Invited contribution to a special issue of REVISTA DE HISTORIA DE LA PSICOLOGÍA devoted to the work of Bob Rescorla, edited by Francisco Javier Bandres Ponce, to appear 2021

Robert Rescorla: Time, Information and Contingency

C.R. Gallistel
Department of Psychology
and
Rutgers Center for Cognitive Science

252 7th Ave Unit 10D
New York, NY 10001
galliste@ruccs.rutgers.edu

Abstract

Rescorla's first theoretical and experimental papers on the truly random control (random, independent presentations of CSs and USs) showed that associative learning was driven by contingency, that is, by the information that events at one time provide about events located elsewhere in time. This discovery has revolutionary neurobiological and philosophical implications. The problem was that Rescorla was unable to derive a function that mapped conditional probabilities into contingencies. Rescorla and Wagner (1972) proposed a hugely influential model for explaining Rescorla's results, but their model ignored his earlier insights about time, temporal order, information and contingency in conditioning. Their paper pioneered an empirically indefensible treatment of time that has continued in associative theorizing down to the present day. A key to a more defensible approach to the cue competition problem (aka the temporal assignment of credit problem) in Pavlovian and instrumental conditioning is to measure the information that cues and responses provide about the wait for reinforcement and the information that reinforcement provides about the recency of a response.

KEYWORDS
associative learning
information theory
probability distributions
trial problem
retrospective contingency

My greatly esteemed former colleague, Bob Rescorla, was the finest experimentalist since Pavlov in the field of learning. I consider him among the very best in the history of experimental psychology and cognitive science. He was also a brilliant theorist, as witness the hugely influential paper co-authored with Alan Wagner (Rescorla & Wagner, 1972). However, he did not do further theory after that dazzling first effort, for reasons I will explain.

That a first-rate scientific mind had come on the scene was apparent in the paper Rescorla published in the *Psychological Review* the year after he finished his graduate work at the University of Pennsylvania and accepted an appointment as Assistant Professor at Yale. In this paper (Rescorla, 1967), he critiqued all 6 control procedures that had been used in Pavlovian conditioning experiments up to that time. As he explained in his opening paragraph, (p.73)

"The operations performed to establish Pavlovian conditioned reflexes require that the presentation of an unconditioned stimulus be contingent upon the occurrence of a conditioned stimulus. Students of conditioning have regarded this contingency between CS and US as vital to the definition of conditioning and have rejected changes in the organism not dependent upon this contingency (such as sensitization or pseudoconditioning) as not being "true" conditioning (i.e., associative)." [scare quotes on 'true' in the original]

Prominent among these control procedures were the *explicitly unpaired control* (sometimes misleadingly called the *random control*), in which USs are presented in the same session as the CSs but never close in time to the CSs and the *Backward conditioning control* in which the US is presented shortly before the CS. Rescorla pointed out that the first of these two "controls" "does not simply remove the contingency between CS and US; rather, it introduces instead a new contingency, such that the US cannot follow the CS for some minimum time interval. Instead of the CS being a signal for the US, it can become a signal for the absence of the US." (p.73)

About the backward conditioning "control:", he wrote:
Q1 "The relevance of this procedure rests upon the assumption that in Pavlovian conditioning not only the CS-US contingency but also their temporal order of presentation is important. It is not clear whether this should be taken as part of the definition of Pavlovian conditioning or as an empirical result." In any event, "The occurrence of the CS predicts a period *free from the US.*" [p. 73, italics in original]

He went on to point out that:
Q2: "...in order for there to be no contingency, the distributions must be such that CS occurrences do not predict the occurrence of USs at any time in the remainder of the session. If the CS predicts the occurrence of a US 30 minutes later in the session, an appropriate random control condition has not been achieved." (p. 74)

These considerations led him to propose the *truly random control* in which the CS and the US
Q3: "...are programmed entirely randomly and independently in such a way that some "pairings" of CS and US may occur by chance alone." (p. 73-74).

Only then does
Q4: " ...the CS provide *no information* about subsequent occurrences of the US." (p. 74, italics in the original)

He then went on to discuss at length possible objections to this control procedure. A common objection was that the chance pairings of the CS and US that are inherent in this control would themselves produce some conditioning. In response to this objection, he notes that
Q5: "It rests upon an assumption, often not made explicit, that the temporal *pairing* of CS and US is the sufficient condition for "true" Pavlovian conditioning." (p. 75, italics and scare quotes in original).
He goes on to explain why contingency is not reducible to temporal pairing. He shows that they make radically different predictions under many circumstances.

Then, he writes:
Q6: "The idea of contingency used here needs explication. By it we mean the degree of dependency which presentation of the US has upon prior presentation of the CS."

so far so good, but he continues:
Q7: "This is clearly a function of the relative proportion of US events which occur during or at some specified time following the CS.... These proportions can be stated in terms of the *probability* of a US occurring given the presence of a CS (or given that the CS occurred at some designated prior time), and the *probability* of a US occurring given the absence of the CS (cf. Prokasy, 1965). The dimension of contingency is then a function of these two probabilities; if Pavlovian conditioning is dependent upon the contingency between CS and US, it, too, will be a function of these two probabilities. However, no attempt is made here to specify a particular function which relates these two probabilities to a continuum of contingencies." [p. 76, italics mine]

I have quoted at length from this historically important paper, partly because the quotes illustrate Rescorla's acute critical intelligence, but mostly because the extent to which the quoted insights have been ignored for more than half a century is the beginning of a tragedy in the history of cognitive neuroscience.

The seeds of the tragedy are in Q7. Everything Rescorla wrote up to Q7 should have led to a revolution in theories of associative learning and in neuroscientific attempts to determine its neurobiological basis. What we have had instead are mostly versions of theories that are now half a century or more old and a search for the physical basis of memory based on conceptual foundations laid by a long line of empiricist philosophers stretching back millennia—from Locke and Hume to Acquinas and Aristotle (Gallistel, 2020).

What Rescorla stresses repeatedly prior to Q7 is the importance of asking what *information* the onset (or, I may add, the offset) of the CS provides about the *time* of occurrence of the US. He makes it clear that answering this question is the key to understanding the contingency that a protocol creates. His arguments imply that to understand contingency, we must turn to the theory of information (Shannon, 1948), because it makes information measurable. Before Shannon, it was not a scientifically useful concept, because: “In physical science a first essential step in the direction of learning any subject is to find principles of numerical reckoning and methods for practically measuring some quality connected with it. When you can measure what you are speaking about, and express it in numbers, you know something about it, when you cannot express it in numbers, your knowledge is of a meager and unsatisfactory kind; it may be the beginning of knowledge, but you have scarcely, in your thoughts advanced to the stage of science.” (Thomson, 1883, later Lord Kelvin, p. 72)

The revolution Rescorla's analysis called for has failed to happen because no practical, mathematically defensible attempt was made for the next 50+ years to relate "...these two *probabilities* to a continuum of contingencies."[italics mine] The attempts that were made were

unsuccessful, because, when the problem of measuring contingency is formulated in terms of finding a function that maps from *probabilities* to contingencies rather than from probability *distributions* to contingencies, it leads into a conceptual wilderness.

We have wandered in that conceptual wilderness because we have continued to think in terms of probabilities rather than in terms of the distributions of probabilities, and we have failed to turn to Shannon to learn how to measure information. Information is measured by computing the *entropies* and *relative entropies* of probability *distributions*. Probabilities have no units, but entropies do (bits or nats). Probability distributions have *parameters*. The parameter of a Bernoulli distribution does double duty, because the Bernoulli p is both the parameter and one of the two complementary probabilities in the probability vector; the other element of the Bernoulli probability vector is $q = 1 - p$.

A probability distribution is a 1-1 mapping between two vectors: One vector numerically represents the support, the things to which the probabilities or probability densities attach. The other vector gives the corresponding probabilities or probability densities. The Bernoulli support vector is [0 1], where 0 represents "failure" and 1 represents "success," while the corresponding probability vector is $[q\ p]$. Conceptually, the support vector for an exponential distribution of wait times consists of the intervals between 0 and infinity. In practice—when graphing it for example—it only contains a discrete sampling from this uncountably infinite set, namely a set of, say, 50 or 100 evenly spaced durations spanning the interval where the probability densities are noticeably above 0. The parameters of distributions, for example, their means and variances, are computed from the probability-weighted support vector and the probability-weighted vector of the squares of its elements (each support element or its square multiplied by the corresponding probability).

The uncertainty in a distribution is measured by its entropy. The entropy may be computed directly from the probability vector or from the value(s) of the distribution's parameter(s). The entropy measures the uncertainty about which element of the support vector will obtain on a future occasion (on a new "draw" from the distribution). The more uncertainty, the greater the *available information*, that is, the more one learns when the occasion arises again and one experiences another draw from the distribution (another success or failure or another wait for reinforcement). Put another way, the available information is the amount of information subjects has to acquire to reduce their uncertainty to 0. Once you the reinforcement arrives, you no longer have any uncertainty about the delay; you have passed from what will happen to what did happen. These facts about information and how it is measured are central to understanding the tragedy.

Although Rescorla's paper in the *Psychological Review* was rich in theoretically important insights, they were offered in pursuit of the goal of improving experimental procedures. This is consistent with what Bob told me decades later when we were colleagues. It came in a conversation that began when I asked him why he continued to use predictions from Hull's theory (Hull, 1952) to motivate his experiments, given that his experiments always showed that the predictions were wrong. He did not dispute my assertion that his experiments always showed Hull's theory to be wrong. He knew Hull's theory was wrong. By the time of this conversation, he was disappointed that so many psychologists and neuroscientists failed to realize that it was wrong (Rescorla, 1988). Bob said he drew on Hull's theory because it was so good at generating

experimentally testable predictions. He added that he was an experimentalist not a theorist—a remarkable thing to say by a co-author of the best-known theoretical paper in the field of associative learning. He said what he loved was conceiving well designed experiments and carrying them out. Among figures at his level of distinction, he was unusual in that he continued throughout his career to run his own experiments.

Rescorla (1968)

Unfortunately, in describing and theorizing about the extraordinarily important results of the experiment that made him well known and earned him early-career election to the National Academy of Sciences, Rescorla led us further into the wilderness. While still a graduate student, he did the first experiment using his truly random control. It was somewhat complicated so I pass on to the justly famous Rescorla (1968) paper, with its elegant methodological simplicity.

The operant phase of training. When rats have been taught to press a lever for food on a variable interval schedule they freeze when the CS comes on that warns of impending shock. Their freezing reduces the rate at which they press the lever. Rescorla used the reduction in their rate of pressing for food as his measure of the extent to which various Pavlovian conditioning protocols produced fear, that is, he used what is called a *conditioned-emotional-response* protocol. To do that, he first trained all his subjects to press a lever to obtain food on a *variable interval* (VI) schedule of food reinforcement.

A VI schedule sets up food rewards at intervals determined by a Poisson process. The scheduling process stops running when it sets up a reward. It restarts only when the subject presses the lever to harvest the reward. The subject gets no signal when a reward is been set up. Put another way, the subject gets no *information* about *when* a reward has been made available (Q4 and Gallistel, Craig, A., Shahan, 2019).

The steady moderate rates of responding that one sees when subjects respond on a VI schedule has long intrigued students of operant conditioning, which these days is often called reinforcement learning. This phenomenon has lessons to teach us about the fundamental role that information plays in conditioning.

When rewards are scheduled by a Poisson process, neither the time at which the last reward was obtained, nor that the fact that the well-trained subject has just made a response provide measurable information about how long the subject will have to wait until a response produces the next reward. A Poisson process produces an exponential distribution of interevent intervals. A unique property of the exponential distribution is that it has a flat hazard function This means that the chance that a reward is available and may be harvested by pressing the lever is the same at every moment after the harvesting of the last reward regardless of how much time has elapsed since that last reward.

Early in training, when the response rate is very low, there is a rate-dependent positive *prospective* contingency between the rate at which a subject presses the lever and the rate at which it obtains rewards: When the average interval between a subject's presses is much longer than the average interval at which the Poisson reward schedular sets up rewards, the subjects' rate of pressing largely determines the rate at which it gets rewards; the faster it presses, the more

rapidly it obtains rewards. This rate-rate contingency soon drives subjects to respond at intervals much shorter than the average interval at which the schedular sets up the rewards. When the response rate, λ_r, is much greater than the rate parameter of the schedule, λ_{VI}, the schedular determines the rate at which the subject obtains reward, that is, $\lambda_R \cong \lambda_{VI}$. When that state of affairs obtains, then the prospective contingency between response rate and reward rate, λ_R, becomes immeasurably small (Gallistel, et al, 2019). The variations in λ_r between reinforcements no longer produce variation in the inter-reinforcement intervals (hereafter, the R–>R intervals).

Moreover, when $\lambda_r >> \lambda_R \cong \lambda_{VI}$, the entropy of the distribution of response-reinforcement intervals (r–>R) is not measurably different from the entropy of the R–>R distribution (Gallistel, et al, 2019). When those two entropies do not differ, pressing the lever does not reduce the subject's uncertainty about when it will get the next reward. Put another way, the *prospective* contingency between r and R is essentially 0.

So, one may ask, why do subjects respond rapidly and steadily on variable interval schedules of reward if each press has no effect on the distribution of waits for reward? They do so because the *retrospective* contingency, the degree to which reward provides information about the events that precede it in time, is 1. The retrospective contingency is 1 because every reinforcement is immediately *preceded* by a response at a very short fixed r<–R interval. Therefore, the distribution of r<–R intervals has 0 entropy. This means that an R time provides the maximal attainable amount of information about how long ago the last r occurred (Gallistel, et al, 2019). One intuitively helpful way to think about why this might be important to a subject is that a high r<–R contingency is a signature of what Aristotle called efficient causality, the causes that make things actually happen.

In backward conditioning, when the USs are presented shortly before the CS onsets, the contingency between CS onset and the US is also retrospective, not prospective. Knowing the time at which a CS is expected to occur tells one the *earlier* time at which a US may be expected to occur. Rescorla called attention to the importance of distinguishing between prospective contingency and retrospective contingency in Pavlovian conditioning (Q1). In 1968, it had not been demonstrated that rats learn the retrospective contingencies in a Pavlovian conditioning protocol, as well as the prospective ones, although Rescorla clearly suspected that they did. It has subsequently been demonstrated that they do (Matzel, Held, & Miller, 1988). Thus, *forward* pairing of the CS and the US is no longer part of a knowledgeable researcher's definition of Pavlovian conditioning (cf Q1). Nonetheless, it remains rare to see this distinction made in contemporary theoretical treatments of conditioning (cf Dayan & Berridge, 2014; Schultz, 2015). Rescorla was far ahead of his time in his recognition of important distinctions. The prospective/retrospective distinction is fundamental to his insight that temporal contingency is about the information that an event provides or fails to provide about *at what remove in time* other events occur.

The Pavlovian phase of training. Returning to Rescorla's (1968) methods: When the rats had learned to press the lever steadily, the lever was removed and Pavlovian conditioning began. The CSs were 2-minute tones. Their onsets were scheduled by a Poisson process with a rate parameter, λ_{toneOn}, of 0.125 tone onset per minute. Thus, the average inter-trial interval, μ, was $\mu = 1/\lambda_{toneOn} = 1/0.125 \text{min}^{-1} = 8$ minutes. Because the distribution of intertrial intervals was

exponential, many intertrial intervals were much shorter than 8 minutes, while some were much longer. Because the exponential has a flat hazard function, when a tone went off, it was impossible to predict the time at which it would come back on. The offset of the tone provided no information about its next onset time; every future moment was as likely as any other.

The shocks were also scheduled by two Poisson processes. One ran during the CSs; the other during the USs. The two rate parameters, $\lambda(\mathrm{US}|\mathrm{CS})$ and $\lambda(\mathrm{US}|{\sim}\mathrm{CS}) \equiv \lambda(\mathrm{US}|\mathrm{ITI})$, varied between groups from as low as 0 for one group to as high as 0.125 shocks per minute for some others. Again—and Rescorla laid particular stress on this—when the shock rates during the CS and during the intertrial intervals were the same, that is, when $\lambda(\mathrm{US}|\mathrm{CS}) = \lambda(\mathrm{US}|{\sim}\mathrm{CS})$, *the "probability"*[1] *that a shock would occur at any moment was the same at every moment*. I put this in italics because theorists have ignored this fundamentally important fact for more than half a century.

The restoration phase. After several sessions of Pavlovian fear conditioning, the levers were returned to the boxes and the shocks ceased, never to occur again. For two sessions, the rats simply pressed the levers to again obtain food on the same VI schedule as in initial training. After these two restorative sessions with no tones, there were 6 test sessions during which the tones sounded at unpredictable intervals, 4 such tones per session. Rescorla measured the extent to which the rats had learned to fear shock during the CSs by the ratio between the rate at which they pressed the lever during a tone and the sum of that rate and the rate at which they pressed during the flanking intertrial intervals. When there is no behavioral evidence of fear, this ratio is .5; when the rat freezes completely throughout each tone, this ratio is 0.

The test phase. The 6 test sessions were extinction sessions, because the tones no longer predicted shock. The rats were expected to gradually lose their fearful reaction to the tone over days, as indeed they did—note the rise toward 0.5 in the curves plotted in Figure 1. That, however, is not what was revolutionary about these results. What was revolutionary was that the more closely the rate of shock during the intertrial intervals approached the rate during the CSs, the less evidence there was that the tone instilled fear, *even though the pairings of tone and shock were identical in every group within each panel*! The shocks during CSs not only occurred at the same rates in different groups within a panel, they occurred at *exactly the same times*, and likewise for the shocks that occurred during the inter-trial intervals. The rats in the groups for which the two rates were equal (one group in each panel) showed no significant evidence of fearing the tone even on the first day of testing. Thus, it was not the temporal pairings of tone and shock that led the rats to fear the tones, it was the contingency between the tones and the shocks, the fact that the tones provided information about the temporal distribution of the shocks.

---

[1] 'probability' is the word Rescorla used. He should have written hazard. A probability cannot be greater than 1, whereas a hazard may rise to infinity, as it does, for example, in the Gaussian hazard function, and in many others. A hazard is not a probability density either, because a probability density is the value at one point in a probability distribution with continuous support. A probability distribution, whether discrete or continuous, must sum or integrate to 1. Hazard functions do not. One could replace 'probability' with 'likelihood' because likelihoods are like hazards in that they live on the interval [0 ∞]. However, the support for a likelihood function is a distribution's parameter vector; whereas the hazard function takes as its argument the support for the distribution itself. This degree of sophistication about mathematical statistics was rare among students of associative learning in the 1960s, myself included. Moving into the territory where Rescorla's insights should have pointed us in 1972 would have required our mastering aspects of mathematical statistics that few of us had been taught.

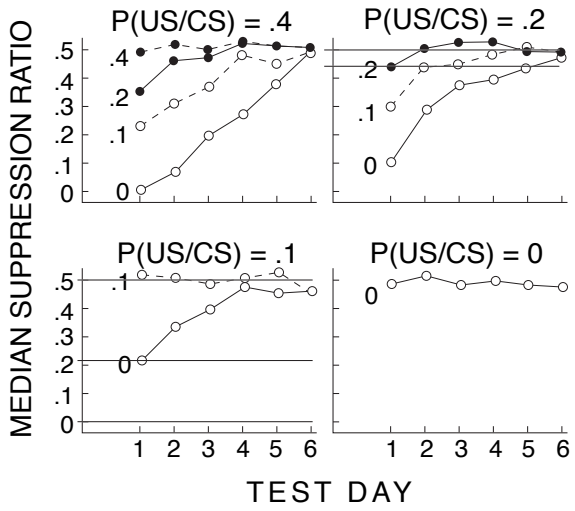

Figure 1. *Median suppression ratio for each of 10 groups of rats during the 6 successive test sessions in which their fear of the previously conditioned tone was measured. For each group, the "probability" of a US (shock) during the conditioning phase of the experiment, given that the tone was present, is at the top of the panel in which the results from that group are plotted. The conditional "probability" of a US given that the tone was not present is given at the left end of the plot for a given group. For explanation of the scare quotes on 'probability', see text and Figure 2. Redrawn from Figure 3 on p. 4 of (Rescorla, 1968) by permission of the publisher*.

The misleading description of the probabilities. The results in Figure 1 are justly famous because they revolutionized thinking about associative learning. However, they were—and still usually are—presented in a misleading way. What varied in this experiment was not the probability of reinforcement given CS or ~CS; what varied were the rates of reinforcement. One is invited to believe that the P(US|CS)'s at the tops of the panels denote the relative frequency with which *a* US occurred during *a* tone. But that is not in fact the quantities Rescorla varied. The quantities he varied were the probabilities that *one or more* USs occurred during a tone or during a CS. This is a necessary consequence of the fact that the US-scheduling processes were Poisson. During some CSs, there were as many as 3 USs and during some ~CSs (intertrial intervals) there were more than 3, while during others, there were none. Put more formally, the P(US}CS) = 0.4 above the upper left panel in Figure 1 should be written $P(n_{US}>0|CS) = .4$, and likewise for all the other conditional probabilities, both in the figure and in Rescorla's text.

I could write a separate paper on what I think explains the fact that Rescorla chose to present the data in terms of probability. I think part of the story was that he correctly judged that most of the intended readers were more comfortable with probabilities than with rates. Then, and even now, the contingencies in Pavlovian conditioning protocols are almost without exception specified as Bernoulli probabilities. This practice both encourages and reflects the view that in some sense contingency reduces to conditional probability (cf Schultz, 2015). That is not the case, as subsequent efforts soon showed (Hallam, Grahame, & Miller, 1992; Hammond, 1980; Hammond & Paynter, 1983). The fact that, generally speaking, conditional probabilities cannot be converted to contingencies is obliquely implied by the final sentence in Q7. It suggests that Rescorla had tried his hand at making the conversion and failed.

I know that Bob was aware of what I have just pointed out about what should have been written above the panels, because once during the years when we were colleagues at Penn, we had roughly the following conversation: It began by my asking, "In your 1968 experiment and in your paper with Wagner, you write only of the probability of reinforcement during the CS and during the ITI. However, you scheduled shocks with a Poisson process, right? And your CSs lasted 2 minutes, right? Therefore, it must be the case that during *some* CSs there were not one but two shocks. And, indeed, for *a few* CSs not one but three shocks, right? So, when you speak of the probability of reinforcement what you really refer to is the probability of *one or more* reinforcements, right?" Bob smiled broadly if somewhat sheepishly and replied, "Right." His broad but sheepish smile indicated to me that we both understood that 'probability of reinforcement', P(US|CS), does not mean the same thing as 'probability of one or more reinforcements', that is, P(US|CS) ≠ P(n_{US}>0|CS).

Rescorla & Wagner (1972)

Rescorla understood that the rats in the groups where the rates of shock were the same during the tones as during the intertrial intervals became conditioned to the experimental chamber rather than to the tone. Anyone who does this experiment will observe that the rats in this condition freeze when put in the chamber at the start of each session. At almost the same time, Kamen published his blocking and overshadowing experiments (Kamin, 1967; Kamin, 1969; Kamin, 1969) and Wagner and his collaborators published their relative validity experiment (Wagner, Logan, Haberlandt, & Price, 1968). These Pavlovian conditioning experiments with rats all showed that what a subject had already learned—in the blocking protocol—or was contemporaneously learning—in the overshadowing and relative validity protocols—strongly interacted with what it learned about a target CS. Earlier in the same decade, Reynolds (1961) showed what was later recognized as overshadowing in pigeons in a operant protocol. These stimulus-interaction findings in Pavlovian and operant experiments were revolutionary. Previous theories of associative learning had treated each CS-US association as an independently developing connection.

The theoretical problem was how to model/understand these interactions. In Pavlovian conditioning, this has come to be called the cue competition problem. In the operant, reinforcement learning and artificial intelligence literatures, it is called the assignment of credit problem (Dayan & Berridge, 2014; Fu & Anderson, 2008; Gallistel, et al, 2019; Minsky, 1961; Staddon & Zhang, 1991; Sutton, 1984).

Rescorla and Wagner (1972) joined forces to publish what continues to be a hugely influential solution. The solution is their differential equation

$$\Delta A_i = \alpha_{i,US}\left(\Lambda_{US} - \sum_{i=1}^{i=n} A_i\right) \qquad (1)$$

where ΔA_i is the change in associative strength of the i^{th} CS on a given trial; $\alpha_{i,US}$ is the associability parameter (aka the rate of learning parameter); $\sum_{i=1}^{i=n} A_i$ is the sum over all the associations between the CSs active on the given trial and the US; and $\Lambda_{i,US}$ is an upper limit on

that sum. The key idea is that the competition between CSs for associative strength is mediated by the upper limit on net associative strength (Λ_{US}).

In Rescorla and Wagner's thinking, the association between a CS and a US was a proxy for how well a CS predicted the US. If a CS predicts the US with probability 1, there is no room for any further increase in the predictability of that US on a trial on which that maximally predictive CS occurs. As Hullians (however temporary in Bob's case), they could not postulate that the brain contained a symbol for a probability. However, if one thinks in terms of synaptic plasticity, as Wagner certainly was thinking, then it is reasonable to imagine that when a single presynaptic spike suffices to fire the postsynaptic neuron, then further increases in synaptic conductance will have no behavioral consequences. Thus, in postulating an upper limit on net associative strength, they were thinking in terms of strength of prediction. Let A_1 be the strength of the CS_1-US association, where CS_1 is a CS that has repeatedly been paired with the US. If $A_1 = \Lambda_{US}$, then $\Lambda_{US} - \sum_{i=1}^{i=n} A_i = 0$. In that case, if a new CS_2 is introduced, as in a blocking protocol, $\Delta A_2 = 0$. Therefore, the strength of the CS_2–US association never grows. Thus, blocking is immediately explained by this formula. A similar but more complex analysis yields Wagner et al's (1968) relative validity results.

Rescorla and Wagner finished with an analysis of Rescorla's (1968) results. The intuition underlying their analysis was that in the truly random control, where USs occurred during the ~CSs (the intertrial intervals), the strength of the association between the *context* (the only CS present during the intertrial intervals) and the shock grows faster than the strength of the association between the tone and the shock. If that is so, then the association to the context will crowd out the association to the tone; the context-US will take up all the possible net associative strength. This intuition accounts for the enduring popularity of the theory. The intuition, however, overlooks three conceptual difficulties that arise when one attempts to apply equation (1) to Rescorla (1968).

One difficulty is the number of free parameters. The subscripts on $\alpha_{i,US}$ indicate that associability (rate of learning) may be unique to each different combination of a CS and a US. That means the theory has as many free parameters as there are CS-US combinations. Free parameters are parameters with unknown values. Their values must be estimated from the data when assessing whether a formula like Equation (1)—aka a model—explains the results of an experiment. The more free parameters a model has, the more data it can "predict." 'Predict' is in scare quotes because when the parameters have to be estimated from the results of each new experiment in order to make the model predict those results, the model does not *pre*dict those results. von Neumann famously said, "With 4 parameters I can fit an elephant, with 5, I can make it wiggle its trunk." Statisticians understand that free parameters give a model room to wiggle. With enough of them, a model can be made to predict random data, which is to say all possible results (Rissanen, 1989, 1999).

In applying their model of associative learning to their respective experimental results, Rescorla and Wagner took advantage of the wiggle room: They assumed one set of values in order to explain Rescorla's (1968) results and a very different set of values to explain the results obtained by Wagner et al (1968)—see Gallistel (1990, pp 412-417). They made these very different

parametric assumptions despite the fact that both experiments used tone and light CSs lasting 2 or three minutes in conditioned emotional response protocols with rat subjects.

The biggest problem with their paper from my perspective, however, is with the treatment of time and contingency. In his paper with Wagner, Rescorla temporarily abandoned the insights he had achieved while still a graduate student. Their paper does not mention contingency; it does not mention information; time does not appear in Equation (1); and their treatment of time led the field further into the wilderness.

For me the greatest element of the tragedy is that in trying to make their theory as close as possible to the Hullian theory that preceded it, they encouraged the field of animal learning and neurobiologists interested in the physical basis of memory to believe that their theory rescued Hullian theory. As a result, the neurobiological community continues to use Hull's theory as the basis for their attempts to discover the engram, the physical basis of memory. The use of this theory as the conceptual foundation for the search for the engram explains why we have still not found it (Gallistel, 2020).

Rescorla and Wagner made a nod toward the nascent movement toward a cognitive psychology, a psychology in which expectations were legitimate postulates. "The central notion suggested here can also be phrased in somewhat more cognitive terms. ....organisms only learn when events violate their expectations. Certain expectations are built up about the event following a stimulus complex; expectations initiated by that complex and its component stimuli are then only modified when consequent events disagree with the composite expectation." --.p. 75. However, they then turned to Hull's theory, with which Wagner was more comfortable[2]. In the paragraph that introduced Equation (1), they wrote "...one way to look at the central notion of this theory is as a modification of Hull's account of the growth of $_{S}H_{R}$. Similarly, one way to view the particular formalization to be proposed is as a modification of the mathematical model most closely related to the Hullian theory, the linear model" -p. 75

In Hullian theory, there are no memories as memories are intuitively understood (facts stored in the brain). A fortiori, there are no records in memory of the many abstract facts about the experienced world that are only very obliquely related to sensory experience—facts such as times, distances, directions, durations, numerosities and probabilities. There is no information in the mind in Hull's theory; there are only conditioned reflexes. The H in $_{S}H_{R}$ stands for habit and 'habit' is another name for conditioned reflexes (Hull, 1930). Hull's concept of a habit is the ancestor of what contemporary reinforcement learning theorists call model-free learning.

In the Hullian theory of memory, there are only the associative bonds that enable stimuli (S) to excite responses or inhibit (R), just as in the neurobiologists' conception of memory, there are only the plastic synapses that enable a presynaptic neuron to excite or inhibit a postsynaptic neuron. This was the message that Hull intended neurobiologists to take from his work. Unlike Skinner, who thought psychologists should leave neurobiology to the neurobiologists, Hull wanted to make his theory of learning neurobiologically plausible. Hull also wanted a radically empiricist theory, a theory that honors the commitment in medieval philosophy to Aristotle's principle that "there is nothing in the mind that was not first in the senses" (Acquinas, 1256-1259, q. 2, a. 3, arg 19).

[2] I know this because I took Wagner's advanced course in learning while a graduate student at Yale in 1965.

The commitments to plastic synapses and to a radical empiricism stand in profound antagonism to Rescorla's insight that temporal information and contingency drive associative learning, not temporal pairing. They do so because time of occurrence, duration, information and contingency are on the long list of intangible abstractions that inform everyday behavior (Gallistel, 2020). They are numerically representable facts acquired from previous experience and stored in memory. The problem for the radical empiricist is that there are no sensory receptors for these aspects of the physical world. If time, information and contingency get into the mind, they do not get there by the route that Aristotle, the peripatetic philosophers, Acquinas, John Locke and Clark Hull thought was the only possible route. Also, these abstract quantities that inform our remembered experiences are not the sort of thing that can be encoded in plastic synapses—unless the mechanisms that make synapses plastic have heretofore unpostulated properties. Indeed, the suggestion that synapses might *encode* anything is not to be found in neurobiological reviews of progress in the search for the engram (cf Poo et al., 2016; Tomonori, Duszkiewicz, & Morris, 2013).

Time and the NoUS in Rescorla-Wagner

In Pavlovian protocols, the target CSs (transient tones, lights, noises, vibrations) are always presented in a context, that is, in a chamber into which the subject is placed at the beginning of each experimental session. As Rescorla realized in consequence of his 1968 experiment, this context is itself a CS; the subjects learn that they get food or water or shock (the usual reinforcement suspects) in that box. Therefore, in the traditional "experimental" condition, where the shocks occur only during the target CS (or, more typically, only at the end of it), both the target CS and the box CS are present. However, as Rescorla's experiment showed, the association between the box and shock is also critical to the behavior that develops in response to the tone CS. In thinking about how to bring the box into play—something no one had previously thought about—they had to consider both the relative frequencies of the tone+box trials and the box-alone trials.

Until Rescorla and Wagner tried to apply their theory to Rescorla's (1968) results, I think it is fair to say that no one had devoted much thought to what constitutes a trial. The events that constituted trials in the minds of researchers up to that time were always demarcated by observable events acting on sensory receptors. A trial was taken to be the interval between the onset and offset of whatever stimulus was used as the CS (or during the interval in which the trace left by the CS in the nervous system faded away).

Given what had always defined a trial, it would seem simple to say what constituted a box-alone trial: A tone+box trial began when the tone came on and ended when it went off; therefore, one would think that a box-alone trial began when the tone went off and ended when it came on. One might also think that in both cases, the probability of a shock during a trial was the total number of shocks that occurred during trials of that kind divided by the number of trials of that kind.

There were 12 CSs in each of the 5 Pavlovian conditioning sessions, so 60 CS trials. Because CS and ~CS trials (tone+box and box-alone trials) alternate, there were also 60 ~CS trials (intertrial intervals). The average number of shocks that occurred during each trial is the shock rate during that kind of trial (shocks/minute) times the cumulative duration of those trials. In the P(US|CS) =

P(US|~CS)= 0.4 group, the shock rate was 0.26 shocks/minute. That rate times 60 gives 31.2 shocks in 60 CS trials for a probability of shock on a CS trial of .52. Notice that this probability does not agree with P(US|CS) = 0.4. Why not? Because, as already mentioned, some CS trials had not 1 but 2 shocks, and others even had 3. For every CS trial that got more than the average number of shocks there had to be other trial that got less than the average. It's the 60% of the trials that got 0 shocks that explain why P(US|CS) equals only 0.4.

The cumulative duration of the box-alone trials was 8x60 = 480 minutes. That cumulative duration multiplied by 0.26 shocks/minute gives 124.8 shocks. That number of shocks divided by the 60 box-alone trials yields a shock probability on box-alone trials of more than 2. Houston, we've got a problem: Probabilities cannot be more than 1. Again, the explanation is that more than one shock occurred on the great majority of box-alone trials (Figure 2); there was a 30% chance of a subject receiving more than 5 shocks during at least one of the 60 intertrial intervals. Less than 13% of the those intervals were shock free.

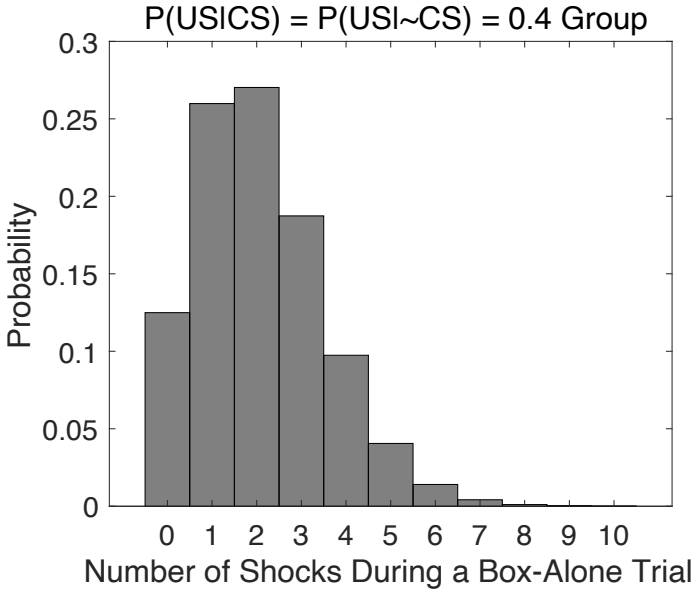

Figure 2. *The distribution of the number of shocks received during box-alone trials for subjects in one of the Rescorla's (1968) experiment when the box-alone trials are taken to be the observable intertrial intervals between CS offs and CS ons*.

Clearly, we cannot compute the probability of *a* shock on a box-alone trial in the usual way *if we define a trial in the usual way*, that is, as an interval demarcated by observable events. In fact, the way one calculates probability in a truly random protocol cannot be done the usual way at all, as witness the discrepancies between Rescorla's (1968) probabilities in Figure 1 and the results just obtained with the usual approach. To get Rescorla's probabilities, one has to first compute the rates at which his Poisson processes scheduled shocks, then use the Poisson distribution to compute the probability of 0 USs given both the shock rates (USs/unit time) and the duration(s) of the trials in question.

Critically, both the computation of the rate and the Poisson part of the computation require one to take into account the duration of the trial in question. Traditionally, researchers did not take

trial duration into account when computing reinforcement probability—for the simple reason that it is not clear how to do so. Probabilities do not have units; whereas rates and durations have temporal units. How is one going to get a unitless quantity on the left side of a formula when the quantities on the right side have units that do not cancel?

To make their theory work in explaining Rescorla's (1968) results, Rescorla and Wagner had to find a way of computing P(US) during box-alone trials that made it comparable to P(US) during tone+box trials. Here's what they did:

"In order to exemplify the application of the model to [the Rescorla's 1968 results], the experimental session was taken to be divisible into time segments the length of the CS duration. Each segment containing the CS is thus treated as a [tone+box] "trial" and each segment not containing the [tone] as a [box-alone] "trial."' (Rescorla & Wagner, 1972, -p. 88; scare quotes in the original)

Decades ago when first I read the just quoted paragraph, I muttered to myself, "How crazy is that?!" I took their scare quotes to indicate they knew it was pretty crazy.

The first problem is that their description of what they did cannot be both complete and accurate. The exponential distribution of the intertrial intervals makes it impossible to carve the time line into a sequence of contiguous 2-minute intervals such that some contain only the CS and the others only the ITIs. There must be many 2-minute intervals that contain some part of a CS and some part of an ITI. In fact, there were a fair number of intertrial intervals that contained CS segments at both ends because they were much shorter than 2 minutes. To be complete, their description would have to say what they did with these *mixed* intervals. I never thought to ask Bob that question.

The second problem was that they posited imaginary events during these imaginary intervals. Rescorla's Poisson scheduling of shocks made it impossible in principle to say when the postulated events should be imagined to occur. The imaginary events in question are sometimes called NoUSs. A NoUS is the failure of a US to occur *at the time it is expected*.

When USs are scheduled by a Poisson process, there is no such time, as Rescorla stressed in 1968. A neurobiological neo-Hullian has to believe that NoUSs are physical causes in the brain, because they are assumed to cause an enduring depotentiation of synaptic conductances. Physical realized events are localized in time. The question is, How can a NoUS cause a change in synaptic conductance at a specifiable time in those cases where it is in principle impossible to say when a reinforcement is expected to occur hence impossible to say when a NoUS has occurred?

There was, however, a third problem—the problem at the core of the tragedy. Rescorla and Wagner were trying to salvage the most conceptually unproblematic kind of temporal pairing, temporal *coincidence*, as the causal event in associative learning. Defining temporal pairing as temporal coincidence has the advantage that one does not have to specify a *critical interval* within which the CS and US—or the response and the reinforcement—must both occur in order to become associated (Gluck & Thompson, 1987; Hawkins & Kandel, 1984). The critical-interval approach to temporal pairing is conceptually problematic because the field has never been able to determine from experiment what that critical interval might be (Rescorla, 1972).

Gallistel and Gibbon (2000) explain that this failure is traceable to the fact that there is no critical interval (Gibbon & Balsam, 1981). Many years later, Rescorla lamented the failure of psychologists to understand what he had written back in 1967 and 1968: "The insufficiency of contiguity for producing Pavlovian conditioning can be illustrated by results that have been available for almost 20 years (e.g., Rescorla, 1968) but that have apparently failed to be integrated into the view of conditioning held by many psychologists." -p.3).

Part of the tragedy is that Rescorla, when teamed with Wagner, tried to build a model based on temporal coincidence even though: 1) Rescorla had explained at length in his 1967 *Psychological Review* paper why this was unlikely to be a successful approach given what was already known about inhibitory conditioning. 2) He had shown in his 1968 experimental paper that it was contingency not temporal pairing that drives the associative learning in excitatory Pavlovian conditioning. 3) They were trying to salvage this idea by imagining trials that had no basis in observable fact, which is why they put scare quotes on 'trials'.

They were forced into this line of theorizing because contingencies *when defined in terms of conditional probabilities* cannot be computed without defining timeless trials. A trial has always been an interval whose duration need not be considered, during which a CSs and USs either do or do not occur. If they both occur within the trial, then an associative bond is strengthened. If the CS occurs and the US does not, whatever bond may exist between them is weakened. When trials are defined in terms of observable events this may not be a problem, at least for the theory of excitatory conditioning. However, Rescorla (1967) was not the first to point out that it has always been a problem when it comes to inhibitory conditioning and extinction (Gleitman, Nachmias, & Neisser, 1954; Kimble, 1961, 318-329).

When theorists allow themselves to imagine trials that have no basis in observable experience, during which imagined events also occur, conceptual problems arise. They arise from the failure to deal with time itself—that philosophically puzzling aspect of our remembered experience that does not come into the brain through the senses. The two problems that arise are: 1) How long do the imagined trials last? 2) When during the imagined intervals do the imagined events occur? The failures of CSs to occur and the failures of USs to occur are both among the events that are imagined to occur. These imagined failures are posited to cause changes at the synaptic level in the brain, so there must be some time at which they have this physical effect, some specifiable time at which the synapses change.

Rescorla and Wagner did not address either question, although both were aware that they lurked. Wagner later tried to extend the theory to cope with them (Vogel, Ponce, & Wagner, 2019; Wagner, 1981). The resulting theory became so rococo that Rescorla no longer wanted to follow where Wagner led. He remarked in a seminar that Wagner's theory was the only theory he knew that was harder to understand and remember than the data it explained. In that same seminar, I trolled Bob about the imagined trials and the imagined events. He did not bridle at my trolling; he freely acknowledged that they were problematic. Indeed, in the same year as the Rescorla-Wagner paper appeared, Rescorla published another chapter, which is unfortunately much less often cited, in which he wrote, "...to be sure, both sequential and temporal variables are important in conditioning and will demand adequate theoretical treatment. But the present data...do not serve as a solid base

for expansion of the theory [to include an explicit treatment of the effects of temporal variables]."
(Rescorla, 1972, p. 40)

Bob and I were on excellent collegial terms. I believe his reluctance to make further theoretical efforts, despite the success of the Rescorla-Wagner theory, arose partly from his inability to see a way to solve the problems posed by the roles of time and information in conditioning and the relations between events separated in time, roles he continued to stress (Rescorla, 1988).

The Enduring Problem with Imagined Events, States and Stimuli

Broadly speaking, the field has remained more comfortable with imagined intervals, imagined states and imagined stimuli than Rescorla was. They play a prominent role in contemporary theorizing (see Dayan & Berridge, 2014; Gershman, Moustafa, & Ludvig, 2014; Kehoe, Ludvig, & Sutton, 2013; Ludvig, Sutton, & Kehoe, 2008; Ludvig, Sutton, & Kehoe, 2012; see Luzardo, Alonso & Mondragón, 2017 for a critical review; for a sampling of the role of imagined trials, states and temporal stimuli, see ; Niv, Daw, & Dayan, 2005;Niv, Daw, Joel, & Dayan, 2007; Schultz, 2015; Schultz, Dayan, & Montague, 1997; Starkwe ather, Babayan, Uchida, & Gershman, 2017; Williams, Todd, Chubala, & Ludvig, 2017). Most of this theorizing is neurobiologically oriented. Like Hull, these theorists care deeply about neurobiological interpretations of their hypothetical constructs; they often spend more time reviewing the neurobiological data than the behavioral data.

Contemporary computational cognitive neuroscientists postulate sequences for brain states and temporal micro-stimuli for the same reasons that Rescorla and Wagner did: the neurobiological representation of these states create something to which a scalar quantity may be associated/connected. Sometimes that quantity is assumed to be an associative strength. In reinforcement learning, it is a "value" (Niv, 2009). In temporal difference learning, it is a prediction error. This theorizing is a continuation of Hull's thinking, couched in more cognitive language. The roles played by the value of a state in reinforcement learning theories and by the prediction error in temporal difference learning are similar to the role played by anticipatory goal stimuli (s_g) in Hull's theory (Hull, 1952). As in Hull's theory, values and prediction errors work their way back through a chain of states or temporal micro-stimuli. In Hull, the elements in the chain are the r_g's, whose sensory consequences are the s_g's (faint copies of stimuli experienced in the goal state). Value and prediction error work their way back through the chain trial by trial, just as do the r_g's in Hull's theory, so as to guide earlier responding in a chain of actions, just as do the s_g's in Hull's theory. Psychologists went down that theoretical path almost a century ago.

The problem with this theorizing is not that it postulates unobserved entities in the brain. That was Skinner's problem; as a logical positivist, he hated hypothetical constructs (Skinner, 1950). Logical positivism died as a guide to how to do science because the history of science shows that postulating entities not yet observed is essential in any reductionist approach. The gene is a classic example of a hypothetical construct. If the geneticists had not posited its existence, molecular biology would not exist. Particles are another example; their postulation and subsequent discovery has played a fundamental role in physics over the last 100+ years.

The problem with this line of theorizing is that it has been tried and failed. It has failed because it has not formulated a coherent theory of how time and information are represented in brains. The

following quote from a 100+ page review by a leading researcher on the neurobiology of learning is indicative of contemporary thinking (Schultz, 2015, Section 3, p. 17):

"The formal treatment of surprise in conditioning employs the concept of prediction error. A reward prediction error PE is the difference between received reward λ and reward prediction V in trial t

$$PE(t) = \lambda(t) - V(t) \quad (1)$$

This formal definition of "error" extends beyond the colloquial meaning of inaccurate behavior. Prediction errors occur whenever Equation 1 applies, irrespective of whether they simply occur during behavioral learning or are actually being used for learning. Animal learning theory aims to explain the role of contingency in conditioning by formalizing how prediction errors update reward predictions from previously experienced rewards (460). The new prediction V for trial t+1 derives from the current prediction V(t) and the prediction error PE(t) weighted by learning rate

$$\alpha\ V(t+1) = V(t) + \alpha * PE(t) \quad (2)$$

Note that V captures the sum of predictions if several stimuli are present."

The enduring influence of the Rescorla-Wagner paper—citation 460 in the quote—is clear. I note in particular that *t* refers to trial, not time. Schultz treats contiguity and contingency as equally important factors in learning, but he fails to define either of them. The propagation for half a century of conceptual errors that Rescorla clearly recognized in 1967 is made possible by the role that various circumlocutions for imaginary trials (aka states or steps) play in temporal difference learning and reinforcement learning.

Gibbon, Berryman and Thompson (1974) reviewed the measures of contingency in conventional (frequentist) statistics that might be applied to conditioning protocols. They stressed that none of them could be applied to instrumental/operant protocols. There is some irony here because temporal difference theories and reinforcement learning models generally attempt to explain results from those protocols (e.g., Niv, et al 2005).

A contingency table contains the four joint "events" that can be regarded as in some sense occurring or not occurring in a protocol with objectively definable trials. On a trial demarcated by observable events, the CS and US may both occur, they may both fail to occur, the CS may occur and not the US, or the US may occur and not the CS. The statistic most often mentioned in the more recent literature is one of the ones discussed by Gibbon et al (1974)

$$\Delta p = p(\mathrm{US}|\mathrm{CS}) - p(\mathrm{US}|{\sim}\mathrm{CS}).$$

I believe the problem with Δp was already apparent to Rescorla. I believe it's why he did not offer a formulation of contingency in 1967 or at any later time. Be that as it may, the problem with Δp soon became apparent to experimentalists who attempted to apply Rescorla's theoretical and empirical demonstrations of the importance of contingency to autoshaping and operant protocols (Gibbon, 1981; Gibbon et al., 1974; Gibbon, Locurto, & Terrace, 1975; Hallam et al., 1992; Hammond, 1980). It is not, however, apparent to many contemporary researchers, as witness the following quote from an anonymous reviewer writing in 2019:
"If the total time is divided into small enough equivalent units (the length of the trial is often used), all 2x2 contingency table values are perfectly measurable individually."

I believe this to be a common belief. There are three reasons why it is a false belief: First, to compute Δp one has to count events that didn't occur, namely, the ~CS's. I leave it to the reader

to try to count how many ~CS's there were in Rescorla's (1968) protocol; Rescorla was not so foolish as to try.

The second problem is subtler. It highlights the conceptual problems that follow from specifying trials and states that have no specified duration and no basis in observable fact. Suppose, we allow the durations of our imagined trials to become arbitrarily short. The anonymous reviewer just quoted thinks this solves the Δp problem. It doesn't. In fact, it highlights the problem.

Suppose that we divide the time line in Rescorla's (1968) protocols into some number of imagined contiguous intervals (imagined "trials") for the purpose of computing a Δp. Let *a* be the count of the intervals that contain a CS and a US, *b* be the count of intervals that contain a ~CS and a US, *c* be the count of the intervals that contain a CS and a ~US, and *d* the count of the intervals that contain a ~CS and a ~US. These are the mutually exclusive and exhaustive possibilities that go into a timeless binary contingency table. Then:

$$\triangle p = \text{p(US|CS)} - \text{p(US|}\sim\text{CS)} = \text{a/(a+c)} - \text{b/(b+d)}$$

Consider next the consequences of allowing the durations of the imagined trials to become arbitrarily short. Rescorla's USs lasted only 0.5 a second. He could have made them much shorter still; a rat fears a nasty shock no matter how short, because what it fears is the onset of the pain in its paws. That's a point event. Therefore, *a and b*, which are, respectively, the count of the intervals in which a US and a CS both appear and the count of the intervals in which a ~CS and a US both appear, do not change; they always add up to the number of USs that actually occurred. The problem is that *c* and *d*, which are, respectively, the count of the intervals in which a CS and a ~US appear and the count of the intervals in which a ~CS and a ~US appear, go to infinity as the durations of our imaginary trials become infinitesimal. These latter two counts, the ones that go to infinity, appear in the denominators of the conditional probabilities in the formula for Δp. Therefore, both conditional probabilities go to 0 as the durations of our imagined trials go to 0—and so does their difference, Δp.

In short, the duration of the imaginary interval within which events are counted determines the value of Δp. When the duration is left unspecified or specified in an arbitrary way, then either Δp is unspecified or its value is arbitrarily determined by the arbitrary choice of a duration for the imagined trials, as those experimentalis who attempted to follow Rescorla's lead (Hallam et al., 1992; Hammond, 1980).

I surmise that Rescorla knew this as early as 1967 when he was careful to write than he did not attempt to specify a formula for *temporal* contingency given *probabilities*. The tragedy is that so few others pondered what such a formula might look like and, if they could not devise one, what other approaches there might be to quantifying *temporal* contingency, given that the atemporal formula cannot be made to work.

Available Information, Transmitted Information and Temporal Contingency

When it is realized that numerosities (event counts), durations, times of occurrence and probability distributions are represented in brains, it is not hard to formulate a mathematically explicit parameter-free model that explains the cue competition phenomena that the Rescorla

Wagner model is erroneously thought to explain (Gallistel, 1990). The key is to focus on rates (numerosities divided by durations) not probabilities (subset numerosities divided by superset numerosities). A parameter-free matrix equation,

$$\lambda_{\mathrm{c}} = \lambda_{\mathrm{r}}\mathbf{T}^{-1} \qquad (2)$$

takes the place of the Rescorla-Wagner formula (Equation 1). The inputs to this formula are counts of observable events and measurable durations (Gallistel & Gibbon, 2000; Wilkes & Gallistel, 2016, 2017). These actually observed quantities determine λ_{r}, a row vector whose elements are the raw estimates of reinforcement rates (cumulative counts of USs during a CS divided by the cumulative duration of that CS. Different observed states of the world are defined by the different combinations of CSs that occur in a protocol. The ratios of the cumulative durations of those observed states determine the entries in $\mathbf{T}$, the temporal coefficient matrix. The product of the raw rate vector and the inverse of the temporal coefficient matrix gives the true (corrected) rates of reinforcement to be ascribed to the different CSs (e.g., the tone+box state and the box-alone state). As explained below, the rates that figure in this simple formula determine the information that the various the observed changes in state transmit to the subjects. Because it has no free parameters and no imagined states, this model has no wiggle room.

It is also not hard to formulate an information-theoretic measure of temporal contingency based only on measurable intervals and the counts of observable events (Balsam, Drew, & Gallistel, 2010; Gallistel,et al, 2019). The information available to a subject placed in a box where shocks occur at random is the entropy, H_{S}, of the exponential distribution of the inter-shock intervals:

$$H_{\mathrm{S}} = 1 - \ln\lambda_{\mathrm{S}}, \qquad (3)$$

where λ_{S} is the rate parameter of the exponential distribution of inter-shock intervals.

The formula for computing the entropy of an exponential distribution from its rate parameter (Equation 3) must be used with caution, because rates have units, as do means and variances. The choice of a temporal unit determines the value for entropy that one obtains from the formula. Unlike with means and variances, however, the temporal unit does not attach to H_{S}, the value for the entropy. The failure of the units on the right of Equation 3 to appear on the left makes it is easy to overlook this important fact. Moreover, for some choices of a temporal unit, the Equation (3) yields nonsense. For example, if one chooses a time unit such that $\lambda_{\mathrm{S}} > \mathrm{e}$, it yields negative entropies.

Entropy is always positive because it measures uncertainty, and there is no such thing as negative uncertainty. In information theory, information and uncertainty are two names for the same quantity. That takes some getting used to, because, intuitively, they are antithetical. When the unit of time is made arbitrarily small—in technical language, as it tends to 0—the entropy of a continuous distribution like the exponential distribution becomes arbitrarily large. This happens for a reason that psychologists should love: the upper limit on the uncertainty a

subject can have about the value of a stochastic variable depends on the number of possible values for it that they can distinguish. As that number goes to infinity, so does their uncertainty.

Fortunately, for actual subjects, their ability to distinguish differences in quantities such as weights, sound intensities, brightnesses, numerosities, durations, distances and directions is limited. Weber measured some of those limits at the dawn of experimental psychology. He found that they scaled with the reference quantity. We now know that, for every quantity that the mind reckons with, there is a Weber constant, w, such that, for every reference quantity, Q_r, we must increase or decrease Q by wQ_r in order to find a quantity that a subject can just distinguish as bigger or smaller than Q_r. Put more simply, the change in a Q required to make a *just noticeable difference* is proportional to Q. The constant of proportionality is w, and w is an inverse measure of a subject's ability to resolve differences; the smaller w is, the more differences a subject can distinguish within a given range.

The w's for duration and numerosity have been repeatedly determined. They generally fall in the range 0.1-0.3 for rats, mice, pigeons and humans (Gibbon, Malapani, Dale, & Gallistel, 1997). Knowing the w for duration and one other very general fact enables us to estimate how much uncertainty brains have about the quantities on which their computational mechanisms operate. The other general fact is that neurobiological measurement mechanisms, like all physically realizable measurement mechanisms, have limited dynamic range. Therefore, these quantity-sensing mechanisms autoscale their sensitivity. By autoscaling their sensitivity they put the current Qs within the range to which the measuring mechanism is currently sensitive, thereby maximizing the information the sensing mechanism delivers to the rest of the brain (Bialek & Rieke, 1992; Brenner, Bialek, & de Ruyter van Steveninck, 2000; Rieke, Warland, de Ruyter van Steveninck, & Bialek, 1997). Quantities that are too great or too small—quantities outside the dynamic range—cannot be measured until the range is again adjusted. Thus, the number of distinguishable quantities—the number of recognizable differences—tiles the dynamic range. Therefore, this number may be estimated from w.

If we assume that subjects in experiments that measure w autoscale so as to put Q_r in the middle of their dynamic range of the mechanism that measure Q, then a Weber fraction of 0.125 implies that there are 8 discriminable values for Q below Q_r and 8 above it, so 16 in all. On the other hand, a subject whose w = .25 can recognize only 8 different durations within any given dynamic range. Thus, the temporal information available to a subject with w = .125 is 4 bits ($2^4 = 16$), while the information available to the less discriminating subject is 3 bits.

The temporal information available to a subject in the absence of an information bearing signal measures the subject's uncertainty. The rats in Rescorla's truly random control conditions very soon knew with certainty that when placed in that box, they were going to get shocked at unpredictable intervals. Given what we have subsequently learned about the temporal Weber fraction in rodents, Rescorla's rats had between 3 and 4 bits of uncertainty about how soon they would get the next shock. For the rats in the truly random conditions, CS onsets did nothing to reduce that uncertainty. Consider, however, the rats where the CS signaled an increase in the rate of shock. The amount of information transmitted to these rats by CS onsets

is the difference in their uncertainty about how soon the next shock might come in the presence and absence of the tone CSs:

$$(1 - \ln\lambda_{\text{box}}) - (1 - \ln\lambda_{\text{tone\&box}}) = \ln\lambda_{\text{tone\&box}} - \ln\lambda_{\text{box}}$$
$$= \ln\left(\frac{\lambda_{\text{tone\&box}}}{\lambda_{\text{box}}}\right) = \ln\left(\frac{\mu_{\text{box}}}{\mu_{\text{tone\&box}}}\right) \quad (4)$$

Balsam and Gallistel (2009) call the ratio of the mean inter-shock intervals, μ_{box} and $\mu_{\text{tone\&box}}$, the *informativeness* of a protocol, because, as shown by the final expression in Equation (4), the log of that ratio is the amount of information transmitted by the CSs to the different groups of rats in Rescorla's (1968) experiment.

Informativeness is a unitless quantity, because the time units attaching to the means in Equation (4) cancel. This is an instance of a more general fact, which makes Shannon's theory so generally useful, even with continuous distributions like the exponential: Both the *differences* in the entropies of continuous distributions and their *relative entropies* are defined—they are computable from their parameters—even though the individual entropies are not (because the individual entropies become arbitrarily large as the time units become arbitrarily small). In the transmission of information, it is the differences in entropies that are most often important, because the information communicated to a subject by a signal (e.g., a tone onset or offset) is the difference between the subject's uncertainty before receiving the signal and their uncertainty after receiving it.

Figure 3 plots the degree of suppression that Rescorla (1968) observed as a function of the bits of information his CSs conveyed about how soon the next shock might occur

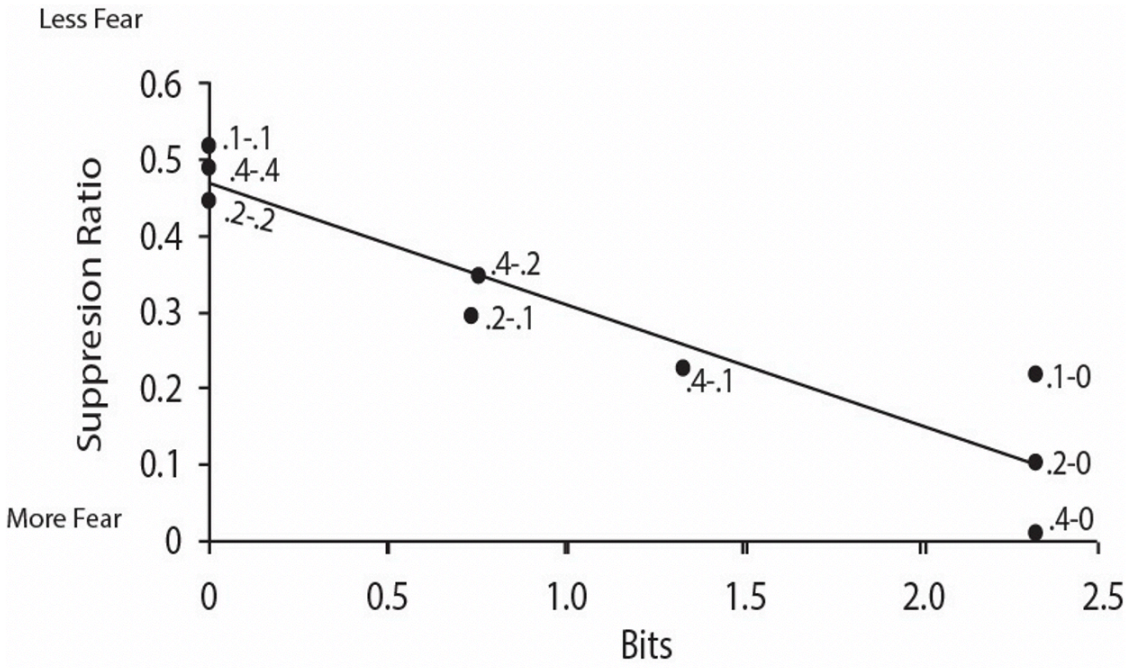

Figure 3. S*uppression ratio in Rescorla (1968) as a function of the number of bits of information the CS transmitted about how soon the next shock would occur. Figure courtesy of Peter Balsam, based on data in Figure 1. The numbers by the data points refer to the conditions in Figure 1.*

In a Pavlovian protocol, a subject's background level of uncertainty is determined by the overall rate of reinforcement while in the experimental chamber (total number of reinforcements divided by total time in the context). In an excitatory protocol, CS onset reduces this uncertainty because it signals an increase in rate, and the greater the rate, the lower the entropy of the exponential distribution. CS offset increases the subject's uncertainty because it signals a return to the lower background rate. In an inhibitory protocol, the roles of CS onset and offset are reversed: CS offset signals an increase in the rate of reinforcement, hence a reduction in uncertainty, while CS onset signals a return to the contextual rate, hence an increase in uncertainty. The contingency is the *unsigned* difference between the uncertainty produced by the higher rate and uncertainty produced by the lower rate divided by the background uncertainty, which is always the lower rate, hence the higher uncertainty. As may be seen in Equation (4), the temporal units cancel out, leaving a time-scale invariant indicator of the degree of "association" between the CS events and the reinforcement events. 'Association' is now in scare quotes because it refers to an objectively measurable property of the observable events, not to a hypothetical construct.

Gallistel, et al (2019) show how to apply the time-scale invariant, information-theoretic measure of contingency to operant protocols to solve the temporal assignment of credit problem in instrumental/operant learning/reinforcement learning. The prospective contingency between a response and a reinforcement is the difference between the entropy of the distribution of waits for reinforcement given a response and the distribution given a randomly chosen moment in time. The retrospective contingency is the difference between the entropy of the distribution of intervals looking back in time from a reinforcement to the most recent response and the entropy of the distribution looking back from randomly chosen moments in time.

In a variable-ratio (VR) protocol, the inter-reinforcement interval is a scalar function of the average inter-response interval since the last reinforcement. Therefore, the prospective contingency between rate of responding and rate of reinforcement is 1; the faster the pigeon pecks, the sooner it gets reinforced. In a variable-interval (VI) protocol, by contrast, the function relating rate of reinforcement asymptotes at the rate determined by the schedule. No matter how fast the pigeon pecks during an inter-reinforcement interval, it only gets reinforcement when the Poisson process times out. As previously explained, when subjects respond fast enough to produce the schedule-determined asymptotic rate of reinforcement, the prospective contingency between the subject's rate of responding and the its rate of reinforcement becomes unmeasurably small, as does the prospective contingency, $C_{\mathrm{R|r}}$ between a response, r, and the wait for reinforcement, R:

$$C_{\mathrm{R|r}} = 1 - \frac{\lambda_{\mathrm{VI}} \mathrm{e}^{-\lambda_{\mathrm{VI}}(\ln\lambda_{\mathrm{r}} - \ln\lambda_{\mathrm{VI}})/(\lambda_{\mathrm{r}} - \lambda_{\mathrm{VI}})}}{\lambda_{\mathrm{VI}}} \qquad (5)$$

where λ_{VI} is the parameter of the VI schedule and λ_r is the response rate. This function is plotted in Figure 4.

The difference between VR schedules and VI schedules in the functions relating the prospective contingencies to the rates of reinforcement and waits for reinforcement explains why a pigeon pecking on a VR schedule pecks much faster than a pigeon pecking on a VI schedule, even when the rates of reinforcement are equated (Catania, Matthews, Silverman, & Yohalem, 1977; cf Niv et al., 2005).

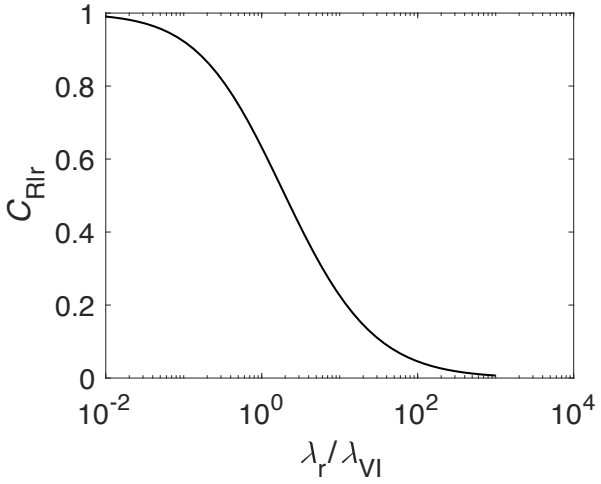

Figure 4. *The prospective contingency,* $C_{R/r}$ *between making a response, r, and the wait for reinforcement, R, on a VI schedule of reinforcement, as a function of the ratio between the rate of responding,* λ_r*, and the VI rate parameter,* λ_{VI}*. Semilogx scale.*

In protocols where reinforcement is triggered by the first response after a fixed interval has elapsed since the last reinforcement (FI schedules), the discrete prospective contingency between r and R is contingent on the time elapsed since the last reinforcement, t_r: When t_r < FI, $p(R|r) = 0$; when t_r > FI, $p(R|r) = 1$. Both distributions have 0 entropy, so the subject's uncertainty arises only from its uncertainty about t_r. That uncertainty is a function only of its Weber fraction, w. In protocols, where reinforcement is triggered by the n^{th} response following the last reinforcement (FR schedules), the prospective discrete contingency between r and R is contingent on the number of responses, n_r made since the last reinforcement: When n_r <FR, $p(R|r) = 0$; when n_r > FR, $p(R|r) = 1$. Again, both distributions have 0 entropy, so the subject's uncertainty arises only from its uncertainty about n_r. However, t_r covaries with n_r, because it takes more time to make more responses. Therefore, the t_r provides considerable information about n_r. Thus, it is not surprising that subjects' behavior is informed by both sources of information in a response-counting protocol (Light, et al, 2019).

In the 4 basic operant protocols, VR, VI, FR and FI, the retrospective contingency between R and r is 1. One might conjecture that this maximal retrospective contingency is encoded in the brain by a symbolic expression to the effect that r is the efficient cause of R. That 'efficient cause' might be among the categories by which brains encode experience is, of course, a hypothetical construct. Skinner would hate it (cognitive science = creation science). Hull would reject it because it is not a conditioned reflex, the only kind of knowledge he assumed brains had (Hull, 1930).

Perhaps it is time to move beyond Skinner and Hull, as Rescorla urged us to do. Rescorla was right: Associative learning is driven by temporal information—in the mathematically rigorous sense, the sense that makes it a measurable aspect of the experienced world. The time of occurrence of one event often provides measurable information about the temporal remove of another event. More technically, events are often stochastically connected through time. Insofar as the temporal relations between events persist over time, predictions and retrodictions may be used to anticipate future occurrences of the same events.

Brains have computational machinery that enables them to take advantage of the informative temporal relations among events to plan and organize behavior. Brains record where in time events occur (Crystal & Suddendorf, 2019; C. R. Gallistel, 1990; Panoz-Brown et al., 2016; Danielle Panoz-Brown et al., 2018) to create a temporal map of the subject's experience (Balsam & Gallistel, 2009). A temporal map of past experience enables a brain to compute the parameters of distributions and their entropies. To move ahead, the field need only realize that information is measured by the entropies of distributions, not by Bernoulli probabilities, and that the temporal information resides in the brain's metric temporal map.

Acknowledgments

I am grateful to Peter Balsam, Fuat Balci, David Freestone, Marc Howard, Vijay Namboodiri and Yael Niv for the critical comments that informed my revisions of an initial draft. The writing of this paper and the cited research from the Balsam lab was supported by NIMH Grant RO MH068073 to Peter Balsam and Eleanor Simpson.

References

Acquinas, T. (1256-1259). *Disputed questions on truth.* http://www.corpusthomisticum.org/qdv02.html

Balsam, P. D., Drew, M. R., & Gallistel, C. R. (2010). Time and Associative Learning. *Comparative Cognition & Behavior Reviews*, 5, 1-22. doi:10.3819/ccbr.2010.50001

Balsam, P. D., & Gallistel, C. R. (2009). Temporal maps and informativeness in associative learning. *Trends in Neurosciences*, 32(2), 73-78. doi:10.1016/j.tins.2008.10.004

Bialek, W., & Rieke, F. (1992). Reliability and information transmission in spiking neurons. *Trends in Neurosciences,* 15(11), 428-434. https://doi:10.1016/0166-2236(92)90005-S

Brenner, N., Bialek, W., & de Ruyter van Steveninck, R. (2000). Adaptive rescaling maximizes information transmission. *Neuron*, 26(3), 695-702. doi:10.1016/S0896-6273(00)81205-2

Catania, A. C., Matthews, T. J., Silverman, P. J., & Yohalem, R. (1977). Yoked variable-ratio and variable-interval responding in pigeons. *Journal of the Experimental Analysis of Behavior*, 28(2), 155-161. doi:10.1901/jeab.1977.28-155

Crystal, J. D., & Suddendorf, T. (2019). Episodic memory in nonhuman animals? *Current Biology,* 29(Dec 16), R1269–R1300. doi:10.1016/j.cub.2019.10.045

Dayan, P., & Berridge, K. C. (2014). Model-based and model-free Pavlovian reward learning: Revaluation, revision, and revelation. *Cognitive, Affective, & Behavioral Neuroscience,* 14(2), 473-492. doi:10.3758/s13415-014-0277-8

Fu, W.-T., & Anderson, J. R. (2008). Solving the credit assignment problem: Explicit and implicit learning of action sequences with probabilistic outcomes. *Psychological Research*, 72(3), 321-330. doi:10.1007/s00426-007-0113-7

Gallistel, C. R. (1990). *The organization of learning*. Cambridge, MA: Bradford Books/MIT Press.

Gallistel, C. R. (2020). The physical basis of memory. *Cognition*, doi:10.1016/j.cognition.2020.104533

Gallistel, C. R., Craig, A., Shahan, T.A. (2019). Contingency, Contiguity and Causality in Conditioning: Applying Information Theory and Weber's Law to the Assignment of Credit Problem. *Psychological Review, 126*(5), 761-773. doi:10.1037/rev0000163

Gallistel, C. R., & Gibbon, J. (2000). Time, rate, and conditioning. *Psychological Review, 107*(2), 289-344. doi:10.1037/0033-295X.107.2.289

Gershman, S. J., Moustafa, A. A., & Ludvig, E. A. (2014). Time representation in reinforcement learning models of the basal ganglia. *Frontiers in Computational Neuroscience, 7*, 8. doi:10.3389/fncom.2013.00194

Gibbon, J. (1981). The contingency problem in autoshaping. In C. M. Locurto, H. S. Terrace, & J. Gibbon (Eds.), *Autoshaping and conditioning theory* (pp. 285-308). New York: Academic.

Gibbon, J., & Balsam, P. D. (1981). Spreading associations in time. In C. M. Locurto, H. S. Terrace, & J. Gibbon (Eds.), *Autoshaping and conditioning theory* (pp. 219-253). New York: Academic.

Gibbon, J., Berryman, R., & Thompson, R. L. (1974). Contingency spaces and measures in classical and instrumental conditioning. J*ournal of the Experimental Analysis of Behavior*, 21(3), 585-605. doi:10.1901/jeab.1974.21-585

Gibbon, J., Locurto, C., & Terrace, H. S. (1975). Signal-food contingency and signal frequency in a continuous trials auto-shaping paradigm. *Animal Learning and Behavior*, 3, 317-324. doi:10.3758/BF03213453

Gibbon, J., Malapani, C., Dale, C. L., & Gallistel, C. R. (1997). Toward a neurobiology of temporal cognition: Advances and challenges. *Current Opinion in Neurobiology*, 7(2), 170-184. doi:10.1016/s0959-4388(97)80005-0

Gleitman, H., Nachmias, J., & Neisser, U. (1954). The S-R reinforcement theory of extinction. *Psychological Review*, 61(1), 23-33. doi:10.1037/h0062623

Gluck, M. A., & Thompson, R. F. (1987). Modeling the neural substrates of associative learning and memory: a computational approach. *Psychological Review*, 94(2), 176-191. doi:10.1037/0033-295X.94.2.176

Hallam, S. C., Grahame, N. J., & Miller, R. R. (1992). Exploring the edges of Pavlovian contingency space: An assessment of contignency theory and its various metrics. *Learning and Motivation*, 23, 225-249. doi:10.1016/0023-9690(92)90007-9

Hammond, L. J. (1980). The effect of contingency upon the appetitive conditioning of free operant behavior. *Journal of the Experimental Analysis of Behavior*, 34, 297-304. doi:10.1901/jeab.1980.34-297

Hammond, L. J., & Paynter, W. E. (1983). Probabilistic contingency theories of animal conditioning: A critical analysis. *Learning and Motivation*, 14, 527-550. doi:10.1016/0023-9690(83)90031-0

Hawkins, R. D., & Kandel, E. R. (1984). Is there a cell-biological alphabet for simple forms of learning? *Psychological Review*, 91, 375-391. doi:10.1037/0033-295X.91.3.375

Hull, C. L. (1930). Knowledge and purpose as habit mechanisms. *Psychological Review,* 37, 511-525. doi:10.1037/h0072212

Hull, C. L. (1952). *A behavior system*. New Haven, CT:: Yale University Press.

Kamin, L. J. (1967). "Attention-like" processes in classical conditioning. In M. R. Jones (Ed.), *Miami symposium on the prediction of behavior: aversive stimulation* (pp. 9-33). Miami: Univ. Miami Press.

Kamin, L. J. (1969). Predictability, surprise, attention, and conditioning. In B. A. Campbell & R. M. Church (Eds.), *Punishment and aversive behavior* (pp. 276-296). New York: Appleton-Century-Crofts.

Kamin, L. J. (1969). Selective association and conditioning. In N. J. Mackintosh & W. K. Honig (Eds.), *Fundamental issues in associative learning* (pp. 42-64). Halifax: Dalhousie University Press.

Kehoe, E., Ludvig, E. A., & Sutton, R. S. (2013). Timing and cue competition in conditioning of the nectitating membrane response of the rabbit (Oryctolagus cuniculus). *Learning and Memory*, 20, 97-102. doi:10.1101/lm.028183.112

Kimble, G. A. (1961). *Hilgard and Marquis' conditioning and learning*. NY: Appleton-Century-Crofts.

Light, K. R., Cotten, B., Malekan, T., Dewil, S., Bailey, M. R., Gallistel, C. R., & Balsam, P. D. (2019). Evidence for a mixed timing and counting strategy in mice performing a Mechner counting task. *Frontiers in Behavioral Neuroscience*, 13(109). doi:10.3389/fnbeh.2019.00109

Ludvig, E. A., Sutton, R. S., & Kehoe, E. J. (2008). Stimulus representation and the timing of reward-prediction errors in models of the dopamine system. *Neural Computation, 20*, 3034–3054. doi:10.1162/neco.2008.11-07-654

Ludvig, E. A., Sutton, R. S., & Kehoe, E. J. (2012). Evaluating the TD model of classical conditioning. *Learning & Behavior*, 40(3), 305-319. doi:10.3758/s13420-012-0082-6

Luzardo, A., Alonso, E., & Mondragón, E. (2017). A Rescorla-Wagner drift-diffusion model of conditioning and timing. *PLoS Computational Biology*, 13. doi:10.1371/journal.pcbi.1005796

Matzel, L. D., Held, F. P., & Miller, R. R. (1988). Information and expression of simultaneous and backward associations: Implications for contiguity theory. *Learning and Motivation*, 19, 317-344. doi:10.1016/0023-9690(88)90044-6

Minsky, M. (1961). Steps toward artificial intelligence. *Proceedings of the IRE, 49*(1), 8-30. doi:10.1109/JRPROC.1961.287775

Niv, Y. (2009). Reinforcement learning in the brain. *Journal of Mathematical Psychology*, 53, 139-154. doi:10.1016/j.jmp.2008.12.005

Niv, Y., Daw, N. D., & Dayan, P. (2005). How fast to work: response vigor, motivation and tonic dopamine. In Y. Weiss, B. Schölkopf, & J. R. Platt (Eds.), *NIPS 18* (pp. 1019–1026). Cambridge, MA: MIT Press.

Niv, Y., Daw, N. D., Joel, D., & Dayan, P. (2007). Tonic dopamine: Opportunity costs and the control of response vigor. *Psychopharmacology*, 191(3), 507-520. doi:10.1007/s00213-006-0502-4

Panoz-Brown, D., Corbin, H. E., Dalecki, S. J., Sluk, C. M., Wu, J.-E., & Crystal, J. D. (2016). Rats Remember Items in Context Using Episodic Memory. *Current Biology*. doi:10.1016/j.cub.2016.08.023

Panoz-Brown, D., Iyer, V., Carey, L. M., Sluka, C. M., Rajic, G., Kestenman, J., . . . Crystal, J. D. (2018). Replay of Episodic Memories in the Rat. *Current Biology, 28*(10), 1628-1634.e1627. doi:https://doi.org/10.1016/j.cub.2018.04.006

Poo, M.-m., Pignatelli, M., Ryan, T. J., Tonegawa, S., Bonhoeffer, T., Martin, K. C., . . . Stevens, C. (2016). What is memory? The present state of the engram. *BMC Biology*. doi:10.1186/s12915-016-0261-6

Prokasy, W. F. (1965). Stimulus fluctuation, reactive inhibition, and time betwen trials in classical eyelid conditioning. *Journal of Experimental Psychology*, 70, 464-472. doi:10.1037/h0022539

Rescorla, R. A. (1967). Pavlovian conditioning and its proper control procedures. *Psychological Review*, 74, 71-80. doi:10.1037/h0024109

Rescorla, R. A. (1968). Probability of shock in the presence and absence of CS in fear conditioning. *Journal of Comparative and Physiological Psychology*, 66(1), 1-5. doi:10.1037/h0025984

Rescorla, R. A. (1972). Informational variables in Pavlovian conditioning. In G. H. Bower (Ed.), *The psychology of learning and motivation* (Vol. 6, pp. 1-46). New York: Academic.

Rescorla, R. A. (1988). Pavlovian conditioning: It's not what you think it is. *American Psychologist*, 43, 151-160. doi:10.1037/0003-066X.43.3.151

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* (pp. 64-99). New York: Appleton-Century-Crofts.

Reynolds, G. S. (1961). Attention in the pigeon. *Journal of the Experimental Analysis of Behavior*, 4, 203-208. doi:10.1901/jeab.1961.4-203

Rieke, F., Warland, D., de Ruyter van Steveninck, R., & Bialek, W. (1997). *Spikes: Exploring the neural code*. Cambridge, MA: MIT Press.

Rissanen, J. (1989). *Stochastic complexity in statistical inquiry*. Singapore: World Scientific Publishing Company.

Rissanen, J. (1999). Hypothesis selection and testing by the MDL principle. T*he Computer Journal*, 42, 260–269. doi:10.1093/comjnl/42.4.260

Schultz, W. (2015). Neuronal reward and decision signals: From theories to data. *Physiological Reviews, 95*(3), 853-951. doi:10.1152/physrev.00023.2014

Schultz, W., Dayan, & Montague, P. R. (1997). A neural substrate of prediction and reward. *Science*, 275, 1593-1599. doi:10.1126/science.275.5306.1593

Shannon, C. E. (1948). A mathematical theory of communication. *Bell Systems Technical Journal*, 27, 379-423, 623-656. doi:10.1002/j.1538-7305.1948.tb01338.x

Skinner, B. F. (1950). Are theories of learning necessary? *Psychological Review*, 57, 193-216. doi:10.1037/h0054367

Staddon, J. E., & Zhang, Y. (1991). On the assignment-of-credit problem in operant learning. In *Commons, Michael L (Ed); Grossberg, Stephen (Ed); et al (1991) Neural network models of conditioning and action Quantitative analyses of behavior series (pp 279-293) xx, 359pp*.

Starkweather, C. K., Babayan, B. M., Uchida, N., & Gershman, S. J. (2017). Dopamine reward prediction errors reflect hidden-state inference across time. *Nature Neuroscience*, 20, 581- 589. doi:10.1038/nn.4520

Sutton, R. S. (1984). *Temporal credit assignment in reinforcement learning.* (PhD). University of Massachusetts, Amherst, Amherst, MA. Retrieved from http://scholarworks.umass.edu/dissertations/AAI8410337

Thomson, W. l. L. K. (1883). Electrical units of measurement. In *Popular Lectures and Addresses* (Vol. 1, pp. 73-460). New York: Macmillan and Company.

Tomonori, T., Duszkiewicz, A. J., & Morris, R. G. M. (2013). The synaptic plasticity and memory hypothesis: storage and persistence. *Philosophical Transactions of the Royal Society* B, 369. doi:10.1098/rstb.2013.0288

Vogel, E. H., Ponce, F. P., & Wagner, A. R. (2019). The development and present status of the SOP model of associative learning. *Quarterly Journal of Experimental Psychology., 72*(2), 346-374. doi:10.1177/1747021818777074

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). Hillsdale, NJ: Lawrence 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. doi:10.1037/h0025414

Wilkes, J. T., & Gallistel, C. R. (2016). Information Theory, Memory, Prediction, and Timing in Associative Learning (original long version). doi:10.3758/s13420-016-0240-3

Wilkes, J. T., & Gallistel, C. R. (2017). Information theory, memory, prediction, and timing in associative learning. In A. Moustafa (Ed.), *Computational Models of Brain and Behavior* (pp. 481-492). New York: Wiley/Blackwell.

Williams, D. A., Todd, T. P., Chubala, C. M., & Ludvig, E. A. (2017). Intertrial unconditioned stimuli differentially impact trace conditioning. *Learning & Behavior*, 45, 49-61. doi:10.3758/s13420-016-0240-3