definition for 17 seconds; Generate condition consisted of presenting the word to be learned for 10 seconds and, after this time, the participant had to define it with another word. After issuing the response, the correct definition was presented for 7 seconds. Finally, the Choice condition consisted of the presentation of the word to be learned and the presentation of two possible definitions simultaneously, one correct and the other incorrect. After the selection, participants received 7 seconds of feedback. Finally, the test consisted of presenting the 60 words randomly with four answer options (one correct and three incorrect). In the Choice condition, one of the incorrect answers was the one that had been trained in the study phase. Results showed that, in all experiments, Generate condition proved better performance remembering definition words on the list. More concretely, in Experiment 1, participants showed a higher percentage of correct answers in the Generate condition, even if the definition suggested by the participants was incorrect. These results were replicated using different training time parameters (Experiment 2a), using another foreign language (Experiment 2b) or using experimental groups where alternative explanations were controlled (Experiment 3). Pott and Shanks (2014) suggested as an explanation that the discrepancy between the response that the participants make, and the correct response generated a surprise that ultimately results in an improvement in later learning (see also Seabrooke, Hollins, Kent, Wills, & Mitchell, 2019).

In summary, there are several animal and human studies in the literature that suggest that experiencing interference may facilitate subsequent new learning. These positive results of learning facilitation have two characteristics: First, those studies used complex learning phenomena; and second, associative interference is always involves eliminating a previously established learning about a relationship between the cue and the outcome. Related to the first feature, learning facilitation after associative interference have been found in temporal conditioning (Alcalá et al., 2019), spatial learning (Alcalá et al., 2020), consummatory Successive Negative Contrast (De la Casa et al., 2018), memory tasks (Potts & Shanks, 2014; Seabrooke et al., 2019). Related to the direction of the associative interference, all of those studies involved going from some form of predicting the outcome to breaking that prediction or predicting a new outcome. However, the increase in the prediction error could be also produced by treatment involving increasing the predictive value of a cue, and the effects of that type of interference have not been yet tested, as far as we know.

Additionally, learning facilitation after associative interference results are not ubiquitous. This phenomenon has not been found in a simpler procedure, such as the one that involves the direct association between one cue and its consequence (Alcalá et al., 2018). Alcalá et al. (2018) reported that the experience of reversal discrimination in appetitive conditioning led to an improvement in contextual conditioning, but not in learning about the relationship between a new CS and the US. This study was conducted with rats in Skinner boxes and consisted in an appetitive conditioning task. Specifically, rats performed a simple discrimination in a first phase (CS1-US, CS2-No US). In the second phase, a reversal discrimination was introduced for the Reversal group (CS1-No US; CS2-US), while the Control group remained being trained with the original discrimination. Training of a CS3-US relationship during phase two proceeded similarly regardless of the group, though the interference experience led to an increase on direct responding to the context. These increase in responding may be caused by an increase in the exploration pattern of behavior that has been suggested occurs when the organism experience associative interference, but also by the introduction of novelty (Beesley, Nguyen, Pearson, & Le Pelley, 2015). The new CS could have led to an increase in the exploration responding in both groups, regardless of the associative interference treatment. Additionally, the new stimulus may draw enough attention because of its novelty as not to be increased by the associative interference experience. Results of this experimental series cast doubts with respect to whether facilitation of subsequent learning after an interference experience may occur when simple cue-outcome relationships are trained.

It is possible that, to observe a facilitation of subsequent learning, the use of pre-exposed cues could be crucial. Whether as a latent inhibitor or a conditioned inhibitor, the pre-exposure of the cues solves the problem of novelty and, in addition, it would provide a more sensitive measure to detect the facilitation of later learning after associative interference. So, the studies of this thesis will try to shed some light on the difficulties of finding a facilitation of learning on different cues, expand the benefits of interference treatment and address new procedures to test associative interference and its consequences. In addition, the studies with prediction error manipulation and associative interference would allow to establish new boundaries on the scope of these phenomena.

References

Alcalá, J. A., Callejas-Aguilera, J., Lamoureux, J. A., & Rosas, J. M. (2019). Discrimination reversal facilitates subsequent acquisition of temporal discriminations in rats' appetitive conditioning. *Journal of Experimental Psychology: Animal Learning and Cognition, 45*, 446-463. doi: 10.1037/xan0000216

Alcalá, J. A., Callejas-Aguilera, J., Nelson, J. B., & Rosas, J. M. (2020). Reversal training facilitates acquisition of new learning in a morris water maze. *Learning & Behavior, 48*(2), 208-220. doi: 10.3758/s13420-019-00392-7

Alcalá, J. A., González, G., Aristizabal, J. A., Callejas-Aguilera, J. E., & Rosas, J. M. (2018). Discrimination reversal facilitates contextual conditioning in rats' appetitive conditioning. *Psicológica, 39*(1), 64-87. doi: 10.2478/psicolj-2018-0004

Aston-Jones, G., Rajkowski, J., & Kubiak, P. (1997). Conditioned responses of monkey locus coeruleus neurons anticipate acquisition of discriminative behavior in a vigilance task. *Neuroscience, 80*(3), 697-715. doi: 10.1016/S0306-4522(97)00060-2

Aston-Jones, G., Rajkowski, J., Kubiak, P., & Alexinsky, T. (1994). Locus coeruleus neurons in monkey are selectively activated by attended cues in a vigilance task. Journal of *Neuroscience, 14*(7), 4467-4480. doi: 10.1523/JNEUROSCI.14-07-04467.1994

Balsam, P. D., Drew, M. R., & Yang, C. (2002). Timing at the start of associative learning. *Learning and Motivation, 33*(1), 141-155. doi: 10.1006/lmot.2001.1104

Beesley, T., Nguyen, K. P., Pearson, D., & Le Pelley, M. E. (2015). Uncertainty and predictiveness determine attention to cues during human associative learning. *Quarterly Journal of Experimental Psychology, 68*(11), 2175-2199. doi: 10.1080/17470218.2015.1009919

Bernal-Gamboa, R., Rosas, J. M., & Nieto, J. (2018). Extinction makes acquisition context-specific in conditioned taste aversion regardless of the context where acquisition and testing take place. *Journal of Experimental Psychology: Animal Learning and Cognition, 44*(4), 385-395. doi: 10.1037/xan0000183

Berns, G. S., Cohen, J. D., & Mintun, M. A. (1997). Brain regions responsive to novelty in the absence of awareness. *Science, 276* (5316), 1272-1275. doi: 10.1126/science.276.5316.1272

Bouton M. E. (1993). Context, time, and memory retrieval in the interference paradigms of Pavlovian learning. *Psychological Bulletin, 114,* 80-99. doi: 10.1037//0033-2909.114.1.80

Bouton M. E. (1994). Conditioning, remembering, and forgetting. *Journal of Experimental Psychology: Animal Behavior Processes, 20*, 219-231. doi: 10.1037//0097-7403.20.3.219

Bouton, M. E. (1997). Signals for whether versus when an event will occur. In M. S. Fanselow & M. E. Bouton, *Learning, motivation, and cognition: The functional behaviorism of Robert C. Bolles* (pp. 385–409). Washington, DC: American Psychological Association.

Bouton, M. E., & Brooks, D. C. (1993). Time and context effects on performance in a Pavlovian discrimination reversal. *Journal of Experimental Psychology: Animal Behavior Processes, 19*(2), 165-179. doi: 10.1037/0097-7403.19.2.165

Bouton, M. E., & Peck, C. A. (1992). Spontaneous recovery in cross-motivational transfer (counterconditioning). *Animal Learning & Behavior, 20*(4), 313-321. doi: 10.3758/BF03197954

Carter, C. S., Braver, T. S., Barch, D. M., Botvinick, M. M., Noll, D., & Cohen, J. D. (1998). Anterior cingulate cortex, error detection, and the online monitoring of performance. *Science, 280* (5364), 747-749. doi: 10.1126/science.280.5364.747

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*(11), 1055-1069. doi: 10.1016/S0005-7967(99)00135-7

De la Casa, L. G., Mena, A., Ruiz-Salas, J. C., Quintero, E., & Papini, M. R. (2018). Reward devaluation disrupts latent inhibition in fear conditioning. *Learning & behavior, 46*(1), 49-59. doi: 10.3758/s13420-017-0282-1

Denniston, J. C., H. I. Savastano, & R. R. Miller. (2001). Learning by contiguity, responding by relative strength: The extended comparator hypothesis. In *Handbook of Contemporary Learning Theories,* ed. RR Mowrer, SB Klein, (pp. 65–117). Mahwah, NJ: Erlbaum.

Frey, P. W., & Ross, L. E. (1968). Classical conditioning of the rabbit eyelid response as a function of interstimulus interval. *Journal of comparative and physiological psychology, 65*(2), 246-250. doi: 10.1037/h0025555

Gilbert, P. F. C., & Thach, W. T. (1977). Purkinje cell activity during motor learning. *Brain research, 128*(2), 309-328. doi: 10.1016/0006-8993(77)90997-0

Hall, G., & Rodriguez, G. (2010). Attentional learning. In C. Mitchell & M. Le Pelley (Eds.), *Attention and associative learning* (pp. 41-70). Oxford, England: Oxford University Press.

Hollerman, J. R., & Schultz, W. (1998). Dopamine neurons report an error in the temporal prediction of reward during learning. *Nature neuroscience, 1*(4), 304-309. doi: 10.1038/1124

Holmes, N. M., Leung, H. T., & Westbrook, R. F. (2016). Counterconditioned fear responses exhibit greater renewal than extinguished fear responses. *Learning & Memory, 23*(4), 141-150. doi: 10.1101/lm.040659.115

Kawato, M., & Gomi, H. (1992). A computational model of four regions of the cerebellum based on feedback-error learning. *Biological cybernetics, 68*(2), 95-103. doi: 10.1007/BF00201431

Kaye, H., & Pearce, J. M. (1984). The strength of the orienting response during Pavlovian conditioning. *Journal of Experimental Psychology: Animal Behavior Processes, 10*(1), 90-109. doi: 10.1037/0097-7403.10.1.90

Kettner, R. E., Mahamud, S., Leung, H. C., Sitkoff, N., Houk, J. C., Peterson, B. W., & Barto, A. G. (1997). Prediction of complex two-dimensional trajectories by a cerebellar model of smooth pursuit eye movement. *Journal of neurophysiology, 77*(4), 2115-2130. doi: 10.1152/jn.1997.77.4.2115

Kim, J. J., Krupa, D. J., & Thompson, R. F. (1998). Inhibitory cerebello-olivary projections and blocking effect in classical conditioning. *Science, 279*(5350), 570-573. doi: 10.1126/science.279.5350.570

Kitazawa, S., Kimura, T., & Yin, P. B. (1998). Cerebellar complex spikes encode both destinations and errors in arm movements. *Nature, 392*(6675), 494-497. doi: 10.1038/33141

Krank, M. D. (1985). Asymmetrical effects of Pavlovian excitatory and inhibitory aversive transfer on Pavlovian appetitive responding and acquisition. *Learning and Motivation, 16*(1), 35-62. doi: 10.1016/0023-9690(85)90003-7

Larrauri, J. A., & Schmajuk, N. A. (2008). Attentional, associative, and configural mechanisms in extinction. *Psychological review, 115*(3), 640-676. doi: 10.1037/0033-295x.115.3.640

Le Pelley, M. E., Mitchell, C. J., Beesley, T., George, D. N., & Wills, A. J. (2016). Attention and associative learning in humans: an integrative review. *Psychological Bulletin*, *142*(10), 1111-1140. doi: 10.1037/bul0000064

Ljungberg, T., Apicella, P., & Schultz, W. (1992). Responses of monkey dopamine neurons during learning of behavioral reactions. *Journal of neurophysiology, 67*(1), 145-163. doi: 10.1152/jn.1992.67.1.145

Lubow, R. E. (1989). *Latent inhibition and conditioned attention theory* (Vol. 9). Cambridge University Press.

Lubow, R. E., & Moore, A. U. (1959). Latent inhibition: the effect of nonreinforced pre-exposure to the conditional stimulus. *Journal of comparative and physiological psychology, 52*(4), 415-419. doi: 10.1037/h0046700

Mackintosh, N. J. (1975). A theory of attention: variations in the associability of stimuli with reinforcement. *Psychological review, 82*(4), 276-298. doi: 10.1037/h0076778

Marr, D. (1969). A theory of cerebellar cortex. *The Journal of Physiology*, *202*, 437-470. doi: 10.1113/jphysiol.1969.sp008820

Mays, L. E., & Sparks, D. L. (1980). Dissociation of visual and saccade-related responses in superior colliculus neurons. *Journal of Neurophysiology*, *43*(1), 207-232. doi: 10.1152/jn.1980.43.1.207

Miller, R. R., Barnet, R. C., & Grahame, N. J. (1995). Assessment of the Rescorla-Wagner model. *Psychological bulletin, 117*(3), 363-386. doi: 10.1037/0033-2909.117.3.363

Miller, R. R., & Matzel, L. D. (1988). The comparator hypothesis: A response rule for the expression of associations. In *Psychology of learning and motivation* (Vol. 22, pp. 51-92). Academic Press.

Mirenowicz, J., & Schultz, W. (1994). Importance of unpredictability for reward responses in primate dopamine neurons. *Journal of neurophysiology*, *72*(2), 1024-1027. doi: 10.1152/jn.1994.72.2.1024

Mirenowicz, J., & Schultz, W. (1996). Preferential activation of midbrain dopamine neurons by appetitive rather than aversive stimuli. *Nature, 379*(6564), 449-451. doi: 10.1038/379449a0

Mitchell, S. J., Richardson, R. T., Baker, F. H., & DeLong, M. R. (1987). The primate globus pallidus: neuronal activity related to direction of movement. *Experimental Brain Research, 68*(3), 491-505. doi: 10.1007/bf00249793

Nelson, J. B. (2002). Context specificity of excitation and inhibition in ambiguous stimuli. *Learning and Motivation*, *33*(2), 284-310. doi: 10.1006/lmot.2001.1112

Nelson, J. B., Fabiano, A. M., & Lamoureux, J. A. (2018). The effects of extinction-aroused attention on context conditioning. *Learning & Memory, 25*, 165-175. doi: 10.1101/lm.046201.117

Niki, H., & Watanabe, M. (1979). Prefrontal and cingulate unit activity during timing behavior in the monkey. *Brain research, 171*(2), 213-224. doi: 10.1016/0006-8993(79)90328-7

Nobre, A. C., Coull, J. T., Frith, C. D., & Mesulam, M. M. (1999). Orbitofrontal cortex is activated during breaches of expectation in tasks of visual attention. *Nature neuroscience, 2*(1), 11-12. doi: 10.1038/4513

North, A. J., & Stimmel, D. T. (1960). Extinction of an instrumental response following a large number of reinforcements. *Psychological Reports, 6*(2), 227-234. doi: 10.2466/pr0.1960.6.2.227

O'Boyle, E. A., & Bouton, M. E. (1996). Conditioned inhibition in a multiple category learning task. *The Quarterly Journal of Experimental Psychology Section B*, *49*(1b), 1-23. doi: 10.1080/713932616

Ogállar, P. M., Ramos-Álvarez, M. M., Alcalá, J. A., Moreno-Fernández, M. M., & Rosas, J. M. (2017). Attentional Perspectives on Context-dependence of Information Retrieval. *International Journal of Psychology and Psychological Therapy,* *17*, 121-136. Retrieved from https://www.ijpsy.com/volumen17/num1/462/attentional-perspectives-on-context-dependence-EN.pdf

Ogállar, P. M., Rosas, J. M., Ramos-Álvarez, M. M., Alcalá, J. A., Nelson, J. B., Aranzubia, M., & Callejas-Aguilera, J. E. (2019). Prior extinction increases acquisition context specificity in human predictive learning. *Behavioural Processes, 169*, 103984. doi: 10.1016/j.beproc.2019.103984

Ojakangas, C. L., & Ebner, T. J. (1992). Purkinje cell complex and simple spike changes during a voluntary arm movement learning task in the monkey. *Journal of neurophysiology, 68*(6), 2222-2236. doi: 10.1152/jn.1992.68.6.2222

Pavlov, I. P. (1927). *Conditioned reflexes*. London: Oxford University Press. Retrieved from http://s-f-walker.org.uk/pubsebooks/pdfs/Conditioned-Reflexes-Pavlov.pdf

Pearce, J. M., & Dickinson, A. (1975). Pavlovian countercondition: Changing the suppressive properties of shock by association with food. *Journal of Experimental Psychology: Animal Behavior Processes, 1*(2), 170-177. doi: 10.1037/0097-7403.1.2.170

Pearce, J. M., & Hall, G. (1980). A model for Pavlovian learning: variations in the effectiveness of conditioned but not of unconditioned stimuli. *Psychological review, 87*(6), 532. doi: 10.1037/0033-295X.87.6.532

Polack, C. W., Jozefowiez, J., & Miller, R. R. (2017). Stepping back from 'persistence and relapse'to see the forest: Associative interference. *Behavioural processes, 141*, 128-136. doi: 10.1016/j.beproc.2017.03.014

Potts, R., & Shanks, D. R. (2014). The benefit of generating errors during learning. Journal of Experimental Psychology: General, *143*(2), 644-667. doi: 10.1037/a0033194

Rescorla, R. A. (1969). Pavlovian conditioned inhibition. *Psychological Bulletin, 72*, 77-94. doi: 10.1037/h0027760

Rescorla, R. A. (2000). Experimental extinction. In *Handbook of contemporary learning theories* (pp. 129-164). Psychology Press.

Rescorla, R. A., & Lolordo, V. M. (1965). Inhibition of avoidance behavior. *Journal of comparative and physiological psychology, 59*(3), 406-412. doi: 10.1037/h0022060

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 research and theory.* New York: Appleton-Century-Crofts.

Richardson, R. T., & DeLong, M. R. (1986). Nucleus basalis of Meynert neuronal activity during a delayed response task in monkey. *Brain research, 399*(2), 364-368. doi: 10.1016/0006-8993(86)91529-5

Richardson, R. T., & DeLong, M. R. (1990). Context-dependent responses of primate nucleus basalis neurons in a go/no-go task. *Journal of Neuroscience, 10*(8), 2528-2540. doi: 10.1523/JNEUROSCI.10-08-02528.1990

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), 461-474. doi: 10.1037/0278-7393.32.3.461

Rosas, J. M., Callejas-Aguilera, J. E., Ramos-Álvarez, M. M., & Abad, M. J. (2006). Revision of retrieval theory of forgetting: What does make information context-specific? *International Journal of Psychology and Psychological therapy, 6*, 147-166. Retrieved from https://www.ijpsy.com/volumen6/num2/136/revision-of-retrieval-theory-of-forgetting-EN.pdf.

Schmajuk, N. A., Lam, Y. W., & Gray, J. A. (1996). Latent inhibition: A neural network approach. *Journal of Experimental Psychology: Animal Behavior Processes, 22*(3), 321-349. doi: 10.1037/0097-7403.22.3.321

Schmajuk, N. A., & Larrauri, J. A. (2006). Experimental challenges to theories of classical conditioning: appli cation of an attentional model of storage and retrieval. *Journal of Experimental Psychology: Animal Behavior Processes, 32*(1), 1-20. doi: 10.1037/0097-7403.32.1.1

Schultz, W. (2016). Dopamine reward prediction-error signalling: a two-component response. *Nature Reviews Neuroscience, 17*(3), 183-195. doi: 10.1038/nrn.2015.26

Schultz, W., Apicella, P., & Ljungberg, T. (1993). Responses of monkey dopamine neurons to reward and conditioned stimuli during successive steps of learning a delayed response task. *Journal of neuroscience, 13*(3), 900-913. doi: 10.1523/JNEUROSCI.13-03-00900.1993

Schultz, W., & Dickinson, A. (2000). Neuronal coding of prediction errors. *Annual review of neuroscience, 23*(1), 473-500. doi: 10.1146/annurev.neuro.23.1.473

Seabrooke, T., Hollins, T. J., Kent, C., Wills, A. J., & Mitchell, C. J. (2019). Learning from failure: Errorful generation improves memory for items, not associations. *Journal of Memory and Language, 104*, 70-82. doi: 10.1016/j.jml.2018.10.001

Shanab, M. E., & Cotton, J. W. (1970). Effects of runway training on behavior in the T-maze. *Psychonomic Science, 19*(3), 129-130. doi: 10.3758/bf03335510

Siegel, S., & Allan, L. G. (1996). The widespread influence of the Rescorla-Wagner model. *Psychonomic Bulletin & Review, 3*(3), 314-321. doi: 10.3758/BF03210755

Spear, N. E., Smith, G. J., Bryan, R. G., Gordon, W. C., Timmons, R., & Chiszar, D. A. (1980). Contextual influences on the interaction between conflicting memories in the rat. *Animal Learning & Behavior, 8(2)*, 273-281. doi: /10.3758/bf03199606

Stout, S. C., & Miller, R. R. (2007). Sometimes-competing retrieval (SOCR): A formalization of the comparator hypothesis. *Psychological Review, 114*(3), 759-783. doi: 10.1037/0033-295X.114.3.759

Sutton, R. S., & Barto, A. G. (1981). Toward a modern theory of adaptive networks: expectation and prediction. *Psychological review, 88*(2), 135-170. doi: 10.1037/0033-295X.88.2.135

Thomas, D. R. (1970). Stimulus selection, attention, and related matters. In J. H. Reynierse (Ed.), *Current issues in animal learning*. Lincoln: University of Nebraska Press.

Thomas, D. R., Moye, T. B., & Kimose, E. (1984). The recency effect in pigeons' long-term memory. *Animal Learning & Behavior, 12*(1), 21-28. doi: 10.3758/BF03199809

Thompson, R. F, & Gluck, M. (1991). Brain substrates of basic associative learning. In H. Weingartner & R. Lister (Ed.), *Cognitive neuroscience* (pp. 24 – 45). New York: Oxford University Press.

Tremblay, L., & Schultz, W. (1999). Relative reward preference in primate orbitofrontal cortex. *Nature, 398* (6729), 704-708. doi: 10.1038/19525

Umeno, M. M., & Goldberg, M. E. (1997). Spatial processing in the monkey frontal eye field. I. Predictive visual responses. *Journal of neurophysiology, 78* (3), 1373-1383. doi: 10.1152/jn.1997.78.3.1373

Vadillo, M. A., Castro, L., Matute, H., & Wasserman, E. A. (2008). Short Article: Backward Blocking: The Role of Within-Compound Associations and Interference between Cues Trained Apart. *Quarterly journal of experimental psychology, 61*(2), 185-193. doi: 10.1080/17470210701557464

Wagner, A. R. (1981). SOP: A model of automatic memory processing in animal behavior. In *Information processing in animals: Memory mechanisms* (eds. N.E. Spears and R.R. Miller) pp. 5–47. Erlbaum, Hillsdale, NJ.

Widrow B., & Hoff, M. E. (1960). Adaptive switching circuits, *1960 WESCON Convention Record, Part IV* (pp. 96–104).

Williams, D. A., Overmier, J. B., & LoLordo, V. M. (1992). A reevaluation of Rescorla's early dictums about Pavlovian conditioned inhibition. *Psychological Bulletin, 111*(2), 275-290. doi: 10.1037/0033-2909.111.2.275

CHAPTER II: GOALS AND HYPOTHESES

Animal and human studies suggested that associative interference plays a role on facilitation of subsequent learning (Alcalá, Callejas-Aguilera, Lamoureux, & Rosas, 2019; Alcalá, Callejas-Aguilera, Nelson, & Rosas, 2020; Nelson, Fabiano, & Lamoureux, 2018; Potts & Shanks, 2014; Seabrooke, Hollins, Kent, Wills, & Mitchell, 2019). It seems like the experience of associative interference may produce an elevation of the attentional resources that, finally, translates into a general improvement in subsequent learning. Furthermore, this phenomenon has been described in tasks of a diverse nature such as avoidance, memory, spatial learning, or temporal conditioning tasks.

In addition, most of these studies test the facilitation of learning after associative interference in complex learning phenomena: Spatial learning (Alcalá et al., 2020), temporal conditioning (Alcalá et al., 2019), memory tasks (Potts & Shanks, 2014; Seabrooke et al., 2019), and the associative interference procedure used to generate the increase in the prediction error has involved some form of extinction, that is, it has involved eliminating previously learned information (Bouton, 1993, 1994, 1997; Rosas, Callejas-Aguilera, Ramos-Álvarez and Abad, 2006). Finally, facilitation of learning has not been found when tested in a simple cue-consequence conditioning (Alcalá, González, Aristizabal, Callejas-Aguilera, & Rosas 2018). Knowing the scope of an interference treatment on later learning, and its limits, requires the evaluation of this phenomenon with new procedures and under different conditions. Specifically, it is important to dissociate between the attention that comes from the interference treatment from the changes in attention that introduces the presentation of the cue itself (novelty). In this sense, a possible control would be to eliminate the increase in attention that the novel cue introduces just because it is novel. To do this, it would be enough to pre-expose the cue on which the learning will fall before the interference treatment takes place. This strategy can be useful to find results that help extend the effect that an interference treatment has on later learning. Furthermore, the introduction of pre-exposure, either as a latent inhibitor or a conditioned inhibitor, makes it possible to facilitate the measurement of learning improvement.

General goal and main Hypothesis

General goal: Testing whether the experience of associative interference produces a generalized improvement in subsequent predictive learning in humans. Specifically, our interest was to test whether the experience with different types of associative interference with a cue would attenuate retardation of learning about a different cue that had been pre-exposed in a human predictive learning situation.

General hypothesis: The experience of interference was expected to facilitate subsequent learning about new cue-outcome relationships that should be shown as an attenuation of the retardation effects of pre-exposing a cue in human predictive learning, regardless of the type of interference experienced.

Specific Goals and Hypotheses

Specific goal 1 (Study 1): Evaluating whether the cue pre-exposure in our task makes it a conditioned inhibitor or a latent inhibitor using the retardation (Experiment 1a) and summation tests (Experiment 1b) (Rescorla, 1969).

Specific hypothesis 1 (Study 1): Pre-exposed cue will pass retardation test (Experiment 1a) acquiring inhibitory properties. Then, preexposed cue will pass both, retardation (Experiment 1a) and summation test (Experiment 1b), so that the procedure would be useful to test the effects of interference on reversing conditioned inhibition.

Specific goal 2 (Study 1): Testing whether the interference experience generated by extinguishing a cue facilitates reversing the meaning of another cue that has been previously preexposed without the outcome.

Specific hypothesis 2 (Study 1): The experience of extinction of a cue was expected to facilitate reversal training of a previously pre-exposed target cue, regardless of the value as a latent inhibitor or as a conditioned inhibitor acquired during the pre-exposure phase, attenuating the retardation of learning with respect to a control group that had no experience of interference.

Specific goal 3 (Study 1): To eliminate possible alternative explanations based on rules learning, additional methodological controls were introduced.

Specific hypothesis 3 (Study 1): The facilitating effect of learning after associative interference will remain after the elimination of possible alternative rule-based explanations.

Specific goal 4 (Study 2): Extend the previous results using a different interference procedure. More specifically, testing whether an associative interference produced by an increase in the cue predictive value (conditioning a pre-exposed cue) would affect the subsequent learning in a different cue.

Specific hypothesis 4 (Study 2): Pairing a pre-exposed cue with the outcome was expected to produce a facilitation of learning about the reversal of the training of a different conditioned inhibitor.

Specific goal 5 (Study 2): Replicating in a single experiment the results reported in Experiment 1 and those reported in Study 1. This, in addition, will enable direct comparison of the effect based on different interference treatments.

Specific hypothesis 5 (Study 2): Improvement in subsequent learning will appear regardless of the interference treatment performed (conditioning or extinction of a previously trained cue) compared with control group. Furthermore, if the different interference treatments affect the prediction error significantly differently, the subsequent learning about a conditioned inhibitor will be facilitated by the condition where the prediction error is the greatest.

References

Alcalá, J. A., Callejas-Aguilera, J., Lamoureux, J. A., & Rosas, J. M. (2019). Discrimination reversal facilitates subsequent acquisition of temporal discriminations in rats' appetitive conditioning. *Journal of Experimental Psychology: Animal Learning and Cognition, 45*, 446-463. doi: 10.1037/xan0000216

Alcalá, J. A., Callejas-Aguilera, J., Nelson, J. B., & Rosas, J. M. (2020). Reversal training facilitates acquisition of new learning in a morris water maze. *Learning & Behavior, 48*(2), 208-220. doi: 10.3758/s13420-019-00392-7

Alcalá, J. A., González, G., Aristizabal, J. A., Callejas-Aguilera, J. E., & Rosas, J. M. (2018). Discrimination reversal facilitates contextual conditioning in rats' appetitive conditioning. *Psicológica, 39*(1), 64-87. doi: 10.2478/psicolj-2018-0004

Bouton M. E. (1993). Context, time, and memory retrieval in the interference paradigms of Pavlovian learning. *Psychological Bulletin, 114,* 80-99. doi: 10.1037//0033-2909.114.1.80

Bouton M. E. (1994). Conditioning, remembering, and forgetting. *Journal of Experimental Psychology: Animal Behavior Processes, 20*, 219-231. doi: 10.1037//0097-7403.20.3.219

Bouton, M. E. (1997). Signals for whether versus when an event will occur. In M. S. Fanselow & M. E. Bouton, *Learning, motivation, and cognition: The functional*

behaviorism of Robert C. Bolles (pp. 385–409). Washington, DC: American Psychological Association.

De la Casa, L. G., Mena, A., Ruiz-Salas, J. C., Quintero, E., & Papini, M. R. (2018). Reward devaluation disrupts latent inhibition in fear conditioning. *Learning & behavior, 46*(1), 49-59. doi: 10.3758/s13420-017-0282-1

Nelson, J. B., Fabiano, A. M., & Lamoureux, J. A. (2018). The effects of extinction-aroused attention on context conditioning. *Learning & Memory, 25*, 165-175. doi: 10.1101/lm.046201.117

Potts, R., & Shanks, D. R. (2014). The benefit of generating errors during learning. *Journal of Experimental Psychology: General, 143*(2), 644-667. doi: 10.1037/e502412013-787

Rosas, J. M., Callejas-Aguilera, J. E., Ramos-Álvarez, M. M., & Abad, M. J. (2006). Revision of retrieval theory of forgetting: What does make information context-specific? *International Journal of Psychology and Psychological therapy, 6*, 147-166. Retrieved from https://www.ijpsy.com/volumen6/num2/136/revision-of-retrieval-theory-of-forgetting-EN.pdf.

Seabrooke, T., Hollins, T. J., Kent, C., Wills, A. J., & Mitchell, C. J. (2019). Learning from failure: Errorful generation improves memory for items, not associations. *Journal of Memory and Language, 104*, 70-82. doi: 10.1016/j.jml.2018.10.001

CHAPTER III: EMPIRICAL STUDIES

Two different empirical studies were developed with the foal of evaluating the subsequent learning facilitation of a pre-exposed cue in humans using a predictive judgment task.

More specifically, in study 1, the goal was to test whether associative interference (extinction) facilitated the subsequent learning of a cue pre-exposed without outcome training regardless of whether it was a conditioned inhibitor or latent inhibitor. This measurement was made through a retardation test and alternative explanations were avoided.

Study 2 was developed with the aim of expanding the previous findings. More specifically, study 2 presented two different goals: the first one was to use an alternative procedure to generate associative interference, precisely, the conditioning of a pre-exposed cue; the second one, evaluated the different associative interference procedures (extinction and conditioning) with the goal of comparing whether different types of manipulation of the prediction error would produce a different facilitation of later learning.

Experimental Study I: Experiencing Extinction with a Non-Target Cue Facilitates Reversal of a Target Conditioned Inhibitor in Human Predictive Learning.

González, G., Alcalá, J. A., Callejas-Aguilera, J., & Rosas, J. M. (2019). Experiencing extinction with a non-target cue facilitates reversal of a target conditioned inhibitor in human predictive learning. *Behavioural Processes, 166*, 103898. http://dx.doi.org/10.1016/j.beproc.2019.103898

Abstract

Three experiments tested the effect of experiencing extinction on learning about a differential conditioned inhibitor that was trained as an excitor. A human predictive learning task was used in which participants had to evaluate the probability of different colored fertilizers (Cues) leading plants to flourish or not (Outcome). Experiment 1 found that presenting the target cue without outcome while other cues were followed by the outcome made the target cue a conditioned inhibitor, passing both, retardation (Experiment 1a) and summation (Experiment 1b) tests of conditioned inhibition. Subsequent extinction of a different cue facilitated reversing the relationship between the conditioned inhibitor and the outcome regardless of whether the situation could be solved by using simple rules (Experiment 2) or not (Experiment 3). Results are discussed in terms of attentional theories that suggests extinction produces a nonspecific increase in attention that facilitates learning.

Keywords: Associative interference; Attention; Conditioned inhibition; Extinction; Predictive judgments.

Experimental Study II: Reversing the Meaning of Non-Target Cues Facilitates Reversal of a Target Conditioned Inhibitor in Human Predictive Learning.

González, G., Alcalá, J. A., Ogállar, P. M., Rosas, J. M., & Callejas-Aguilera, J. E. (2021). Reversing the relationship between a nontarget cue and the outcome facilitates subsequent human predictive learning. *Behavioural Processes, 193*, 104529. https://doi.org/10.1016/j.beproc.2021.104529

Abstract

Two experiments were conducted to test the effect of experiencing associative interference on later learning. A predictive learning task was used in which human participants had to evaluate whether plants would grow or not (Outcome) after being watered with different fertilizers (Cues). Experiment 1 found that pairing with the outcome a cue that had been presented alone before (inhibition reversal) facilitated subsequent learning about the relationship with the outcome of a different cue that had been exposed without outcome. Experiment 2 compared the effect on subsequent learning of experiencing two types of interference, inhibition reversal and extinction (presenting the cue by itself after being paired with the outcome). Regardless of its type, the experience of associative interference with non-target cues facilitated inhibition reversal of the target cue. It is suggested that the increase in prediction error produced by the experience of associative interference leads to a general increase in attention that facilitates subsequent learning.

Keywords: Associative interference; Attention; Extinction; Human predictive learning; Inhibition.

CHAPTER IV: GENERAL DISCUSSION AND CONCLUSIONS

General discussion

In the following pages, the goals set out in the empirical studies and their relationship with the goals and hypotheses raised in this doctoral thesis will be analyzed and discussed. This analysis will continue through the theoretical framework that makes it possible, encompassing the results obtained as extension of the different situations in which prediction error has implication in subsequent learning and the relevance of the interference procedure as useful tools to manipulate prediction error.

Specifically, interference is a key factor in studying the role of context in learning. In this sense, it is on the basis of the explanation of the coding of the information as dependent on the training context, either specifically (Bouton 1993, 1994, 1997) or in a non-specific way (all the information learned after associative interference becomes dependent on the training context, e.g., Ogállar, Ramos-Álvarez, Alcalá, Moreno-Fernández & Rosas, 2017; Rosas, Callejas-Aguilera, Ramos-Álvarez & Abad, 2006, for a recent review). However, there are studies that suggest that contextual dependence produced by associative interference is a specific case of a more general phenomenon (e.g, Nelson & Lamoureux, 2015). Using a human avoidance task, Nelson and Lamoureux (2015) demonstrated that those groups in which one visual cue was extinguished, the subsequent extinction of a different cue was performed more quickly than in a control group. This study suggests, therefore, that associative interference can facilitate subsequent learning of, at least, when context is involved in learning situation. However, facilitation of learning after experience with associative interference has also been found in designs in which context is not involved and it could be a signal of a broader effect (see e.g., Alcalá, Callejas-Aguilera, Lamoreux & Rosas, 2019).

The evidence around the idea that associative interference generates an improvement in later learning is growing. There is an increasing number of studies supporting this idea in both, animals (Alcalá et al., 2019, Alcalá, Callejas-Aguilera, Nelson & Rosas, 2020; De la Casa, Mena, Ruiz-Salas, Quintero & Papini, 2018) and humans (Potts & Shanks, 2014; Seabrooke, Hollins, Kent, Wills & Mitchell, 2019). These studies have been carried out under different types of tasks such as: Spatial learning, temporal discrimination, and consummatory Successive Negative Contrast (animals' studies) and avoidance and memory tasks (humans' studies) respectively. In addition, all these previous studies have been developed in complex learning tasks (such as temporal discrimination or memory tasks). In parallel, multiple positive results in terms of facilitation of later learning and associative interference have been obtained using extinction (Bouton, 1993, 1994, 1997; Rosas et al., 2006), that is, the increase of the prediction error through the elimination of previously learned learning. This phenomenon has never been described through associative interference in the other direction, that is, increasing prediction error through conditioning of pre-exposed information. Finally, this phenomenon has not been replicated using a simple procedure in which a cue is paired with an outcome (Alcalá, González, Aristizábal, Callejas-Aguilera & Rosas, 2018).

The main goal of this doctoral thesis was to assess the generality of the learning facilitation effect using a different procedure (a human predictive learning task) and under different learning conditions (on previously pre-exposed stimuli) using different interference procedures (extinguishing a trained cue or conditioning a pre-exposed cue) in a simple conditioning cue-outcome training. With this main goal in mind, we carried out two studies using a human predictive learning task in which different fertilizers were used as cues and flower growth as the outcome. In each trial, participants gave a predictive judgment about the probability that the fertilizer (cue) produced flower growth (consequence). To do this, participants responded on a scale of 0 to 100 and received feedback immediately. As for the experimental design, both studies were similar. During the first phase, all participants received the training of three different cues followed or not by the outcome. The difference between the studies took place during the second phase of the task, where the target cue received different associative interference treatments. In both studies the subsequent learning facilitation was evaluated by testing the impact of the experience of associative interference in reversing the meaning of a pre-exposed cue through a retardation test. More specifically, in study 1, extinction of the nontarget cue was used as the associative interference procedure. In study 2, associative interference consisted of training a cue that had previously been pre-exposed. The test was identical in both studies consisting of reversing the relationship with the outcome of a cue that was pre-exposed during phase 1.

Two different goals were set in each of the studies. For study 1, the goal was to evaluate whether the associative interference produced by the extinction of a previously trained cue facilitated the subsequent training of a pre-exposed cue. The goal of study 2 was twofold: Evaluating whether the associative interference produced by the conditioning of a pre-exposed cue would facilitate subsequent learning of a different pre-exposed cue and, establishing a direct comparison between the effect of the two types of associative interference used in the two studies, extinction and reversal of conditioned inhibition.

Regarding the results, study 1 found that the pre-exposure treatment received by cues in this experimental arrangement render the pre-exposed cues as conditions inhibitors, as they passed both, retardation and summation tests (Experiments 1a and 1b). Related to the main goal of the dissertation, study 1 also found facilitation of learning in the group that received the associative interference treatment (Extinction) both, in Experiments 2 and 3. In study 2, this improvement in later learning was also obtained in the retardation test. In this case, the group that received the associative interference treatment (reversal of inhibition) showed faster learning than the control group. Likewise, in Experiment 2, the direct comparison between associative interference treatments (extinction of a previously trained cue versus conditioning of a pre-exposed cue) yielded interesting results. Both associative interference treatments showed facilitation in the testing phase with no differences between the two associative interference treatments. In other words, the type of associative interference procedure does not seem to be relevant to obtain this later learning improvement since it is similar regardless of the type of interference used (study 2, Experiment 2).

These results provide new evidence on the generality of the use of associative interference as a procedure to produce improvement of later learning. Results of previously described studies already showed a facilitation of learning using a memory task in humans (Potts & Shanks, 2014; Seabrooke et al., 2019). This improvement in learning also has been observed in avoidance tasks (Nelson & Lamoreaux, 2015). In this sense, the results described here add a new situation in which the learning facilitation is enhanced by associative interference. In parallel, the results obtained in animal learning suggest that learning facilitation after associative interference is a mechanism that transcends the animal species used, as much as the task used (Alcalá et al., 2019; 2020).

Attentional changes seem to be at the base for the explanation of these results. The associative interference treatment seems to become as an attention activator. This assumption is in line with the idea that the effect of associative interference in contextual dependency of retrieval of the information is also based on associative interference leading to attentional increases that affect specifically to context processing (e.g., Bouton, 1993, 1994, 1997; Rosas et al., 2006; Rosas & Callejas-Aguilera, 2007). In fact, it has been proposed that the increase in context processing after interference may be a side effect of the experience of associative interference producing a general increase on attention (Rosas & Nelson, 2019).

Attentional associative models have also included the error term as the trigger for attention to fluctuate throughout learning (Mackintosh, 1975; Pearce & Hall, 1980). The error term is high when learning begins and is gradually decreasing as learning progresses. Learning models such as Pearce and Hall (1980) sustain that the decrease in the error term is accompanied by a decrease in the attention the stimuli receive. In a more recent approach, Le Pelley, Mitchell, Beesley, George & Wills, (2016) proposed the existence of two different attentional mechanisms that enter play under different circumstances, and at different moments in training: Attentional exploitation mechanism an attentional exploration mechanism. The second, attentional exploration mechanism, is used when the relationship between cues and consequences is uncertain. Accordingly, some reports suggested that organisms spent more time observing those cues that are bad predictors than to the good ones (see Koenig, Kadel, Uengoer, Schubö & Lachnit, 2017; Luque, Vadillo, Le Pelley & Beesley, 2017). The break in the established predictions involved in associative interference produces a sudden increase in the error term that is assumed to be responsible of the facilitation of learning that is reported in the studies presented in this doctoral thesis. The results presented here could be an example of the attentional exploration mechanism where associative interference plays an activating role of the attentional exploration mechanism (Beesley, Nguyen, Pearson & Le Pelley, 2015).

Additionally, the results reported here may be considered as an indirect support for the interpretation given by Alcalá et al. (2018) for the failure to obtain an improvement in learning after the experience of interference. In Alcalá et al. (2018) study, the improvement of learning does not affect to a new CS-US relationship, but to context processing. In the studies conducted in this dissertation improvement of learning does affect new learning about a Cue-Outcome relationship, but about a familiar cue. It is reasonable to assume that the novelty of the CS-US relationship may have activated attention to the CS and to the CS-US relationship so that no further improvements can be detected (Schmajuk, Lam & Gray, 1996). The use of consistently preexposed cues in these experimental series would have avoided reaching the asymptote of attention to the target cues at the time of testing.

In sum, interference seems to facilitate subsequent acquisition of learning, and this has been found regardless of the type of interference used (extinction or preexposure and conditioning). However, we need more studies to amplify the range of tasks, procedures, and types of interference to describe the generalities and the boundaries in which this learning phenomenon takes place. Currently, the number of studies is quite limited, and future research is needed to determine the scope and its relevance as well as the use of more precise attention measures that are independent from behavioral measures, such as gaze tracking, or orientation responding.

Conclusions

1. Associative interference treatments (extinction a previously trained cue or conditioning a pre-exposed cue) facilitated, in humans, subsequent learning about a Cue-Outcome relationship that had been previously pre-exposed in a predictive judgments task. This learning improvement replicates the results of previous studies with animals (Alcalá et al., 2019, 2020; De la Casa et al., 2018) and, also, using memory tasks (Potts & Shanks, 2014; Seabrooke et al., 2019) and an avoidance task with humans (Nelson & Lamoureux, 2015).

2. The learning improvement found in these studies is in agreement with the idea of the increase in the prediction error caused by the experience of associative interference leading to a general increase in attention that facilitates learning (Rosas & Nelson, 2019; see also LePelley et al., 2016).
3. The two associative interference procedures used (extinction of a previously trained cue or conditioning of a previously pre-exposed cue) led to the same facilitation of learning in the retardation test used. Although this result may suggest that different types of interference produce equivalent increases in attention, it is also possible that the measure used in the task is not sensitive enough to detect small differences across interference treatments. Thus, future research should complement behavioral measures with other independent measures of attention such as the ones provided by eye-tracking devices.

4. Improvement of learning after associative interference is consistent with the idea that sudden increases in the prediction error activate the attentional exploration mechanism facilitating learning because of a general increase in attention to the situation (Beesley et al., 2015; LePelley et al., 2016).

References

Alcalá, J. A., Callejas-Aguilera, J., Lamoureux, J. A., & Rosas, J. M. (2019). Discrimination reversal facilitates subsequent acquisition of temporal discriminations in rats' appetitive conditioning. *Journal of Experimental Psychology: Animal Learning and Cognition, 45*, 446-463. doi: 10.1037/xan0000216

Alcalá, J. A., Callejas-Aguilera, J., Nelson, J. B., & Rosas, J. M. (2020). Reversal training facilitates acquisition of new learning in a morris water maze. *Learning & Behavior, 48*(2), 208-220. doi: 10.3758/s13420-019-00392-7

Alcalá, J. A., González, G., Aristizabal, J. A., Callejas-Aguilera, J. E., & Rosas, J. M. (2018). Discrimination reversal facilitates contextual conditioning in rats' appetitive conditioning. *Psicológica, 39*(1), 64-87. doi: 10.2478/psicolj-2018-0004

Beesley, T., Nguyen, K. P., Pearson, D., & Le Pelley, M.E. (2015). Uncertainty and predictiveness determine attention to cues during human associative learning. *Quarterly Journal of Experimental Psychology (2006), 68*, 2175-2199. doi: 10.1080/17470218.2015.1009919

Bouton M. E. (1993). Context, time, and memory retrieval in the interference paradigms of Pavlovian learning. *Psychological Bulletin, 114,* 80-99. doi: 10.1037//0033-2909.114.1.80

Bouton M. E. (1994). Conditioning, remembering, and forgetting. *Journal of Experimental Psychology: Animal Behavior Processes, 20*, 219-231. doi: 10.1037//0097-7403.20.3.219

Bouton, M. E. (1997). Signals for whether versus when an event will occur. In M. S. Fanselow & M. E. Bouton, *Learning, motivation, and cognition: The functional*

behaviorism of Robert C. Bolles (pp. 385–409). Washington, DC: American Psychological Association.

De la Casa, L. G., Mena, A., Ruiz-Salas, J. C., Quintero, E., & Papini, M. R. (2018). Reward devaluation disrupts latent inhibition in fear conditioning. Learning & behavior, 46(1), 49-59. doi: 10.3758/s13420-017-0282-1

Koenig, S., Kadel, H., Uengoer, M., Schubö, A., & Lachnit, H. (2017). Reward draws the eye, uncertainty holds the eye: Associative learning modulates distractor interference in visual search. *Frontiers in behavioral neuroscience, 11*, 128. doi: 10.3389/fnbeh.2017.00128

Le Pelley, M. E., Mitchell, C. J., Beesley, T., George, D. N., & Wills, A. J. (2016). Attention and associative learning in humans: An integrative review. *Psychological Bulletin, 142*(10), 1111 1140. doi: 10.1037/bul0000064

Luque, D., Vadillo, M. A., Le Pelley, M. E., & Beesley, T. (2017). Prediction and uncertainty in associative learning: Examining controlled and automatic components of learned attentional biases. *The Quarterly Journal of Experimental Psychology, 70*(8), 1485-1503. doi: 10.1080/17470218.2016.1188407

Mackintosh, N. J. (1975). A theory of attention: variations in the associability of stimuli with reinforcement. *Psychological review, 82*(4), 276. doi: 10.1037/h0076778

Nelson, J. B., & Lamoureux, J. A. (2015). Contextual control of conditioning is not affected by extinction in a behavioral task with humans. *Learning & Behavior, 43*(2), 163-178. doi: 10.3758/s13420-015-0170-5

Ogállar, P. M., Ramos-Álvarez, M. M., Alcalá, J. A., Moreno-Fernández, M. M., & Rosas, J. M. (2017). Attentional Perspectives on Context-dependence of Information Retrieval. *International Journal of Psychology and Psychological Therapy, 17*, 121-136. Retrieved from https://www.ijpsy.com/volumen17/num1/462/attentional-perspectives-on-context-dependence-EN.pdf

Pearce, J. M., & Hall, G. (1980). A model for Pavlovian learning: variations in the effectiveness of conditioned but not of unconditioned stimuli. *Psychological review, 87*(6), 532. doi: 10.1037/0033-295X.87.6.532

Potts, R., & Shanks, D. R. (2014). The benefit of generating errors during learning. *Journal of Experimental Psychology: General, 143*(2), 644-667. doi: 10.1037/e502412013-787

Rosas, J. M., & Callejas-Aguilera, J. E. (2007). Acquisition of a conditioned taste aversion becomes context dependent when it is learned after extinction. The Quarterly Journal of Experimental Psychology, 60, 9-15. https://doi.org/10.1080/17470210600971519

Rosas, J. M., Callejas-Aguilera, J. E., Ramos-Álvarez, M. M., & Abad, M. J. (2006). Revision of retrieval theory of forgetting: What does make information context-specific? International Journal of Psychology and Psychological therapy, 6, 147-166. Retrieved from https://www.ijpsy.com/volumen6/num2/136/revision-of-retrieval-theory-of-forgetting-EN.pdf.

Rosas, J. M., & Nelson, J. B. (2019). Context dependency as a function of prediction error-based attention. *Psicológica, 40*(2), 34-45. doi: 10.2478/psicolj-2019-0003

Schmajuk, N. A., Lam, Y., & Gray, J. A. (1996). Latent inhibition: A neural network approach. *Journal of Experimental Psychology: Animal Behavior Processes, 22,* 321–349. doi: 10.1037//0097-7403.22.3.321

Seabrooke, T., Hollins, T. J., Kent, C., Wills, A. J., & Mitchell, C. J. (2019). Learning from failure: Errorful generation improves memory for items, not associations. *Journal of Memory and Language, 104*, 70-82. doi: 10.1016/j.jml.2018.10.001

CHAPTER V: RESUMEN EXTENDIDO

La supervivencia de cualquier organismo requiere de su adaptación al medio en el que vive. En este sentido, es vital aprender la regularidad con la que se suceden algunos eventos en el ambiente, en especial los relacionados con la conducta sexual, la reproducción, el miedo, la agresividad, o la alimentación, entre otros (Domjan, 2005). Según lo anterior podríamos definir el aprendizaje como la capacidad para organizar el ambiente mediante la detección y predicción de regularidades en diferentes situaciones. Ahora bien, en ocasiones la regularidad con la que transcurren determinados acontecimientos cambia, bien porque se rompe la sucesión en la que ocurrían, o porque se modifica su significado al anunciar nuevos acontecimientos. En este sentido, la capacidad de adaptación y supervivencia, además de la detección y predicción de regularidades, requiere de un mecanismo para la corrección del *error de predicción,* entendido este como la discrepancia entre lo que el organismo espera y lo que finalmente ocurre en el ambiente.

Desde los modelos clásicos del aprendizaje asociativo (Mackintosh, 1975; Pearce y Hall, 1980; Rescorla y Wagner, 1972; Sutton y Barto, 1981) se ha considerado el error de predicción como el mecanismo que guía al aprendizaje en un intento de reducir la discrepancia entre lo que el organismo espera y lo ocurre; en definitiva el mecanismo que permite la adaptación con éxito a un medio en el que algunas de sus relaciones cambian con el tiempo.

El error de predicción varía a medida que transcurre el aprendizaje; es máximo al comienzo del entrenamiento, disminuyendo su valor conforme aumenta la certeza sobre lo que va a ocurrir. Las teorías asociativas del aprendizaje se diferencian entre sí en la explicación de cómo se reduce el error de predicción, en parte a partir del papel que atribuyen a los diferentes elementos del aprendizaje. Algunas teorías (e.g., Rescorla y Wagner, 1972; Wagner, 1981) atribuyen su reducción a la relevancia del estímulo incondicionado (EI), en concreto a que éste resulte más o menos esperado. Otros modelos (e.g., Mackintosh, 1975; Pearce y Hall, 1980), ponen el énfasis en el estímulo condicionado (EC), en concreto en la capacidad que tiene para predecir la aparición del EI, lo que determinará la atención que recibe el EC. Desde una perspectiva distinta Larrauri y Schmajuk (2008) sugieren una relación directa entre la atención y la novedad, si bien proponen un incremento generalizado de la atención, no sólo al estímulo novedoso, que será mayor conforme también lo sea la novedad que introduce el estímulo. Por último, Le Pelley, Mitchell, Beesley, George y Wills (2016), proponen un modelo híbrido capaz de reconciliar las predicciones contrapuestas entre Mackintosh (1975) y Pearce y Hall (1980), al postular un doble mecanismo atencional: uno de explotación, activo en ambientes estables con el objetivo de maximizar las ganancias, por el que se presta mayor atención a los EC que son buenos predictores del EI, compatible, por tanto, con la propuesta de Mackintosh. El otro mecanismo, exploratorio, se activaría ante cambios producidos en el ambiente, promoviendo la atención a cualquier estímulo, con el objetivo de encontrar buenos predictores de la consecuencia, y por tanto compatible con la propuesta de Pearce y Hall.

El término error de predicción también ha recibido apoyo empírico desde el ámbito de las neurociencias. En este sentido Widrow y Hoff (1960) propusieron la existencia de correlatos biológicos para el error de predicción que, con posterioridad, ha dado lugar al desarrollo de estudios desde diferentes estrategias de investigación (neurotransmisores, áreas y sistemas cerebrales). Independientemente de la función específica de neuronas, áreas, y sistemas, todos parecen actuar como un sistema estructurado que utiliza el error de predicción como elemento organizador de la función neuronal. De la misma forma que el error de predicción sirve como activador de la atención a nivel conductual, el error de predicción parece jugar un papel regulador en el procesamiento neuronal (Schultz y Dickinson, 2000).

Desde el punto de vista de la conducta, el estudio empírico del error de predicción se ha abordado a partir de procedimientos de interferencia asociativa. Los procedimientos más utilizados han sido la extinción, el contracondicionamiento y la inversión de la discriminación. Todos ellos se caracterizan por la modificación de la asociación evento-consecuencia previamente establecida, lo que genera por tanto una discrepancia entre lo que el organismo espera y lo que recibe. El empleo de estos procedimientos ha estimulado durante las últimas décadas el estudio del aprendizaje bajo situaciones de incertidumbre. Una de las líneas de investigación más fructíferas es la iniciada por Bouton (1993, 1994, 1997), quien considera que el aprendizaje inicial no se borra como consecuencia de un tratamiento de interferencia, sino que en su lugar ambos aprendizajes (interferido e interferente) se almacenan en la memoria, y en el que el contexto desempeña un papel clave para resolver la situación de incertidumbre. En concreto, Bouton (1993) sugiere que la interferencia asociativa convierte a la información inhibitoria, o aprendida en segundo lugar, en dependiente del contexto (no obstante, véase Nelson, 2002). Posteriormente, Rosas, Callejas-Aguilera, Ramos-Álvarez y Abad (2006), extienden esta idea, y muestran que no solo la información interferente se convierte en dependiente del contexto, sino que, una vez ocurre la interferencia, toda la información aprendida se codifica como dependiente del contexto, independientemente de que se trate de información aprendida en primer o en segundo lugar, o de su carácter excitatorio o inhibitorio.

Más aún, estudios recientes sugieren que el incremento en la atención que genera un procedimiento de interferencia asociativa podría suponer una mejora en el aprendizaje que ocurre de manera posterior. En este sentido, el papel que desempeñan los contextos en fenómenos como el de renovación, sería tan solo la manifestación particular de un efecto más general. En concreto, la idea que subyace a esta hipótesis es que el aumento en el error de predicción que genera un tratamiento de interferencia produce un aumento general de la atención, de modo que el organismo comenzaría a atender a toda la información presente en la situación de aprendizaje (e.g., Alcalá, Callejas-Aguilera, Lamoreux y Rosas, 2019; Alcalá, Callejas-Aguilera., Nelson y Rosas, 2020; Nelson, Fabiano y Lamoureux, 2018; Shanab y Cotton, 1970). De este modo, la interferencia asociativa, el error de predicción y la atención estarían relacionados. No obstante, este efecto facilitador no ha resultado infalible, mostrándose esquivo cuando se ha puesto a prueba en el aprendizaje de una adquisición simple entre un Estímulo Condicionado (EC) y un Estímulo Incondicionado (EI) (Alcalá, González, Aristizábal, Callejas-Aguilera y Rosas, 2018).

La presente tesis doctoral trata de continuar el análisis de las implicaciones y límites de la interferencia asociativa como procedimiento facilitador de un aprendizaje posterior, lo que permitirá ampliar el conocimiento acerca de los beneficios del tratamiento de interferencia sobre el nuevo aprendizaje. Asimismo, evalúa el empleo de nuevos procedimientos de interferencia asociativa que permitirán el análisis de sus consecuencias, y de las condiciones en las que se observa este fenómeno.

Para la consecución de este objetivo, la tesis doctoral se ha estructurado en cuatro capítulos: el Capítulo 1 consiste en una revisión de los estudios más relevantes en cuanto al error de predicción, interferencia asociativa y atención, y sus implicaciones sobre el efecto facilitador del aprendizaje. El Capítulo 2 incluye los objetivos, tanto generales como específicos, así como las hipótesis planteadas en cada uno de los estudios experimentales que componen esta tesis doctoral. El Capítulo 3 contiene dos estudios experimentales que exploran la implicación de la interferencia asociativa sobre el aprendizaje de una clave pre-expuesta en una tarea de aprendizaje predictivo humano.

Finalmente, el Capítulo 4 concluye con una discusión general de los resultados obtenidos y de sus implicaciones teóricas, así como una serie de conclusiones principales que se derivan del presente trabajo de tesis doctoral.

5.2. Resumen de los estudios empíricos.

Estudio 1: La extinción de una clave no objetivo facilita la inversión de un inhibidor condicionado en una tarea de aprendizaje predictivo con humanos

Objetivo específico 1: Evaluar si la clave pre-expuesta, en la tarea utilizada en este estudio, adquiere propiedades como inhibidor latente (Experimento 1a), o bien como inhibidor condicionado (Experimento 1b).

Hipótesis específica 1: Si la clave pre-expuesta adquiere propiedades como inhibidor latente, entonces se observará un retraso en la asociación posterior de esa clave con la consecuencia, en comparación a la asociación entre una clave nueva con la consecuencia (Experimento 1a). Si la clave pre-expuesta adquiere propiedades como inhibidor condicionado, entonces se observará una disminución en la respuesta cuando la clave pre-expuesta se presenta junto a otra que previamente adquirió propiedades excitatorias.

Objetivo específico 2: Evaluar si la experiencia de interferencia generada por la extinción de una clave facilita la inversión del significado de otra clave previamente pre-expuesta.

Hipótesis especifica 2: El tratamiento de extinción sobre una clave facilitará el aprendizaje posterior de otra clave que fue previamente pre-expuesta, en comparación a un grupo de control que no recibe el tratamiento de extinción.

Objetivo 3: Replicar los resultados encontrados en el Experimento 2, al mismo tiempo que se excluye una explicación alternativa fundada en la utilización de una estrategia basada en reglas.

Hipótesis 3: El efecto facilitador en el aprendizaje encontrado en el Experimento 2 se mantendrá con un diseño que elimina la posibilidad de resolver la tarea a partir del empleo de reglas.

Con estos objetivos se diseñó una serie experimental en aprendizaje predictivo humano, en la que diferentes productos de jardinería fueron utilizados como claves que predecían, o no, el crecimiento de unas plantas (consecuencia). Como respuesta se registraron los juicios predictivos emitidos por los participantes de la relación entre cada una de las claves con la consecuencia. La serie experimental se compuso de tres experimentos. Los Experimentos 1a y 1b se realizaron para caracterizar las propiedades adquiridas por la clave objetivo (Ci) como consecuencia de su exposición previa en ausencia de la consecuencia. En el Experimento 1a se utilizó un diseño intrasujeto en el que se comparó la velocidad con la que Ci se asoció con la consecuencia en una fase posterior, en comparación con una clave nueva (N) que no había recibido entrenamiento previo alguno. En el Experimento 1b, también intrasujeto, se evaluó la capacidad de la clave pre-expuesta Ci para reducir la respuesta de otra clave (P) previamente asociada con la consecuencia. Los resultados mostraron un retraso en el aprendizaje de Ci comparado con N en el Experimento 1a, y una disminución de la respuesta a P cuando fue presentada junto a Ci (CiP), lo que sugiere que la clave pre-expuesta adquirió propiedades como inhibidor condicionado. El Experimento 2 evaluó si la interferencia asociativa facilita el aprendizaje posterior de la clave pre-expuesta Ci. Para ello, una clave fue asociada con la consecuencia (E+) y posteriormente extinguida (E-) en dos fases consecutivas. Finalmente, durante la fase de prueba, se evaluó si el tratamiento de extinción facilitó el condicionamiento de la clave pre-expuesta Ci, en comparación a un grupo de control que no recibió tratamiento de interferencia. Los resultados mostraron una facilitación en el aprendizaje de Ci en el grupo de extinción en comparación a un grupo de control que no recibió tratamiento de interferencia. Por último, el Experimento 3 se diseñó con el objetivo de excluir una explicación alternativa de los resultados anteriores basada en el uso de reglas. Con este objetivo, se incorporaron nuevas claves entrenadas de manera consistente con la consecuencia (F1- y F2+) durante todo el entrenamiento (durante la adquisición y la extinción de E), eliminando así una posible regla según la cual durante la fase de extinción se habría modificado el significado de todas las claves que previamente estuvieron asociadas con la consecuencia. Los resultados en el Experimento 3 replicaron los informados en el Experimento 2, lo que sugiere que la facilitación en el aprendizaje de una adquisición simple entre la clave pre-expuesta y la consecuencia fue debida al tratamiento de interferencia que recayó sobre otra clave, superando así los problemas informados por Alcalá et al. (2018), y extendiendo a una tarea de aprendizaje predictivo humanos los resultados informados previamente en la literatura (e.g., Alcalá et al., 2019, 2020, Nelson et al., 2018; Potts y Shanks, 2014; Seabrooke, Hollins, Kent, Wills, & Mitchell, 2019; Shanab y Cotton, 1970).

Estudio 2: Invertir el significado de las señales no objetivo facilita la inversión de un inhibidor condicionado objetivo en una tarea de aprendizaje predictivo humano

Objetivo 1: Extender los resultados informados en el Estudio 1 utilizando un procedimiento de interferencia diferente. Más específicamente, probar si la interferencia asociativa generada por el aumento en el valor predictivo de una clave (condicionando excitatoriamente una clave pre-expuesta) facilita el aprendizaje posterior de otra clave diferente que también fue pre-expuesta.

Hipótesis 1: El empleo de este nuevo tratamiento de interferencia facilitará el aprendizaje posterior de otra clave que fue previamente pre-expuesta, en comparación a un grupo de control que no recibe tratamiento de interferencia.

Objetivo 2: Replicar en un único experimento los resultados informados en el Experimento 1 de este estudio, así como en los Experimentos 2 y 3 del Estudio 1. Al mismo tiempo, comparar de forma directa si los distintos tratamientos de interferencia analizados afectan de manera distinta la facilitación del aprendizaje.

Hipótesis 2: La mejora en el aprendizaje posterior aparecerá independientemente del tratamiento de interferencia empleado (condicionamiento o extinción de una clave previamente entrenada) en comparación con el grupo control. Además, si los diferentes tratamientos de interferencia afectan al error de predicción de forma significativamente diferente, el aprendizaje posterior sobre un inhibidor condicionado se verá facilitado por la condición en la que el error de predicción sea subjetivamente mayor.

Con estos objetivos, se realizó una serie experimental con dos experimentos utilizando la misma tarea de juicios predictivos informada en el Estudio 1. El diseño del Experimento 1 fue similar al utilizado en el Experimento 3 del Estudio 1, con la salvedad que la interferencia fue generada por el condicionamiento de una clave previamente pre-expuesta (N-/N+). Al igual que ocurriera en el Estudio 1, los resultados mostraron facilitación en el aprendizaje de una adquisición simple entre otra clave pre-expuesta y la consecuencia, en comparación a un grupo de control sin tratamiento de interferencia. En el Experimento 2 se añadió un grupo más a los dos anteriores, en el que se introdujo la extinción simple como tratamiento de interferencia. Los resultados nuevamente mostraron una facilitación en el aprendizaje en ambos los grupos de interferencia en comparación al grupo de control, si bien ninguno de los tratamientos de interferencia se mostró mejor que el otro en cuanto a la mejora que producían sobre el aprendizaje posterior. Estos resultados replican los obtenidos previamente y, además, extienden el uso del condicionamiento de una clave pre-expuesta como un procedimiento válido para generar interferencia asociativa, al menos, con participantes humanos en una tarea de juicios predictivos.

5.3. Conclusiones

1. Los tratamientos de interferencia asociativa (extinción de una señal previamente entrenada o condicionamiento de una señal pre-expuesta) facilitaron el condicionamiento posterior de una clave que había sido pre-expuesta, en una tarea de aprendizaje predictivo humano. Estos resultados replican los informados previamente en estudios con animales (Alcalá et al., 2019, 2020; De la Casa, Mena, Ruiz-Salas, Quintero, & Papini, 2018), y en humanos en tareas de memoria (Potts & Shanks, 2014; Seabrooke et al., 2019), y de evitación (Nelson & Lamoureux, 2015).
2. La mejora del aprendizaje encontrada en estos estudios va en consonancia con la idea de que el aumento del error de predicción provocado por la experiencia de interferencia asociativa conduce a un incremento general de la atención facilitando, en última estancia, el aprendizaje (Rosas & Nelson, 2019; ver también LePelley et al., 2016).
3. Los dos procedimientos de interferencia asociativa utilizados (extinción de una señal previamente entrenada o condicionamiento de una señal previamente pre-expuesta) condujeron a una facilitación similar en el aprendizaje. Aunque este resultado puede sugerir que diferentes tipos de interferencia generan aumentos

equivalentes en la atención, también es posible que la medida utilizada en la tarea no fuera lo suficientemente sensible para detectar pequeñas diferencias entre los tratamientos de interferencia. Por lo tanto, la investigación futura debería complementar las medidas conductuales con otras medidas independientes de la atención, como las proporcionadas por los dispositivos de seguimiento de la mirada (*eyetracker*).

4. La mejora del aprendizaje después de la interferencia asociativa es consistente con la idea de que un cambio repentino en el error de predicción activa el mecanismo de exploración atencional, facilitando el aprendizaje como consecuencia de un incremento generalizado de la atención a la situación de aprendizaje (Beesley, Nguyen, Pearson, & Le Pelley, 2015; LePelley et al., 2016).

5.4. Referencias

Alcalá, J. A., Callejas-Aguilera, J. E., Lamoureux, J. A., & Rosas, J. M. (2019). Discrimination reversal facilitates subsequent acquisition of temporal discriminations in rats' appetitive conditioning. *Journal of Experimental Psychology: Animal Learning and Cognition, 45*, 446-463. doi: 10.1037/xan0000216

Alcalá, J. A., Callejas-Aguilera, J. E., Nelson, J. B., & Rosas, J. M. (2020). Reversal training facilitates acquisition of new learning in a morris water maze. *Learning & Behavior, 48*(2), 208-220. doi: 10.3758/s13420-019-00392-7

Alcalá, J. A., González, G., Aristizabal, J. A., Callejas-Aguilera, J. E., & Rosas, J. M. (2018). Discrimination reversal facilitates contextual conditioning in rats' appetitive conditioning. *Psicológica, 39*(1), 64-87. doi: 10.2478/psicolj-2018-0004

Beesley, T., Nguyen, K. P., Pearson, D., & Le Pelley, M. E. (2015). Uncertainty and predictiveness determine attention to cues during human associative learning. *Quarterly Journal of Experimental Psychology, 68*(11), 2175-2199. doi: 10.1080/17470218.2015.1009919

Bouton M. E. (1993). Context, time, and memory retrieval in the interference paradigms of Pavlovian learning. *Psychological Bulletin, 114,* 80-99. doi: 10.1037//0033-2909.114.1.80

Bouton M. E. (1994). Conditioning, remembering, and forgetting. *Journal of Experimental Psychology: Animal Behavior Processes, 20*, 219-231. doi: 10.1037//0097-7403.20.3.219

Bouton, M. E. (1997). Signals for whether versus when an event will occur. In M. S. Fanselow & M. E. Bouton, *Learning, motivation, and cognition: The functional*

behaviorism of Robert C. Bolles (pp. 385–409). Washington, DC: American Psychological Association.

De la Casa, L. G., Mena, A., Ruiz-Salas, J. C., Quintero, E., & Papini, M. R. (2018). Reward devaluation disrupts latent inhibition in fear conditioning. *Learning & behavior, 46*(1), 49-59. doi: 10.3758/s13420-017-0282-1

Domjan, M. (2005). Pavlovian conditioning: A functional perspective. *Annual Review of Psychology, 56*, 179-206. doi: 10.1146/annurev.psych.55.090902.141409

Larrauri, J. A., & Schmajuk, N. A. (2008). Attentional, associative, and configural mechanisms in extinction. *Psychological review, 115*(3), 640-676. doi: 10.1037/0033-295x.115.3.640

Le Pelley, M. E., Mitchell, C. J., Beesley, T., George, D. N., & Wills, A. J. (2016). Attention and associative learning in humans: an integrative review. *Psychological Bulletin, 142(10)*, 1111-1140. doi: 10.1037/bul0000064

Mackintosh, N. J. (1975). A theory of attention: variations in the associability of stimuli with reinforcement. *Psychological review, 82*(4), 276-298. doi: 10.1037/h0076778

Nelson, J. B. (2002). Context specificity of excitation and inhibition in ambiguous stimuli. *Learning and Motivation*, *33*(2), 284-310. doi: 10.1006/lmot.2001.1112

Nelson, J. B., Fabiano, A. M., & Lamoureux, J. A. (2018). The effects of extinction-aroused attention on context conditioning. *Learning & Memory, 25*, 165-175. doi: 10.1101/lm.046201.117

Nelson, J. B., & Lamoureux, J. A. (2015). Contextual control of conditioning is not affected by extinction in a behavioral task with humans. *Learning & Behavior, 43*(2), 163-178. doi: 10.3758/s13420-015-0170-5

Pearce, J. M., & Hall, G. (1980). A model for Pavlovian learning: variations in the effectiveness of conditioned but not of unconditioned stimuli. *Psychological review, 87*(6), 532. doi: 10.1037/0033-295X.87.6.532

Potts, R., & Shanks, D. R. (2014). The benefit of generating errors during learning. *Journal of Experimental Psychology: General, 143*(2), 644-667. doi: 10.1037/e502412013-787

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 research and theory.* New York: Appleton-Century-Crofts.

Rosas, J. M., Callejas-Aguilera, J. E., Ramos-Álvarez, M. M., & Abad, M. J. (2006). Revision of retrieval theory of forgetting: What does make information context-specific? *International Journal of Psychology and Psychological therapy, 6*, 147-166. Retrieved from https://www.ijpsy.com/volumen6/num2/136/revision-of-retrieval-theory-of-forgetting-EN.pdf.

Rosas, J. M., & Nelson, J. B. (2019). Context dependency as a function of prediction error-based attention. *Psicológica, 40*(2), 34-45. doi: 10.2478/psicolj-2019-0003

Schultz, W., & Dickinson, A. (2000). Neuronal coding of prediction errors. *Annual review of neuroscience, 23*(1), 473-500. doi: 10.1146/annurev.neuro.23.1.473

Seabrooke, T., Hollins, T. J., Kent, C., Wills, A. J., & Mitchell, C. J. (2019). Learning from failure: Errorful generation improves memory for items, not associations. *Journal of Memory and Language, 104*, 70-82. doi: 10.1016/j.jml.2018.10.001

Shanab, M. E., & Cotton, J. W. (1970). Effects of runway training on behavior in the T-maze. *Psychonomic Science, 19*(3), 129-130. doi: 10.3758/bf03335510

Sutton, R. S., & Barto, A. G. (1981). Toward a modern theory of adaptive networks: expectation and prediction. *Psychological review, 88*(2), 135-170. doi: 10.1037/0033-295X.88.2.135

Wagner, A. R. (1981). SOP: A model of automatic memory processing in animal behavior. In *Information processing in animals: Memory mechanisms* (eds. N.E. Spears and R.R. Miller) pp. 5–47. Erlbaum, Hillsdale, NJ.

Widrow B., & Hoff, M. E. (1960). Adaptive switching circuits, *1960 WESCON Convention Record, Part IV* (pp. 96–104).