

## Why is the Rescorla-Wagner model so influential?

Fabian A. Soto<sup>a,\*</sup>, Edgar H. Vogel<sup>b</sup>, Yerco E. Uribe-Bahamonde<sup>c</sup>, Omar D. Perez<sup>d,e</sup>

<sup>a</sup> Florida International University, Florida, USA

<sup>b</sup> Research Center on Cognitive Sciences and Applied Psychology Center, Faculty of Psychology, University of Talca, Chile

<sup>c</sup> University of Talca, Chile

<sup>d</sup> Department of Industrial Engineering, University of Chile

<sup>e</sup> Instituto Sistemas Complejos de Ingeniería, Chile

### ARTICLE INFO

#### Keywords:

Rescorla-Wagner  
Prediction error  
Learning rules  
Computational  
Pavlovian conditioning  
Reinforcement  
Supervised learning  
Learning algorithm

### ABSTRACT

The influence of the Rescorla-Wagner model cannot be overestimated, despite that (1) the model does not differ much computationally from its predecessors and competitors, and (2) its shortcomings are well-known in the learning community. Here we discuss the reasons behind its widespread influence in the cognitive and neural sciences, and argue that it is the constant search for general-process theories by learning scholars which eventually produced a model whose application spans many different areas of research to this day. We focus on the theoretical and empirical background of the model, the theoretical connections that it has with later developments across Marr's levels of analysis, as well as the broad variety of research that it has guided and inspired.

### 1. Introduction

Over half a century ago, Robert Rescorla and Allan Wagner from Yale University published two book chapters in which they offered a formal theory of Pavlovian conditioning that came to be known as the Rescorla-Wagner (RW) model (Rescorla & Wagner, 1972). These ideas were first presented in two specialized conferences at McMaster University, USA, in 1969 on "classical conditioning", and at the University of Sussex, England, in 1971 on "learning and inhibition". The conferences were attended by a good part of the cream of the so-called "learning theorists"; among them, Abe Black, Donald Blough, Leon Kamin, Elliot Hearst, Robert Boakes, Herb Jenkins, Nick Mackintosh, Shepard Siegel, Werner Honing, and William Prokasi.

Learning theory specialists reacted almost immediately to the theory. For example, the core principles and nomenclature of the RW model were used to propose a few very influential alternative theories (e.g., Blough, 1975; Frey & Sears, 1978; Mackintosh, 1975) and several studies soon showed that some of the predictions of the RW model were accurate (e.g., Gillan & Domjan, 1977; Levitan, 1975; Mackintosh, 1976; Saavedra, 1975; Wasserman, 1974), while others were incorrect (e.g., A. G. Baker, 1974; Clarke et al., 1979; Scavio & Gormezano, 1974; Zimmer-Hart & Rescorla, 1974). All of this happened in the 1970s, when learning theory had a second boost after its initial flourishing during the reigns

of Hull, Tolman, and Guthrie. Many laboratories conducted theoretically-oriented programs of research instead of isolated studies, and ties between theory and data became a "must" among members of this scientific community. Moreover, the first independent volume of *The Journal of Experimental Psychology: Animal Behavior Processes* was published in 1975. Allan Wagner, in its first editorial, stated that "more emphasis should be placed on integrative rather than piecemeal reporting of research" and that "We will be more insistent on this point, to give preference to integrated sets of experiments and substantial blocks of research". It appears that the RW model was proposed in the right place at the right time but, of course, this is only part of the story.

Over the coming years, a progressive consensus was built on the importance of the model. By the end of the 1990s, for instance, reviewers proclaimed that the RW model meant "the inauguration of an important era in conditioning research" (Siegel & Allan, 1996) and that it was "the primary export of traditional learning theory to other areas of psychology" (Miller et al., 1995). A preliminary bibliometric analysis shows the extent of this export to other areas of not only psychology, but all cognitive sciences. A search in the Web of Science scientific database shows that, between 1975 and 2022, the chapter by Rescorla and Wagner (1972) was cited by 4,500 articles, positioning it among the 10 most cited papers in experimental psychology. To provide a reference, consider that, in the same period, the very influential papers on

\* Corresponding author.

E-mail address: [fasoto@fiu.edu](mailto:fasoto@fiu.edu) (F.A. Soto).

<https://doi.org/10.1016/j.nlm.2023.107794>

Received 25 January 2023; Received in revised form 30 May 2023; Accepted 26 June 2023

Available online 19 July 2023

1074-7427/© 2023 Elsevier Inc. All rights reserved.

habituation published by [Thompson and Spencer \(1966\)](#) and by [Groves and Thompson \(1970\)](#) were cited 1,750 and 1,470 times, respectively.

[Figs. 1 and 2](#) present some data that emerge from the bibliometric analysis of the [Rescorla and Wagner \(1972\)](#) chapter. [Fig. 1](#) shows that the number of papers citing the RW model has grown proportionately to the growth of the total number of publications in experimental psychology, suggesting that its relative influence has been stable over the years. It is noticeable that most of the citations come from the aforementioned *JEP:ABP* ( $n = 239$ ), but there are also many citations from general-scope journals such as *Psychological Review* ( $n = 71$ ), *Psychonomic Bulletin and Review* ( $n = 43$ ) and *Scientific reports* ( $n = 43$ ).

[Fig. 2](#) presents the results of a bibliometric network analysis conducted with the 5,674 different keywords within the articles citing the chapter by [Rescorla and Wagner \(1972\)](#). The predominant topics are, of course, “learning”, “associative learning”, “conditioning” and “extinction”, to all extents the major targets of the article. However, less obvious topics also appear, such as “dopamine”, “hippocampus”, “amygdala” and “decision making”. Roughly, one might view the network in [Fig. 2](#) as being formed by three major clusters. The first cluster (green circles) is dominated by the topic of “learning” and can be said to represent the original concerns of Rescorla and Wagner, namely “associative learning”, “attention”, “blocking”, “overshadowing” and “cue competition”. Note, however, that this cluster also includes additional topics within cognitive psychology, such as “causal learning”, “spatial learning”, “object recognition”, and “visual search”. This suggests that the model first spread its influence to other areas of cognitive psychology, where the discussion focused on topics similar to those that originally preoccupied Rescorla, Wagner, and other associative learning theorists. In line with this observation, in the first few sections of this paper, we will start by presenting the RW model, its empirical and theoretical background, and the reasons behind its frequent use as an explanatory mechanism for a variety of forms of cognition.

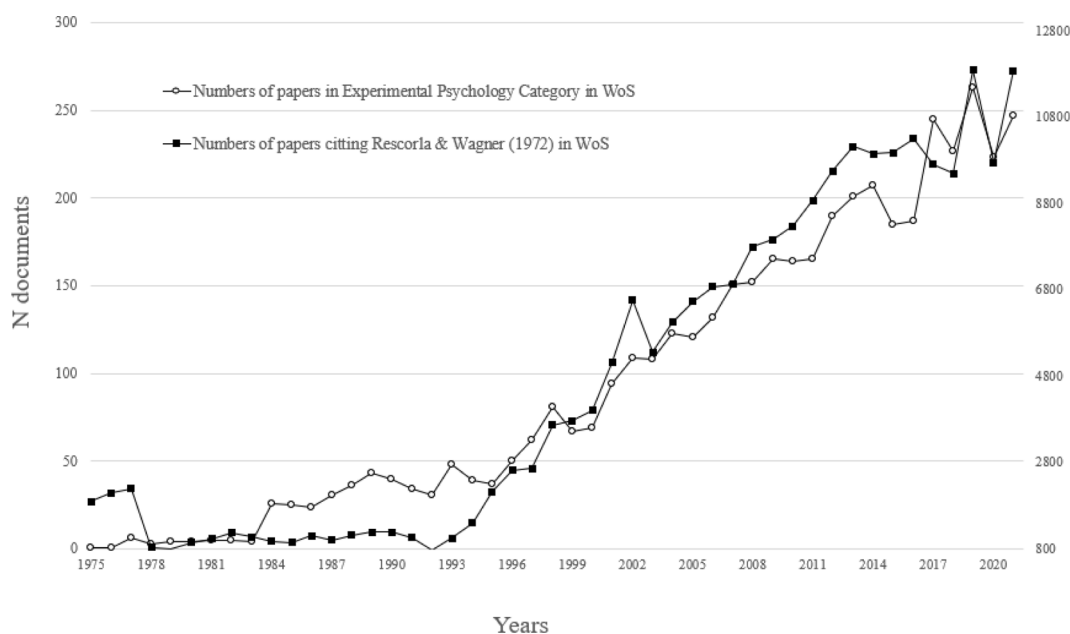
A second important cluster (red circles) comprises topics such as “dopamine”, “fMRI”, “reinforcement learning”, and “prediction error”, which seem to deal with a transition both from behavioral to neural science, and from pavlovian to instrumental conditioning. We will cover this second wave of spreading influence in the last section of our manuscript, explaining the theoretical connection between reinforcement learning models, which deal mainly with instrumental conditioning, and the RW model.

Lastly, an additional cluster (blue circles) is dominated by extinction and seems to deal with fear (“fear” and “fear conditioning”), the processes and determinants of decremental learning (e.g., “context”, “renewal”, “reinstatement” and “spontaneous recovery”), and their applications (“exposure therapy”, “relapse”, “reconsolidation”). We will not cover this vast literature here (but for recent reviews, see [Bouton et al., 2021](#); [Delamater & Westbrook, 2014](#); [Dunsmoor et al., 2015](#)), but we note in passing that it focuses on some of the failures of the RW model, which notoriously provides a poor description of extinction and other so-called interference learning paradigms. The large size of this cluster likely stems from the importance of extinction as an experimental model for the unlearning of fear and anxiety.

Despite the positive reception of the RW model, the theory has by no means been free of severe criticism. For instance, [Miller et al. \(1995\)](#) state that “... given the current status of the Rescorla-Wagner model, it is not appropriate for models outside the framework of traditional learning theory to claim any measure of success because their assumptions, processes, or predictions are concordant with those of the Rescorla-Wagner model”. More recently, [Gallistel \(2021\)](#) regarded that the conceptual framework laid down by the RW model was “the beginning of a tragedy in the history of cognitive neuroscience” (p. 8).

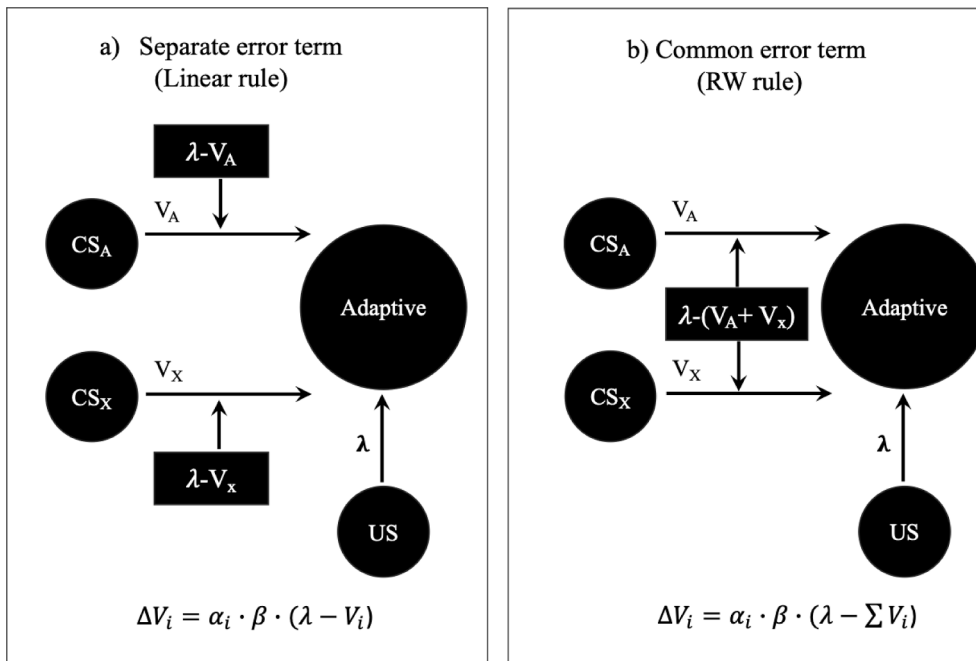
Taken as the cardinal example of the associative approach to learning and cognition ([Wasserman & Castro, 2022](#)), the RW model has also been the target of attack by authors who argue against the assumption that associative learning plays any role at all in human associative learning, let alone in causal learning and other forms of higher-level cognition. From this perspective, learning in causal and predictive tasks can only be captured under the assumption that propositional reasoning is involved ([Mitchell et al., 2009](#)).

Why, then, is the RW model so influential? We argue that the widespread influence of the RW model is not due to the model being unique or better than other models of associative learning, but rather to the commitment of the associative learning community to a “general process” approach and the assumption that the model captures *some* key mechanisms that are at work across many forms of learning and cognition. We will develop these ideas further after a more detailed presentation of the model itself and its theoretical and empirical background.



**Fig. 1.** Number of documents citing the article (left y-axis) and the total number of published papers in experimental psychology (right y-axis) as a function of year.





**Fig. 3.** Two approaches to error computation in Pavlovian conditioning. It is assumed that the presentation of experimental stimuli (CSs A and X and the US) provokes the activation of their respective representational units (A, X and US). The adaptive unit represents the output responsible for the generation of the CR and is unconditionally activated by the US, via  $\lambda$ , and conditionally by the CSs via  $V_A$  and  $V_X$ . According to the linear rule (panel a) the error is computed independently for each CS, while for the RW rule (panel b) the error term is common. The terms  $\alpha$  and  $\beta$  in the equations represent the salience of the CSs and the US, respectively.

The networks depicted in Fig. 3 contain the three basic structural elements of any connectionist theory for conditioning: stimulus representation units, response generation units, and links between units. A fundamental condition for learning is concurrent CS-US processing (Sutton and Barto, 1981). However, although necessary, this simple Hebbian mechanism is not sufficient to account for the fact that over a number of CS-US pairings, the size of the increments in CS-US association becomes smaller as the accumulated associative strength of the CS increases towards an asymptote. This fact, commonly referred to as a negatively accelerated learning curve, was managed by early theoreticians of Pavlovian conditioning (e.g., Atkinson & Estes, 1963; Bush & Mosteller, 1955; Hull, 1943) by assuming what Wagner (1971) called a “saturation principle”, in which each CS can acquire only a limited amount of association with the US, which is provided by the US itself.

The saturation principle is equivalent to stating that learning is due to an error correction mechanism in which the US acts as a supervisor that teaches the CS how to produce the CR via a progressive diminution of the difference between  $\lambda$  and  $V$  (Sutton & Barto, 1998). The left panel of Fig. 3 depicts a graphical representation of this mechanism, where learning accrued to each CS, A and B, depends only on the amount of their respective error terms. Formally, the change in the associative strength accrued to CS  $i$ ,  $\Delta V_i$ , is given by the equation  $\Delta V_i = \alpha\beta(\lambda - V_i)$ , known as “linear rule” (or “local error correction rule”). Here,  $\alpha$  and  $\beta$  represent learning rate parameters associated with the salience or intensity of the CS and US, respectively, and  $\lambda - V_i$  is the error term or reinforcement. According to this equation, CR acquisition occurs in those trials in which the CS and the US are present, and the error term or reinforcement is positive. With the linear rule it is possible to predict the negatively accelerated acquisition curve, since the increments in  $V$  decrease as the net  $V$  value increases; learning reaches its asymptote when the error equals zero. The error correction mechanism involved in the linear rule easily accounts also for extinction. In this case, since  $\lambda = 0$  and  $V > 0$ , the error term becomes negative and  $V_i$  results in a negatively accelerated decrement towards an asymptote of zero.

A relatively evident issue with the linear rule is that it cannot account for the formation of inhibitory associations. Of course, it can be assumed that inhibition develops through other mechanisms, independent from those in charge of the development of excitation (Estes, 1994; Hull, 1943). Another possibility is to assume that CSs interact with one another during learning, such that the amount of associative strength

accrued to a CS depends on the amount of associative strength acquired by all concurrently active CSs. In a typical conditioned inhibition paradigm, reinforced presentations of one CS, A+, are interspersed with nonreinforced presentations of a compound of the same CS and a second stimulus, AB- (Rescorla, 1969). In order to learn this A+ AB- discrimination, B could develop negative associative strength in the AB- trials to counteract the excitation obtained by A in the A+ trials.

Several experiments published in the late sixties provided compelling support for the later possibility. The most commonly cited study is the demonstration of the so-called “blocking phenomenon”, in which the prior reinforcement of A alone reduces (blocks) the acquisition of conditioned response to another CS, B, if it is reinforced together with A (A+/AB+). The conclusion suggested by blocking and similar findings, like overshadowing (Pavlov, 1927) and relative validity (Wagner et al., 1968), is that what is learned to one of the cues on a trial appears to depend not only upon its own current associative value, but also upon the associative value of the other cues present in the trial. This fact is sometimes called stimulus selection or stimulus competition (Wagner, 1969).

The RW model was the first formal explanation proposed for this competition among CSs. The general assumption was that the change in the associative value of a given CS does not depend only on its own associative value, but upon the aggregate value of all CSs present on the trial. This is illustrated in panel (b) of Fig. 3, where it is shown that the only difference with the saturation principle is in the use of a common, instead of an independent, error term for all CSs. Formally, the RW rule posits that the change in associative strength accrued to CS  $i$ ,  $\Delta V_i$ , is given by the equation  $\Delta V_i = \alpha_i\beta(\lambda - \sum_j V_j)$  where  $j$  indexes all stimuli present on a trial, including stimulus  $i$ .

The way in which the aggregated error term works can be illustrated by an experimental situation reported by Wagner (1969), in which animals learned that a compound of two CSs, A and X, was reinforced with the presentation of the US (AX+). Wagner observed that interspersing trials with A alone reinforced (A+) decreased the response to X and that interspersing trials with A alone nonreinforced (A-) increased the response to X. This finding cannot be explained by the separated-error-term notion (Fig. 3a) because  $\Delta V_X$  is independent of  $V_A$ . On the other hand, the RW model predicts that although in all three conditions AX will approach an asymptote of  $\lambda$ , the share of it apportioned to X would

depend on the associative value gained by A alone. When A+ trials are interspersed with AX+ trials, the RW model predicts that  $V_A$  increases in both types of trial. On the other hand,  $V_X$  will change only on AX+ trials, first increasing as long as the predictive value of the compound  $V_A + V_X$  is less than  $\lambda$ , and decreasing once that value becomes larger than  $\lambda$ . That is, as  $V_A$  approaches  $\lambda$  in A+ trials, X will progressively lose the associative strength gained by A. On the contrary, when A- trials are interspersed with AX+ trials, the RW model predicts that as  $V_A$  approaches 0 in A- trials, X will progressively gain the associative strength lost by A.

The RW rule represents a very simple modification in the computation of the error term used in the linear rule, but it meant an enormous conceptual change in the way in which Pavlovian conditioning was interpreted. After this model, Pavlovian conditioning began to be considered as one of the fundamental mechanisms by which animals learn predictive relationships among stimuli. Since the activation of the adaptive unit by the US becomes ineffective in producing reinforcement when similar amounts of activation are obtained by conditioned stimuli, the RW model is regarded as using a “variable reinforcement mechanism”, which also has some implications for instrumental conditioning and (as we will see later) the neuroscience of reward learning. This seemingly trivial addition of a summed prediction in the error term should not be underappreciated.

### 3. The general processes approach and the RW model

The intellectual background of the RW model includes a deep-rooted commitment to the so-called “general processes” approach to the study of behavioral and cognitive processes (Bitterman, 2000; Macphail & Bolhuis, 2001; Papini, 2002; Soto & Wasserman, 2012b). From this perspective, some principles of behavior or cognitive processing are both widely distributed across species and highly useful for the solution of various environmental tasks by any single species.

The main evolutionary argument for this view has two parts (Papini, 2002). First, because of shared ancestry, a common mechanism could have evolved and underlie seemingly variable behavior across species in a given task. Second, some environmental pressures are so common that they affect the selection of similar mechanisms across many species. In the case of Pavlovian conditioning, this would be the pressure to predict biologically significant outcomes from the presence of informative environmental cues. Because the main goal set by generalists is to discover the principles of behavior that are common to many species, they tend to study how distantly related species solve similar environmental tasks. If distantly related species show evidence of using similar behavioral mechanisms to solve a given task, then either these mechanisms have been conserved across evolution or their similarity stems from convergent evolution under similar environmental pressures. Both cases lead to the discovery of “general principles” of behavior, reflecting the general solution to an environmental task displayed by multiple organisms, beyond any adaptations evolved by a specific species to specialize to their own environmental niche. General process theorists, including those studying associative learning in the 1960s and 1970s, choose to focus on the analysis of behavioral mechanisms that are similar across species, sometimes ignoring the exact evolutionary explanation for their similarity (Riley & Langley, 1993).

This way, the empirical phenomena that informed the development of the RW model (i.e., stimulus competition phenomena) came from a variety of species and behaviors. For example, overshadowing was first reported in conditioning of the salivary response in dogs (Pavlov, 1927) and, before the publication of the RW model, also found in rabbits, cats, humans, and pigeons using varied preparations (for a review, see Baker, 1968). Similarly, blocking was first reported using fear conditioning in rats (Kamin, 1969), and later found in pigeons (Mackintosh & Honig, 1970; vom Saal & Jenkins, 1970), and rabbits (Wagner, 1969). The assumption underlying the development of the RW model was that it would be possible to capture the principles underlying associative learning across the variety of species, behaviors, and neurobiological

circuits involved in this prior literature. This assumption proved to be correct: later surveys of the literature have shown that associative learning and stimulus competition phenomena can be found not only among mammals and birds, but across all vertebrates (Macphail, 1982) and some invertebrate phyla including arthropods, platyhelminthes (flatworms), and molluscs (C. Heyes, 2012; Loy et al., 2021), suggesting that associative learning, as described by the RW model, might be common to all nephrozoa (i.e., most animals; Ginsburg & Jablonka, 2010).

The general process approach to the study of animal behavior was not only important for the development of the RW model, but is also key to understanding why the model became as influential as it is today. Learning theorists saw in the RW model a summary of well-documented general principles of associative learning, which should be at play not only in the species and procedures typically used to study that form of learning, but across a wide set of complex behaviors and forms of cognition.

### 4. The RW model spreads its influence to cognitive psychology

Under the influence of the general processes approach, starting in the late 1970s and early 1980s a number of researchers sought to determine to what extent associative mechanisms could be at work in complex forms of cognition, a trend that continues to this day. Two types of evidence are consistently used in this literature to claim that associative learning might be involved in a particular form of cognition, and both are related to the RW model (for a review of the early work, see Siegel & Allan, 1996).

The first type of evidence came from theoretical work. The original RW model or a modified version were used to show that associative learning mechanisms could explain aspects of a particular form of cognition. Examples are paired-associate learning (Rudy, 1974), generalization (Blough, 1975; Gluck, 1991), category learning (Gluck & Bower, 1988; Shanks, 1991), causal learning (Shanks & Dickinson, 1988; Van Hamme & Wasserman, 1994), transitive inference (Wynne, 1995), abductive reasoning (Wang et al., 1999), spatial cognition (Chamizo, 2002; N. Y. Miller & Shettleworth, 2007; Pearce, 2009), perceptual learning (Hall, 2009; Kahnt et al., 2011), object recognition (Soto et al., 2012; Soto & Wasserman, 2010), sequential learning (Gureckis & Love, 2010), social group biases (Vanhoomissen & Van Overwalle, 2010), language learning (Baayen et al., 2011; Nixon & Tomaschek, 2021; Ramsar et al., 2013), imitation (Cooper et al., 2013), word morphological processing (Milin et al., 2017), sound categories (Olejarczuk et al., 2018), speech perception (Arnold et al., 2017), memory modification (Gershman et al., 2017), planning (Lind, 2018), social learning (N. Miller, 2018), semantic learning (Hollis, 2019), and learning about personality traits during social interaction (Frolich et al., 2022), among others.

It is important to note two points regarding these applications of the RW to phenomena outside its original domain. First, when most of these applications were published, a variety of models proposed after the RW model were also available to explain stimulus competition phenomena (Blough, 1975; Mackintosh, 1975; Pearce, 1987; Pearce & Hall, 1980; Wagner, 1981). Indeed, some of these models were used by generalists to make the point that associative learning might underlie a form of cognition, as was the case in the application of Wagner’s SOP model (Wagner, 1981) to causal learning phenomena (e.g., Dickinson & Burke, 1996). Still, the RW model was most often used, perhaps to honor the fact that it was the first to propose a global error term. Second, most of these applications do not strictly apply the original RW model, but rather modifications that keep the assumptions in the original model (learning by reduction of global prediction error) while expanding its scope. Both of these features underscore that the main point of this literature is not to promote the RW model itself, but rather the general process hypothesis that associative learning is involved in many cognitive phenomena, *under the assumption that other mechanisms* (e.g.,

perceptual and cognitive) would also be at work (hence the modifications to the model).

The second type of evidence came from empirical work. Studies were performed to show that the stimulus competition phenomena that inspired the development of the RW model, such as blocking, are also found in a variety of cognitive tasks (e.g., Beesley & Shanks, 2012; Chapman, 1991; Cook et al., 2010; Cramer et al., 1985; Mata et al., 2021; Nixon, 2020; Pearce et al., 2006; Rodrigo et al., 1997; Sanbonmatsu et al., 1994; Shanks, 1985; Soto & Wasserman, 2010, 2012a; Vogel et al., 2015). This extensive literature that sought to find evidence of stimulus competition phenomena serves to further show that it is associative learning in general, and not the RW model, that is being proposed as a general mechanism for cognition.

Thus, the main driving force behind the influence of the RW model in cognitive psychology was not the model itself, but the proposal that associative learning underlies many forms of human cognition. The RW model became influential because of its ability to succinctly capture many of the mechanisms underlying associative learning, summarizing decades of research on animal learning and capturing a variety of phenomena not yet tested when the model was first published.

An advantage of this strong focus on the RW model was that it led to an improved theoretical understanding of the model. From the point of view of cognitive modeling, the RW model functions at the algorithmic level (Marr, 1982), in that it proposes explicitly how knowledge about the environment is represented (as associations between stimulus representations) as well as the algorithm used to acquire such knowledge (global prediction error correction). On the other hand, many of the models competing with the RW model in different fields are proposed at the computational level of analysis (also called rational or normative; see Anderson, 1990). Models at this level focus on formalizing both the goals of the cognitive system and the environmental task to which it is exposed, in order to determine what computations must be carried out to solve that task. Algorithmic and computational models can be viewed as complementary instead of mutually exclusive accounts but, because of the pressure to compare the RW model against computational-level models, much was learned about exactly how the RW model could be described at the computational level of analysis.

More specifically, several authors have shown (Chapman & Robbins, 1990; Cheng, 1997; Danks, 2003; Yuille, 2006) that the long-run behavior of the RW model (the equilibrium reached after long training with the same event contingencies) approximates the predictions of the probabilistic contrast model (Cheng & Novick, 1992), which was developed to predict how a rational observer would judge covariation between a cause and an outcome, while taking into account other potential causes. Formally, the covariation between  $CS_i$  and the US,  $\Delta P_i$ , is calculated as the probability of observing the US in a given context X when  $CS_i$  is present, or  $P(US|X, CS_i)$  minus the probability of observing the US in the same context X in the absence of  $CS_i$ , or  $P(US|X, \neg CS_i)$ . The context X works as a placeholder for all other potential predictors of the US. Because the long-run behavior of the RW model approximates  $\Delta P_i$  in many scenarios, the RW model solves the task of learning how good a cue is at predicting an outcome by computing something akin to the partial correlation between them, removing the effect of other potential cues.

While this computational-level characterization of the RW model is useful, it also strips the theory from most of its theoretical insights: the principle that it is contiguity between events, and not mere covariation, what determines learning (see Papini & Bitterman, 1990), the role of prediction error in driving learning, and the fact that learning is a dynamic process in which the ordering of learning events matters (see Allan, 2003).

A disadvantage of the vast influence of the RW model in cognitive psychology is that it became a straw man representing “associative learning,” a way to prove that this form of learning was not involved in a particular form of cognition, rather than a way to show that it was. For example, in the causal learning literature it became fashionable to

equate the RW model with associative learning theory in general (see Le Pelley et al., 2017), and beyond that, to take the predictions of the probabilistic contrast model as the rational equivalent of the predictions of the RW model. Thus, the typical strategy would be to compare the predictions of a computational-level theory of causal learning against the probabilistic contrast model and, if the former fitted the data better than the latter, then conclude that associative learning theories are not a good account for human causal learning (e.g., Cheng, 1997; Waldmann et al., 2008).

This strategy, of course, is against the original spirit of the general process approach that drove the application of the RW model to areas outside of animal conditioning, and has been criticized by theorists in the associative tradition (Allan, 2003; Le Pelley et al., 2017; Lober & Shanks, 2000). From this perspective, the RW model represents some, but not all mechanisms involved in Pavlovian conditioning and other forms of cognition. Since its inception, it has been recognized that the RW model does not provide a complete explanation for human and animal associative learning (see R. R. Miller et al., 1995), and a number of alternative theories have been developed to account for all its failures (for a review, see Vogel et al., 2004). These developments should be taken into account to reach any general conclusion about the inadequacy of associative learning mechanisms to explain human causal learning or any other form of cognition.

As indicated earlier, most applications of the RW model involve modifications that capture idiosyncrasies of the cognitive phenomena being the target of explanation. Still, the original contribution of the RW model of a global error-correcting learning rule has been implemented in most algorithmic models of associative learning, as well as in algorithmic models of causal learning developed to approximate the predictions of computational-level competitors of the RW model (Danks et al., 2003). It is this feature of the model that seems to be key, rather than its computation of partial covariation.

## 5. The RW model spreads its influence to neuroscience

We have already seen how the RW model quickly spread its influence to cognitive psychology, where it climbed its way up the Marrian ladder, influencing and being influenced by theoretical work at the computational level of analysis. Starting around the 2000s, the model would spread down to Marr’s implementational level of analysis. At the time, the field was ripe for investigating the brain areas that deployed the mechanisms postulated in the RW algorithm.

The most important way in which the RW model spread its influence to neuroscience was indirect, pushed by the growing interest of that community on reinforcement learning (RL; Sutton & Barto, 1998), an area of machine learning motivated by the idea that the same algorithms of prediction error used in associative learning theory could be analyzed in the normative terms previously proposed by Marr.

As noted above, the RW model was formulated to explain Pavlovian arrangements where there was a clear demarcation between trials. The assumption was that the discretization of events and trials would be good enough to capture a wealth of data in animal conditioning. However, multiple findings could not be explained using such a coarse-grained analysis. In second-order conditioning, for example, a stimulus B is first reinforced and then another stimulus A is paired with B, with the result that both stimuli accrue predictive value even though A has never been paired with a reward. As we have seen, the RW model update rule is applied at the end of each episode, with nothing happening within episodes or during the ISIs which separate trials. The model has no mechanism for B to become a reinforcer, or to distinguish between  $A \rightarrow B$  trials,  $B \rightarrow A$  trials, and  $AB$  trials. To overcome this, Sutton and Barto (1981, 1990) modified the RW rule to capture this and other phenomena. They proposed two critical modifications to the RW model. The first one is that the predictions and updates of  $V$  are made in real time, for each discrete time step  $t$  during an experiment. This is partly achieved by assuming that a CS has a so-called *eligibility trace*,

which can be interpreted as the internal representation of a stimulus that weighs how much associative strength it can acquire at any given time. The eligibility trace is highest during presentation of a CS but it remains active for a limited time after the CS has disappeared. As a consequence, the temporal arrangement of stimuli during an episode matters for the modification of  $V$ . This is a critical advancement of TD over RW, as it considers the true temporal contiguity between the internal representations of stimuli and outcomes as the main driver of learning, and not only the contiguity as defined by the experimenter (i.e., trialwise). The second modification is that the prediction errors are not only computed when a reinforcer is presented; animals can also be surprised when another stimulus that has acquired associative strength is presented. In other words, there is a mechanism by which CSs that have acquired associative strength themselves become reinforcing, just as the US. Taken together, these two additions can capture the results of a second-order conditioning experiment and many other phenomena which were not captured by the RW model.

More formally, the temporal-difference algorithm (TD; Sutton & Barto, 1990) for the acquisition of associative strength for each time-step  $t$  is:  $V_t^{new} = V_t^{old} + \alpha \delta_{t+1}$ , where  $\delta_{t+1} = (\lambda_{t+1} + \gamma \sum V_{t+1}^{old}) - \sum V_t^{old}$ , that is, the error term  $\delta_{t+1}$  is the difference between the value of events at time  $t + 1$  (in parentheses) and the value predicted at time  $t$  from all the stimuli present at that time (for clarity, we have omitted the influence of eligibility traces from the equations). Note that the reinforcement value in parentheses considers both the value of direct reinforcement  $\lambda_{t+1}$  as well as the value of events presented at time  $t + 1$ . This means that events that have become valuable through association with a reinforcer themselves become reinforcing, and that value can backpropagate in time from the reinforcer to the onset of a predictive cue, both features that differentiate TD from the RW model. The parameter  $\gamma$  is a number between 0 and 1 that reflects how much the subject discounts the reinforcing value of future events with respect to current values.

There are multiple theoretical connections between the TD and RW models that are worth noting. The first and most obvious one is that TD, as it was originally implemented by Sutton and Barto (see Sutton & Barto, 1990, p. 533) computes expected reinforcement based on the sum of associative strengths of all cues present at a given time step  $t$ . That is, the model preserves the assumption of learning from correction of global prediction error, which is probably the most significant advancement of RW with respect to its predecessors. Many publications in the machine learning and neuroscience literature omit this part of the equations, which obscures the strong influence of the RW model in the development of TD and other reinforcement learning algorithms. A second theoretical connection is that TD extends the logic behind RW by calculating expected reinforcement based not only on a sum of associative strengths of all *present* cues, but also all *past* cues whose eligibility trace is still active at time  $t$ .

A third theoretical connection is that, at the computational level, both the TD and RW algorithms solve the so-called credit assignment problem (see Sammut, 2011). When a learning agent observes a complex series of events or decision process leading to a given outcome, one problem is that it must apportion credit or blame for the outcome to all the different elements leading to it. The temporal credit assignment refers to the problem of assigning credit to multiple actions that have been executed in sequence leading to the outcome. The structural credit assignment refers to the problem of assigning credit to multiple aspects of the environment or system state (e.g., cues or representational features) leading to the outcome. The TD algorithm, and other reinforcement learning models commonly used in the neuroscientific literature, solve the temporal credit assignment problem. The RW model, with its ability to compute the covariation between a cue and an outcome while removing the influence of other cues, is one way in which the structural credit assignment problem can be solved.

A body of work in neuroscience shows that both TD learning and RW seem to be implemented in the brain. Although the neurotransmitter

dopamine had been hypothesized for some time as being implicated in motor function, cognitive skills, and motivation (from evidence of performance deficits in patients suffering from some sort of dopamine depletion, such as that produced by Parkinson's disease), it became clear during the second half of the 1990s that dopamine was also involved in reward processing. In one of the most highly-cited papers in neuroscience, Schultz et al., (1997) demonstrated how midbrain dopamine neurons in the ventral tegmental area (VTA) and substantia nigra followed the predictions of the TD rule. They trained monkeys in a Pavlovian task where one CS would be followed by fruit juice, the reinforcer. Recordings of midbrain neurons showed a remarkable fit to the predictions of the TD rule, in that before learning they would fire at the onset of the reinforcer but these responses would gradually move back to the onset of the CS that predicted the reinforcer. Moreover, when the juice reinforcer was omitted after presentation of a trained CS, neurons would show a dip in responding, just as if the TD term was predicting the unexpected omission of the reinforcer. Later on, it was found that dopamine responses correlated with expected value, that is, the product of probability and magnitude of a potential reward, which is also a prediction embedded in both TD and RW algorithms (Fiorillo et al., 2003).

As we have already discussed, the fact that all CSs presented in a trial contribute to prediction of reinforcement (and thus to prediction error) is probably the main feature of the RW model. Using a classic blocking design in monkeys (A+/B-; AX+, BY+; test X and Y), Waelti, Dickinson and Schultz (2001) found that dopamine neuron activity was directly associated with the summed prediction error, but not with simple pairing of CSs and reinforcement; neural activity could be blocked for a given stimulus even if it was paired with the reinforcer when the associative strength of another accompanying stimulus was already high.

More recent evidence has cast doubt on the idea that dopamine signals carry only information about reward prediction error (see Sharpe et al., 2017; Takahashi et al., 2017). However, it is at present clear that prediction error, and in particular dopamine responses, are sufficient to produce learning to stimuli that would have otherwise not accrued any learning. For example, Steinberg et al. (2013) used optogenetic activation of dopamine neurons in rats to show that a cue could be unblocked by simply generating dopamine activity at the time the reward was presented in the second stage of a blocking design. In practice, what the authors did was to artificially generate a reward prediction error that would give room for learning to occur in spite of the target cue not being relevant for prediction after training. This is direct evidence of reward prediction error encoding by dopamine neurons, and its role in Pavlovian conditioning.

When first introduced in 1972, the RW model was applied to several phenomena in conditioning that were not captured by the linear rule previously used by associative learning theorists. But perhaps one of the most significant factors explaining its influence is that it also made several predictions that were later confirmed by experimental data. The phenomenon of overexpectation is one of them. In an overexpectation experiment, the experimenter arranges for two cues to independently predict a reinforcer (A+; B+) and then trains both of these stimuli with the same level of outcome (AB+). Given that the summed prediction of AB together is higher than that of the US, the RW model predicts that  $V_A$  and  $V_B$  should decrease due to a negative prediction error, rather than staying constant. Chang et al. (2016) used this design together with optogenetics in rats to test if the negative prediction errors could be minimized by inhibiting dopamine release during the second phase and, indeed, they found that such manipulation prevented unlearning during the second phase.

As shown in Fig. 2, the literature linking reinforcement learning and dopamine is one of the areas in which the RW model is most cited. However, the model has also guided research on the neurobiology of associative learning in other systems, including motor conditioning in the cerebellum (e.g., Gluck et al., 2001; Rasmussen, 2020; Rasmussen et al., 2015), and fear conditioning in the amygdala and related areas (e.

g., Fanselow, 1998; Johansen et al., 2010; McNally et al., 2011; McNally & Cole, 2006; Roy et al., 2014; Walker et al., 2020). Nevertheless, it is important to note that the neurobiological data is compatible with any model of associative learning proposing a global error correcting algorithm (e.g., Mackintosh, 1975; Pearce, 1987; Pearce & Hall, 1980; Wagner, 1981; for an alternative explanation, see Witnauer et al., 2014), not exclusively the RW model. In addition, other neurobiological data are more in line with those alternative models (e.g., Holland & Maddux, 2010; Holland & Schiffrino, 2016). Once again, applications of the RW and related models in the neuroscience literature have more to do with what the model *represents* than with a judgment of the model being superior to its competitors. In this case, what the model represents is (1) an assumption that the mechanisms of associative learning are similar across many behavioral tasks, regardless of the specific neural circuits involved, and (2) a relatively strong consensus that among those mechanisms is learning by reduction of global prediction error.

## 6. What can we learn from the success of the RW model?

Our main hypothesis here has been that the widespread influence of the RW model is not due to the model being unique or better than other models of associative learning. Indeed, the associative learning community has known for a long time that the model cannot explain a large body of phenomena (R. R. Miller et al., 1995). Still, members of this community have insisted on exporting the model to other areas of the cognitive sciences. Why? Their underlying assumption is not that the model is complete and consistent with all observations, but rather that it provides insights on *some* of the key mechanisms of associative learning that might be at work across many forms of learning and cognition. The success of the model to explain phenomena across levels of description and fields shows the correctness of this assumption. The failure of the model to explain other phenomena shows that (1) the model does not capture all regularities of associative learning, and (2) associative mechanisms are not the only ones at work in complex cognition.

The widespread influence of the RW model is largely the result of a community of scholars insisting on the value of a general processes approach. In some cases, this is an inheritance from Hull's neo-behaviorism and its push to develop comprehensive theories of behavior, but across the years the community's perspective has expanded well beyond those early ideas. Importantly, it has returned to its Pavlovian roots (Pavlov, 1927), where behavior is seen as a biological process in need of biological/physiological explanations (Konorski, 1948, 1967; and Sokolov, 1963 are further examples of this early tradition). The general processes approach is now understood under the light of evolutionary theory, and the study of general processes is importantly influenced by neurobiology.

There are three important take-home messages from the history of the RW model, if we want to repeat its success in the future. First, it is likely that there are other general processes besides associative learning. We should aim to discover them, study them at great length, and succinctly describe their algorithmic mechanisms in formal models. Second, when a model becomes as influential as the RW model, we should avoid the natural tendency to simplify its history, assumptions, and key mechanisms. Although blocking was among the phenomena that inspired the development of the RW model, many other empirical results supported the model before and after its publication. No model as successful as the RW model is built on the basis of a single experiment, but rather on a strong empirical basis that does not crumble when the results from a single study are reevaluated (Soto, 2018). Similarly, although the RW model (for a given choice of parameters) can be seen as computing partial covariance between events (equivalently to the probabilistic contrast model; Cheng & Novick, 1992), this is not the only or best way to characterize the model. Although brain structures can be seen as implementing the computations proposed by the RW model, similar computations are also proposed by other models (e.g., Mackintosh, 1975; Pearce, 1987; Pearce & Hall, 1980; Wagner, 1981).

The third point is particularly important given that we are currently in the midst of neuroimaging, machine learning, and big data revolutions. With the advent of fMRI, neuroimaging has taken over the research programs of many psychological departments and laboratories around the world. With the development of powerful machine learning algorithms, those algorithms have taken over the theoretical development in many fields of psychology, and there is a trend toward their application to large databases. All these developments have advantages in their own right. However, they are contrary to the spirit behind the success of the RW model, which was developed by putting emphasis on the value of behavioral data, careful design of highly-controlled experiments ("key data" rather than "big data"), and a "data-first" approach to theory development, where data should guide the slow development of basic explanatory mechanisms, which can then be applied to more complex phenomena, rather than using ready-made algorithms and models (e.g., imported from machine learning) to explain complex behavioral or neuroscientific datasets. We claim that formalization is only useful as long as it is simple enough to allow the theory (a) to be easily falsifiable through precisely targeted experimental work and (b) to allow a smooth transition up and down the Marrian ladder. As we have seen, the RW model was very unique in accomplishing both of these purposes successfully.

Fortunately, there are still remnants in our community of the values that gave strength to the RW model. There have been recent calls to put more emphasis on behavioral data in behavioral and cognitive neuroscience (Niv, 2021), and a trend toward more scrutiny and rigorous testing of theoretical frameworks imported from the machine learning literature (Bowers et al., 2022; Jones & Love, 2011). Many in the neuroscientific community, especially those using animal models, have had a key role in preserving the approach to research that gave us the RW model and spread its influence across cognitive science (see for example, Ottenheimer et al., 2020; Sadacca et al., 2018; Sharpe et al., 2017), although they have been more insular than associative learning theorists from the 70s and 80s. Perhaps it is their turn now to carry the torch lit by Rescorla and Wagner.

## CRediT authorship contribution statement

**Fabian A. Soto:** Conceptualization, Writing – original draft, Writing – review & editing. **Edgar H. Vogel:** Conceptualization, Writing – original draft, Writing – review & editing, Visualization. **Yerco E. Uribe-Bahamonde:** Data curation, Formal analysis, Visualization. **Omar D. Perez:** Conceptualization, Writing – original draft, Writing – review & editing.

## Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## Data availability

No data was used for the research described in the article.

## Funding

Omar D. Perez is supported by ANID-SIA 85220023, ANID-FONDECYT 1231027 and ANID-PIA/PUENTE AFB220003.

## Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.nlm.2023.107794>.



## References

- Allan, L. G. (2003). Assessing power PC. *Learning & Behavior*, 31(2), 192–204.
- Anderson, J. R. (1990). *The adaptive character of thought*. Lawrence Erlbaum Associates.
- Arnold, D., Tomaschek, F., Sering, K., Lopez, F., & Baayen, R. H. (2017). Words from spontaneous conversational speech can be recognized with human-like accuracy by an error-driven learning algorithm that discriminates between meanings straight from smart acoustic features, bypassing the phoneme as recognition unit. *PLOS ONE*, 12(4), e0174623.
- Atkinson, R. R., & Estes, W. K. (1963). Stimulus sampling theory. In R. D. Luce, & R. B. Bush (Eds.), *Handbook of Mathematical Psychology* (pp. 212–268). Wiley.
- Baayen, R. H., Milin, P., Durdević, D. F., Hendrix, P., & Marelli, M. (2011). An amorphous model for morphological processing in visual comprehension based on naive discriminative learning. *Psychological Review*, 118(3), 438–481. <https://doi.org/10.1037/a0023851>
- Baker, A. G. (1974). Conditioned inhibition is not the symmetrical opposite of conditioned excitation: A test of the Rescorla-Wagner model. *Learning and Motivation*, 5(3), 369–379.
- Baker, T. W. (1968). Properties of compound conditioned stimuli and their components. *Psychological Bulletin*, 70, 611–625. <https://doi.org/10.1037/h0026744>
- Beesley, T., & Shanks, D. R. (2012). Investigating cue competition in contextual cuing of visual search. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 38, 709–725. <https://doi.org/10.1037/a0024885>
- Bitterman, M. E. (2000). Cognitive evolution: A psychological perspective. In C. M. Heyes, & L. Huber (Eds.), *The evolution of cognition* (pp. 61–79). MIT Press.
- Blough, D. S. (1975). Steady state data and a quantitative model of operant generalization and discrimination. *Journal of Experimental Psychology: Animal Behavior Processes*, 104(1), 3–21.
- Bouton, M. E., Maren, S., & McNally, G. P. (2021). Behavioral and neurobiological mechanisms of pavlovian and instrumental extinction learning. *Physiological Reviews*, 101(2), 611–681. <https://doi.org/10.1152/physrev.00016.2020>
- Bowers, J. S., Malhotra, G., Dujmović, M., Montero, M. L., Tsvetkov, C., Biscione, V., et al. (2022). Deep problems with neural network models of human vision. *Behavioral and Brain Sciences*, 1–74. <https://doi.org/10.1017/S0140525X22002813>
- Bush, R. R., & Mosteller, F. (1955). *Stochastic models for learning* (pp. xvi, 365). John Wiley & Sons, Inc.
- Chamizo, V. D. (2002). Spatial learning: Conditions and basic effects. *Psicológica*, 23, 33–57.
- Chang, C. Y., Esber, G. R., Marrero-García, Y., Yau, H. J., Bonci, A., & Schoenbaum, G. (2016). Brief optogenetic inhibition of dopamine neurons mimics endogenous negative reward prediction errors. *Nature Neuroscience*, 19(1), Article 1. <https://doi.org/10.1038/nn.4191>
- Chapman, G. B. (1991). Trial order affects cue interaction in contingency judgment. *Journal of Experimental Psychology: Learning Memory and Cognition*, 17(5), 837–854.
- Chapman, G. B., & Robbins, S. J. (1990). Cue interaction in human contingency judgment. *Memory & Cognition*, 18(5), 537–545.
- Cheng, P. W. (1997). From covariation to causation: A causal power theory. *Psychological Review*, 104(2), 367–405.
- Cheng, P. W., & Novick, L. R. (1992). Covariation in natural causal induction. *Psychological Review*, 99(2), 365–382.
- Clarke, J. C., Westbrook, R. F., & Irwin, J. (1979). Potentiation instead of overshadowing in the pigeon. *Behavioral and Neural Biology*, 25(1), 18–29. [https://doi.org/10.1016/S0163-1047\(79\)90705-2](https://doi.org/10.1016/S0163-1047(79)90705-2)
- Cook, R., Press, C., Dickinson, A., & Heyes, C. (2010). Acquisition of automatic imitation is sensitive to sensorimotor contingency. *Journal of Experimental Psychology: Human Perception and Performance*, 36, 840–852. <https://doi.org/10.1037/a0019256>
- Cooper, R. P., Cook, R., Dickinson, A., & Heyes, C. M. (2013). Associative (not Hebbian) learning and the mirror neuron system. *Neuroscience Letters*, 540, 28–36. <https://doi.org/10.1016/j.neulet.2012.10.002>
- Cramer, R. E., Weiss, R. F., Steigleder, M. K., & Balling, S. S. (1985). Attraction in context: Acquisition and blocking of person-directed action. *Journal of Personality and Social Psychology*, 49, 1221–1230. <https://doi.org/10.1037/0022-3514.49.5.1221>
- Danks, D. (2003). Equilibria of the Rescorla-Wagner model. *Journal of Mathematical Psychology*, 47(2), 109–121.
- Danks, D., Griffiths, T. L., & Tenenbaum, J. B. (2003). Dynamical causal learning. In S. Becker, S. Thrun, & K. Obermayer (Eds.), *Advances in Neural Information Processing* (pp. 67–74). MIT Press; Google Scholar.
- Delamater, A. R., & Westbrook, R. F. (2014). Psychological and neural mechanisms of experimental extinction: A selective review. *Neurobiology of Learning and Memory*, 108, 38–51. <https://doi.org/10.1016/j.nlm.2013.09.016>
- Dickinson, A., & Burke, J. (1996). Within-compound associations mediate the retrospective reevaluation of causality judgements. *Quarterly Journal of Experimental Psychology*, 49B(1), 60–80.
- Dunsmoor, J. E., Niv, Y., Daw, N., & Phelps, E. A. (2015). Rethinking Extinction. *Neuron*, 88(1), 47–63. <https://doi.org/10.1016/j.neuron.2015.09.028>
- Estes, W. K. (1994). Toward a statistical theory of learning. *Psychological Review*, 101(2), 282–289. <https://doi.org/10.1037//0033-295X.101.2.282>
- Fanselow, M. S. (1998). Pavlovian conditioning, negative feedback, and blocking: Mechanisms that regulate association formation. *Neuron*, 20(4), 625.
- Fiorillo, C. D., Tobler, P. N., & Schultz, W. (2003). Discrete coding of reward probability and uncertainty by dopamine neurons. *Science*, 299(5614), 1898–1902. <https://doi.org/10.1126/science.1077349>
- Frey, P. W., & Sears, R. J. (1978). Model of conditioning incorporating the Rescorla-Wagner associative axiom, a dynamic attention process, and a catastrophe rule. *Psychological Review*, 85(4), 321–340. <https://doi.org/10.1037/0033-295X.85.4.321>
- Frolich, K. M. M., Rosenblau, G., & Korn, C. W. (2022). Incorporating social knowledge structures into computational models. *Nature Communications*, 13(1), Article 1. [10.1038/s41467-022-33418-2](https://doi.org/10.1038/s41467-022-33418-2).
- Gallistel, C. R. (2021). Robert Rescorla: Time, Information and Contingency. *Revista de Historia de La Psicología*, 42(1), 7–21.
- Gershman, S. J., Monfils, M. H., Norman, K. A., & Niv, Y. (2017). The computational nature of memory modification. *ELife*, 6, e23763.
- Gillan, D. J., & Domjan, M. (1977). Taste-aversion conditioning with expected versus unexpected drug treatment. *Journal of Experimental Psychology: Animal Behavior Processes*, 3(4), 297–309. <https://doi.org/10.1037/0097-7403.3.4.297>
- Ginsburg, S., & Jablonka, E. (2010). The evolution of associative learning: A factor in the Cambrian explosion. *Journal of Theoretical Biology*, 266(1), 11–20. <https://doi.org/10.1016/j.jtbi.2010.06.017>
- Gluck, M. A. (1991). Stimulus generalization and representation in adaptive network models of category learning. *Psychological Science*, 2(1), 50–55.
- Gluck, M. A., Allen, M. T., Myers, C. E., & Thompson, R. F. (2001). Cerebellar substrates for error correction in motor conditioning. *Neurobiology of Learning and Memory*, 76(3), 314–341. <https://doi.org/10.1006/nlme.2001.4031>
- Gluck, M. A., & Bower, G. H. (1988). From conditioning to category learning: An adaptive network model. *Journal of Experimental Psychology: General*, 117(3), 227–247.
- Groves, P. M., & Thompson, R. F. (1970). Habituation: A dual-process theory. *Psychological Review*, 77(5), 419–450. <https://doi.org/10.1037/h0029810>
- Gureckis, T. M., & Love, B. C. (2010). Direct associations or internal transformations? Exploring the mechanisms underlying sequential learning behavior. *Cognitive Science*, 34(1), 10–50. <https://doi.org/10.1111/j.1551-6709.2009.01076.x>
- Hall, G. (2009). Perceptual learning in human and nonhuman animals: A search for common ground. *Learning & Behavior*, 37(2), 133–140.
- Heyes, C. (2012). Simple minds: A qualified defence of associative learning. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 367(1603), 2695–2703. <https://doi.org/10.1098/rstb.2012.0217>
- Holland, P. C., & Maddux, J. M. (2010). Brain systems of attention in associative learning. In C. J. Mitchell, & M. E. L. Pelley (Eds.), *Attention and associative learning: From brain to behaviour* (pp. 187–216). Oxford University Press.
- Holland, P. C., & Schiffrino, F. L. (2016). Mini-review: Prediction errors, attention and associative learning. *Neurobiology of Learning and Memory*, 131, 207–215. <https://doi.org/10.1016/j.nlm.2016.02.014>
- Hollis, G. (2019). Learning about things that never happened: A critique and refinement of the Rescorla-Wagner update rule when many outcomes are possible. *Memory & Cognition*, 47(7), 1415–1430. <https://doi.org/10.3758/s13421-019-00942-4>
- Hull, C. L. (1943). *Principles of Behavior: An Introduction to Behavior Theory*. Appleton-Century-Crofts.
- Johansen, J. P., Tarpley, J. W., LeDoux, J. E., & Blair, H. T. (2010). Neural substrates for expectation-modulated fear learning in the amygdala and periaqueductal gray. *Nature Neuroscience*, 13(8), Article 8. [10.1038/nn.2594](https://doi.org/10.1038/nn.2594)
- Jones, M., & Love, B. C. (2011). Bayesian fundamentalism or enlightenment? On the explanatory status and theoretical contributions of bayesian models of cognition. *Behavioral and Brain Sciences*, 34(04), 169–188. <https://doi.org/10.1017/S0140525X10003134>
- Kahnt, T., Grueschow, M., Speck, O., & Haynes, J. D. (2011). Perceptual learning and decision-making in human medial frontal cortex. *Neuron*, 70(3), 549–559.
- Kamin, L. J. (1969). Selective association and conditioning. In N. J. Mackintosh, & W. K. Honig (Eds.), *Fundamental Issues in Associative Learning* (pp. 42–64). Dalhousie University Press.
- Konorski, J. (1948). *Conditioned Reflexes and Neuron Organization*. Cambridge University Press.
- Konorski, J. (1967). *Integrative activity of the brain*. Chicago.
- Le Pelley, M. E., Griffiths, O., & Beesley, T. (2017). Associative accounts of causal cognition. *The Oxford Handbook of Causal Reasoning*, 13–28.
- Levitin, L. (1975). Tests of the Rescorla-Wagner model of Pavlovian conditioning. *Bulletin of the Psychonomic Society*, 6(3), 265–268.
- Lind, J. (2018). What can associative learning do for planning? *Royal Society Open Science*, 5(11), Article 180778. <https://doi.org/10.1098/rsos.180778>
- Lober, K., & Shanks, D. R. (2000). Is causal induction based on causal power? Critique of Cheng (1997). *Psychological Review*, 107(1), 195–212.
- Loy, I., Carnero-Sierra, S., Acebes, F., Muñoz-Moreno, J., Muñoz-Diez, C., & Sánchez-González, J. C. (2021). Where association ends. A review of associative learning in invertebrates, plants and protista, and a reflection on its limits. *Journal of Experimental Psychology: Animal Learning and Cognition*, 47, 234–251. <https://doi.org/10.1037/xan0000306>
- Mackintosh, N. J. (1975). A theory of attention: Variations in the associability of stimuli with reinforcement. *Psychological Review*, 82, 276–298.
- Mackintosh, N. J. (1976). Overshadowing and stimulus intensity. *Animal Learning & Behavior*, 4(2), 186–192.
- Mackintosh, N. J., & Honig, W. K. (1970). Blocking and enhancement of stimulus control in pigeons. *Journal of Comparative and Physiological Psychology*, 73, 78–85. <https://doi.org/10.1037/h0030021>
- Macphail, E. M. (1982). *Brain and intelligence in vertebrates*. Clarendon Press.
- Macphail, E. M., & Bolhuis, J. J. (2001). The evolution of intelligence: Adaptive specializations versus general process. *Biological Reviews*, 76(03), 341–364.
- Marr, D. (1982). *Vision, a Computational Investigation into the Human Representation and Processing of Visual Information*. W. H: Freeman.
- Mata, A., Vaz, A., & Palma, T. A. (2021). Cue competition in mental state inference: Blocking effects in learning to interpret the behaviors of others. *Journal of Experimental Social Psychology*, 95, Article 104122.

- McNally, G. P., & Cole, S. (2006). Opioid receptors in the midbrain periaqueductal gray regulate prediction errors during Pavlovian fear conditioning. *Behavioral Neuroscience*, *120*, 313–323. <https://doi.org/10.1037/0735-7044.120.2.313>
- McNally, G. P., Johansen, J. P., & Blair, H. T. (2011). Placing prediction into the fear circuit. *Trends in Neurosciences*. <https://doi.org/10.1016/j.tins.2011.03.005>
- Millin, P., Feldman, L. B., Ramsar, M., Hendrix, P., & Baayen, R. H. (2017). *Discrimination in lexical decision*. *PLOS ONE*, *12*(2), e0171935.
- Miller, N. (2018). Social learning and associative processes: A synthesis. *Journal of Experimental Psychology: Animal Learning and Cognition*, *44*, 105–113. <https://doi.org/10.1037/xan0000167>
- Miller, N. Y., & Shettleworth, S. J. (2007). Learning about environmental geometry: An associative model. *Journal of Experimental Psychology: Animal Behavior Processes*, *33* (3), 191–212. <https://doi.org/10.1037/0097-7403.33.3.191>
- Miller, R. R., Barnet, R. C., & Grahame, N. J. (1995). Assessment of the Rescorla-Wagner model. *Psychological Bulletin*, *117*(3), 363–386.
- Mitchell, C. J., De Houwer, J., & Lovibond, P. F. (2009). The propositional nature of human associative learning. *Behavioral and Brain Sciences*, *32*(02), 183–198. <https://doi.org/10.1017/S0140525X09000855>
- Niv, Y. (2021). The primacy of behavioral research for understanding the brain. *Behavioral Neuroscience*, *135*(5), 601–609. <https://doi.org/10.1037/bne0000471>
- Nixon, J. S. (2020). Of mice and men: Speech sound acquisition as discriminative learning from prediction error, not just statistical tracking. *Cognition*, *197*, Article 104081. <https://doi.org/10.1016/j.cognition.2019.104081>
- Nixon, J. S., & Tomaschek, F. (2021). Prediction and error in early infant speech learning: A speech acquisition model. *Cognition*, *212*, Article 104697. <https://doi.org/10.1016/j.cognition.2021.104697>
- Olejarczuk, P., Kapatsinski, V., & Baayen, R. H. (2018). Distributional learning is error-driven: The role of surprise in the acquisition of phonetic categories. *Linguistics Vanguard*, *4*(s2), 20170020.
- Ottenheimer, D. J., Bari, B. A., Sutlief, E., Fraser, K. M., Kim, T. H., Richard, J. M., et al. (2020). A quantitative reward prediction error signal in the ventral pallidum. *Nature Neuroscience*, *23*(10), Article 10. <https://doi.org/10.1038/s41593-020-0688-5>
- Papini, M. R. (2002). Pattern and process in the evolution of learning. *Psychological Review*, *109*(1), 186–201.
- Papini, M. R., & Bitterman, M. E. (1990). The role of contingency in classical conditioning. *Psychological Review*, *97*(3), 396–403.
- Pavlov, I. P. (1927). *Conditioned Reflexes: An Investigation of the Physiological Activity of the Cerebral Cortex*. Oxford University Press.
- Pearce, J. M. (1987). A model for stimulus generalization in Pavlovian conditioning. *Psychological Review*, *94*(1), 61–73.
- Pearce, J. M. (2009). The 36th Sir Frederick Bartlett Lecture: An associative analysis of spatial learning. *The Quarterly Journal of Experimental Psychology*, *62*(9), 1665. <https://doi.org/10.1080/17470210902805589>
- Pearce, J. M., Graham, M., Good, M. A., Jones, P. M., & McGregor, A. (2006). Potentiation, overshadowing, and blocking of spatial learning based on the shape of the environment. *Journal of Experimental Psychology: Animal Behavior Processes*, *32*, 201–214. <https://doi.org/10.1037/0097-7403.32.3.201>
- 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.
- Ramscar, M., Dye, M., & McCauley, S. M. (2013). Error and Expectation in Language Learning: The Curious Absence of “Mouses” in Adult Speech. *Language*, *89*(4), 760–793.
- Rasmussen, A. (2020). Graded error signals in eyeblink conditioning. *Neurobiology of Learning and Memory*, *170*, Article 107023. <https://doi.org/10.1016/j.nlm.2019.04.011>
- Rasmussen, A., Zucca, R., Johansson, F., Jirenhed, D. A., & Hesselro, G. (2015). Purkinje cell activity during classical conditioning with different conditional stimuli explains central tenet of Rescorla-Wagner model. *Proceedings of the National Academy of Sciences*, *112*(45), 14060–14065. <https://doi.org/10.1073/pnas.1516986112>
- Rescorla, R. A. (1969). Pavlovian conditioned inhibition. *Psychological Bulletin*, *72*(2), 77–94. <https://doi.org/10.1037/h0027760>
- Rescorla, R. A., & Wagner, A. (1972). A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. *Classical Conditioning II: Current Research and Theory*, *2*, 64–99.
- Riley, D. A., & Langley, C. M. (1993). The logic of species comparisons. *Psychological Science*, *4*(3), 185–189.
- Rodrigo, T., Chamizo, V. D., McLaren, I. P. L., & Mackintosh, N. J. (1997). Blocking in the spatial domain. *Journal of Experimental Psychology: Animal Behavior Processes*, *23*, 110–118. <https://doi.org/10.1037/0097-7403.23.1.110>
- Roy, M., Shohamy, D., Daw, N., Jepma, M., Wimmer, G. E., & Wager, T. D. (2014). Representation of aversive prediction errors in the human periaqueductal gray. *Nature Neuroscience*, *17*(11), Article 11. <https://doi.org/10.1038/nn.3832>
- Rudy, J. W. (1974). Stimulus selection in animal conditioning and paired-associate learning: Variations in the associative process. *Journal of Verbal Learning and Verbal Behavior*, *13*(3), 282–296. [https://doi.org/10.1016/S0022-5371\(74\)80065-4](https://doi.org/10.1016/S0022-5371(74)80065-4)
- Saavedra, M. A. (1975). Pavlovian compound conditioning in the rabbit. *Learning & Motivation*, *6*, 314–326.
- Sadacca, B. F., Wied, H. M., Lopatina, N., Saini, G. K., Nemirovsky, D., & Schoenbaum, G. (2018). Orbitofrontal neurons signal sensory associations underlying model-based inference in a sensory preconditioning task. *eLife*, *7*, e30373.
- Sammot, C. (2011). Credit assignment. In C. Sammot, & G. I. Webb (Eds.), *Encyclopedia of Machine Learning* (pp. 238–242). Springer.
- Sanbonmatsu, D. M., Akimoto, S. A., & Gibson, B. D. (1994). Stereotype-based blocking in social explanation. *Personality and Social Psychology Bulletin*, *20*(1), 71–81. <https://doi.org/10.1177/0146167294201007>
- Scavio, M. J., & Gormezano, I. (1974). CS Intensity effects on rabbit nictitating membrane conditioning, extinction and generalization. *Pavlovian Journal of Biological Science*, *9*(1), 25–34. <https://doi.org/10.1007/BF03000500>
- Schultz, W., Dayan, P., & Montague, P. R. (1997). A neural substrate of prediction and reward. *Science*, *275*(5306), 1593–1599.
- Shanks, D. R. (1985). Forward and backward blocking in human contingency judgement. *The Quarterly Journal of Experimental Psychology*, *37B*(1), 1–21.
- Shanks, D. R. (1991). Categorization by a connectionist network. *Journal of Experimental Psychology: Learning Memory and Cognition*, *17*(3), 433–443.
- Shanks, D. R., & Dickinson, A. (1988). Associative accounts of causality judgment. In G. H. Bower (Ed.), *The Psychology of Learning and Motivation* (Vol. 21, pp. 229–261). Academic Press.
- Sharpe, M. J., Chang, C. Y., Liu, M. A., Batchelor, H. M., Mueller, L. E., Jones, J. L., Niv, Y., & Schoenbaum, G. (2017). Dopamine transients are sufficient and necessary for acquisition of model-based associations. *Nature Neuroscience*, *20*(5), Article 5. <https://doi.org/10.1038/nn.4538>
- Siegel, S., & Allan, L. G. (1996). The widespread influence of the Rescorla-Wagner model. *Psychonomic Bulletin & Review*, *3*(3), 314–321.
- Sokolov, E. N. (1963). *Perception and the Conditioned Reflex*. Pergamon.
- Soto, F. A. (2018). Contemporary associative learning theory predicts failures to obtain blocking: Comment on Maes et al. (2016). *Journal of Experimental Psychology: General*, *147*(4), 597–602. <https://doi.org/10.1037/xge0000341>
- Soto, F. A., Siow, J. Y. M., & Wasserman, E. A. (2012). View-invariance learning in object recognition by pigeons depends on error-driven associative learning processes. *Vision Research*, *62*, 148–161. <https://doi.org/10.1016/j.visres.2012.04.004>
- Soto, F. A., & Wasserman, E. A. (2010). Error-driven learning in visual categorization and object recognition: A common elements model. *Psychological Review*, *117*(2), 349–381.
- Soto, F. A., & Wasserman, E. A. (2012a). A category-overshadowing effect in pigeons: Support for the Common Elements Model of object categorization learning. *Journal of Experimental Psychology: Animal Behavior Processes*, *38*(3), 322–328. <https://doi.org/10.1037/a0028803>
- Soto, F. A., & Wasserman, E. A. (2012b). Visual object categorization in birds and primates: Integrating behavioral, neurobiological, and computational evidence within a “general process” framework. *Cognitive, Affective, & Behavioral Neuroscience*, *12*(1), 220–240. <https://doi.org/10.3758/s13415-011-0070-x>
- Steinberg, E. E., Keiflin, R., Boivin, J. R., Witten, I. B., Deisseroth, K., & Janak, P. H. (2013). A causal link between prediction errors, dopamine neurons and learning. *Nature Neuroscience*, *16*(7), Article 7. <https://doi.org/10.1038/nn.3413>
- Sutton, R. S., & Barto, A. G. (1981). Toward a modern theory of adaptive networks: Expectation and prediction. *Psychological Review*, *88*(2), 135–170.
- Sutton, R. S., & Barto, A. G. (1990). *Time-derivative models of Pavlovian reinforcement* (pp. 497–537). Learning and Computational Neuroscience: Foundations of Adaptive Networks.
- Sutton, R. S., & Barto, A. G. (1998). *Reinforcement learning: An introduction* (Vol. 1, Issue 1). MIT press Cambridge.
- Takahashi, Y. K., Batchelor, H. M., Liu, B., Khanna, A., Morales, M., & Schoenbaum, G. (2017). Dopamine Neurons Respond to Errors in the Prediction of Sensory Features of Expected Rewards. *Neuron*, *95*(6), 1395–1405.e3. <https://doi.org/10.1016/j.neuron.2017.08.025>
- Thompson, R. F., & Spencer, W. A. (1966). Habituation: A model phenomenon for the study of neuronal substrates of behavior. *Psychological Review*, *73*(1), 16–43. <https://doi.org/10.1037/h0022681>
- van Eck, N., & Waltman, L. (2010). Software survey: VOSviewer, a computer program for bibliometric mapping. *Scientometrics*, *84*(2), 523–538. <https://doi.org/10.1007/s11192-009-0146-3>
- Van Hamme, L. J., & Wasserman, E. A. (1994). Cue competition in causality judgments: The role of nonpresentation of compound stimulus elements. *Learning and Motivation*, *25*(2), 127–151.
- Vanhooymissen, T., & Van Overwalle, F. (2010). Me or not me as source of ingroup favoritism and outgroup derogation: A connectionist perspective. *Social Cognition*, *28* (1), 84–109. <https://doi.org/10.1521/soco.2010.28.1.84>
- Vogel, E. H., Castro, M. E., & Saavedra, M. A. (2004). Quantitative models of Pavlovian conditioning. *Brain Research Bulletin*, *63*(3), 173–202.
- Vogel, E. H., Glynn, J. Y., & Wagner, A. R. (2015). Cue competition effects in human causal learning. *The Quarterly Journal of Experimental Psychology*, *68*(12), 2327–2350. <https://doi.org/10.1080/17470218.2015.1014378>
- vom Saal, W., & Jenkins, H. M. (1970). Blocking the development of stimulus control. *Learning and Motivation*, *1*(1), 52–64. [https://doi.org/10.1016/0023-9690\(70\)90128-1](https://doi.org/10.1016/0023-9690(70)90128-1)
- Waelti, P., Dickinson, A., & Schultz, W. (2001). Dopamine responses comply with basic assumptions of formal learning theory. *Nature*, *412*(6842), 43–48.
- Wagner, A. R. (1969). Stimulus selection and a “modified continuity theory”. In G. H. Bower, & J. T. Spence (Eds.), *Psychology of Learning and Motivation* (Vol. 3, pp. 1–41). Academic Press. [https://doi.org/10.1016/S0079-7421\(08\)60396-5](https://doi.org/10.1016/S0079-7421(08)60396-5)
- Wagner, A. R. (1971). Elementary associations. In H. H. Kendler, & J. T. Spence (Eds.), *Essays in neobehaviorism: A memorial volume to Kenneth W. Spence* (pp. 187–213). Appleton-Century-Crofts.
- Wagner, A. R. (1981). SOP: A model of automatic memory processing in animal behavior. In N. E. Spear, & R. R. Miller (Eds.), *Information Processing in Animals: Memory Mechanisms* (pp. 5–47). Erlbaum.
- Wagner, A. R., Logan, F. A., Haberlandt, K., & Price, T. (1968). Stimulus selection in animal discrimination learning. *Journal of Experimental Psychology*, *76*(2), 171–180.
- Waldmann, M. R., Cheng, P. W., Hagmayer, Y., & Blaisdell, A. P. (2008). Causal learning in rats and humans: A minimal rational model. In N. Chater, & M. Oaksford (Eds.),

- The probabilistic mind: Prospects for Bayesian cognitive science* (pp. 453–484). Oxford: University Press.
- Walker, R. A., Wright, K. M., Jhou, T. C., & McDannald, M. A. (2020). The ventrolateral periaqueductal grey updates fear via positive prediction error. *European Journal of Neuroscience*, 51(3), 866–880. <https://doi.org/10.1111/ejn.14536>
- Wang, H., Johnson, T. R., & Zhang, J. (1999). UEcho: A model of uncertainty management in human abductive reasoning. *Proceedings of the 20th Annual Conference of the Cognitive Science Society*, 1113–1118. 10.4324/9781315782416-200.
- Wasserman, E. A. (1974). Stimulus-reinforcer predictiveness and selective discrimination learning in pigeons. *Journal of Experimental Psychology*, 103(2), 284–297.
- Wasserman, E. A., & Castro, L. (2022). The Rescorla-Wagner Model: The culmination of Hume's theory of causation. *Journal of Experimental Psychology: Animal Learning and Cognition*, 48, 315–320. <https://doi.org/10.1037/xan0000325>
- Witnauer, J. E., Urcelay, G. P., & Miller, R. R. (2014). The error in total error reduction. *Neurobiology of Learning and Memory*, 108, 119–135. <https://doi.org/10.1016/j.nlm.2013.07.018>
- Wynne, C. D. L. (1995). Reinforcement accounts for transitive inference performance. *Animal Learning & Behavior*, 23(2), 207–217.
- Yuille, A. (2006). Augmented Rescorla-Wagner and maximum likelihood estimation. *Advances in Neural Information Processing Systems*, 18, 1561.
- Zimmer-Hart, C. L., & Rescorla, R. A. (1974). Extinction of Pavlovian conditioned inhibition. *Journal of Comparative and Physiological Psychology*, 86(5), 837–845.