14,560 words

Information Theory, Memory, Prediction, and Timing in Associative Learning

Jason T. Wilkes*

C. R. Gallistel†

*Department of Psychology
UCSB
Santa Barbara, CA 93106
jason.wilkes@psych.ucsb.edu

†Rutgers Center for Cognitive Science
152 Frelinghuysen Rd
Piscataway, NJ 08854-80210
galliste@ruccs.rutgers.edu

This is the original version of a chapter written for: *Computational Models of Brain and Behavior*. Edited by A. Moustafa. New York, Wiley/Blackwell (2016). The length originally specified by the editors was 15,000 words. Long after this version was submitted, new commissioning editors at Wiley decreed that all the chapters had to be shortened to 7,000 words. To reduce the chapter to that length required substantial rewriting and the elimination of all of the computational details. We post this for those who want to see computational details.

KEYWORDS
minimum description length maximum entropy assignment of credit partial reinforcement extinction stochastic models real-time

Abstract

Two information-theoretic principles—maximum entropy, and minimum description length—dictate a computational model of associative learning that explains cue competition (assignment of credit) and response timing. The theory's primitives are two cue types—state cues and point cues—and two stochastic distributions. The preferred stochastic model gives the relative code lengths for an efficient encoding of the data already seen; it predicts the data not yet seen; and the associated hazard function roughly predicts the observed timing of anticipatory (conditioned) behavior. State cues use the exponential distribution to encode, predict and time; point cues use a form of the Gaussian distribution that allows for event failure. An implementation of the refined minimum-description-length approach to stochastic model selection (Rissanen 1999) determines which stochastic model best compresses the data, and hence which is the best predictive model for a given protocol. The model brings into sharp focus the need to focus neurobiological inquiry on the coding question in memory.

Introduction

It is widely believed that brains compute but not in the way computers compute (Piccinini & Bahar, 2012). The grounds for belief in the uniqueness of neural computation are suspect because we know how computers compute but we don't know how brains compute. A common belief is that all neural computation can be conceived of as signal processing by, for example, convolution networks that have been rewired by experience (LeCun, Bengio et al. 2015).

In pure signal processing computations, transient signals generated by external events (stimuli) carry information into computational operations in the brain, where it is processed on line in real time, so memory is not required. More often than not, however, brains and computers process information extracted from experiences spread out in time and space and slowly accumulated in memory. Computers can do this because they have an addressable read-write memory into which information is put as it becomes available, and from which they retrieve information as it becomes relevant to further computation. By preserving accumulated information and the results of earlier computations in retrievable form, addressable read-write memory liberates computation from the tyranny of the current moment (Gallistel and King 2010).

The computations that enable brains to navigate time and space imply the presence of a similar memory in the brain. Abstract quantities like distance, direction, duration and probability are encoded in retrievable form and kept over long retention intervals. The computations mediating spatial navigation depend on retrieving the global geometry of the experienced environment (Moser, Kripff et al. 2008, Gallistel and Matzel 2013), but the large-scale environment is experienced one small portion at a time, not all at once. Temporal navigation—the anticipation of future events—depends on the remembered durations of the intervals between different events experienced on different occasions in the past (Balsam and Gallistel 2009). Risk assessment in the "switch" paradigm depends on the remembered relative frequencies of short and long trials, which are separately experienced one by one over long stretches of time (Balci, Freestone et al. 2009, Kheifets and Gallistel 2012). What is common to these diverse examples is that the quantities involved—distances, directions, relative rates, relative frequencies, etc.—are summaries of the animal's experience over time: summaries that cannot be computed without some method of preserving information over time in a computationally accessible form.

Hebbian synapses (more generally, plastic synapses)—the commonly assumed medium of memory—are not in and of themselves capable of encoding distances, directions, durations or probabilities (Gallistel and King 2010, Gallistel and Matzel 2013). Alterations in synaptic conductances are the hypothesized physical realization of associative bonds. Associative bonds are not symbols; they affect signal flow through the system, but they do not encode any experiential fact. Thus, they cannot supply computational machinery with information that has been accumulated piecewise over time (Elman 1990). Nor can they supply computations with information that may be needed under circumstances quite different from those in which it was acquired, that is, in

the absence of the original stimuli. In defiance of this obvious fact, many neurally oriented timing models adopt a signal-processing stance. They do not assume that experienced durations are encoded in a temporally stable readable memory medium. Instead, they posit temporally patterned neural activities set in motion by stimuli. These differentially rising and falling traces of past events become selectively associated with anticipatory behavior (Grossberg and Schmajuk 1991, Fiala, Grossberg et al. 1996, Meck 2003, Mauk and Buonomano 2004). In these models, experienced durations are not encoded in altered synaptic conductances. Neither are they encoded in the stimulus traces, because the traces are assumed to be innate properties of the stimulated neurons. The experienced intervals are merely implicit in the experience-altered neural circuitry. If you activate the circuit with the same stimuli that it experienced during the training experience, then it reproduces the interval that it experienced, but it does not encode that interval. In the absence of appropriate input from the environment, input that recapitulates input at the time the information was acquired, there is no way to read out what that interval was.

Timing models that are intended to be neurobiologically plausible rest on the assumption that the kinetics of neuronal responses to stimuli allows the brain to reproduce previously experienced intervals. One may doubt the neural plausibility of such accounts when confronted with the fact that brains remember durations measured in hours (Gibbon, Baldock et al. 1977), but that is not the biggest problem with this approach to memory. A bigger problem is that these accounts are unique to duration. They give no idea how the brain remembers and processes a quantity like distance. From the nervous system's perspective, a distance of several kilometers is not a “stimulus” in any straightforward sense of the term, yet the nervous systems of navigating animals appear to operate on distances of this magnitude regularly. They store, retrieve, and combine them with other variables in the service of navigational computations. Distances measured in kilometers cannot be reproduced within the confines of the 1mm-in-diameter brain of an insect or even the 10mm-in-diameter brain of a rat, but distances are nonetheless remembered—and computed with—even by insects (Menzel, Fuchs et al. 2011, Gallistel and Matzel 2013). The biggest problem is that there would appear to be no way to arithmetically combine quantities that are merely implicit in altered neural circuitry. How can the brain take an interval that is implicit in one neural circuit and subtract from it—or divide it by—an interval that is implicit in another neural circuit to produce a neural circuit in which the difference or ratio of those implicit input intervals is implicit?

We believe that the goal of computational theories of behavior is to guide neurobiological inquiry—to tell us what is to be looked for within the nervous system. Thus, the first goal of a computational theory should be to parsimoniously explain a rich body of behaviorally established fact. A theory that does this can then serve as a guide to what to look for in the brain. Many aspects of the behavior produced by the Pavlovian and operant conditioning protocols commonly used to study associative learning have been shown to depend on differences and ratios between intervals demarcated by different events and experienced at different times in the course of training (Gallistel 1990, Barnet and Miller 1996, Barnet, Cole et al. 1997, Savastano and Miller 1998, Arcediano, Escobar et al. 2003, Balsam and Gallistel 2009, Balsam, Drew et al. 2010).

And, it has repeatedly been shown that earlier experiences are re-encoded in the light of much later experiences (Matzel, Schachtman et al. 1985, Baker and Mercier 1989, Yin, Grahame et al. 1993, Blaisdell, Gunther et al. 1999, Arcediano, Escobar et al. 2003, Urushihara and Miller 2010). Thus, we want a theory that focuses on (i) how past experience is encoded in memory in a manner that allows it to be re-encoded in the light of later experience, and (ii) how that encoding can be used to predict future experience.

The Analytic Theory of Associative Learning

We call our theory the analytic theory of associative learning (TATAL), first because it is rooted in mathematical principles of optimal inference from states of extremely limited information—a computational problem that any well-designed mechanism for associative learning must be able to solve. Second, because it is implemented entirely with analytic, closed-form functions. The theory's core assumption is that the brain encodes inter-event intervals and the cues that predict them, using one or the other of two stochastic model forms, the exponential and the Gaussian. These stochastic models enable compression of the data extracted from experience in such a way that they can be reconstituted in their originally registered precision (lossless compression). The same models enable prediction of future experiences of the same kind.

The theory rests on a fundamental result in probabilistic inference: The model that achieves the greatest compression of data already seen is the model that best predicts data not yet seen (Grünwald, Myung et al. 2005). Thus, by asking not, "Which model is best?" but rather "How can the data of experience be best compressed?", the problem of model selection can be solved essentially for free. By focusing on data compression, the nervous system achieves two highly desirable goals: efficient use of memory, and good prediction of the future.

The theory is unusual among theories of associative learning in that: 1) There are no associative bonds (plastic synapses), hence no window of associability (Hawkins and Kandel 1984, Gluck and Thompson 1987) and no spike-timing dependent plasticity (Gallistel and Matzel 2013). 2) There is no ad hoc parsing of time into discrete trials prior to feeding experience to the model, as there is in most associative models (for example, Rescorla and Wagner 1972, Mackintosh 1975, Wagner 1981, McClaren, Kaye et al. 1989, Dickinson 2001). 3) The computational model does not parse time into sequences of discrete states, as do reinforcement learning models (see Gershman, Moustafa et al. 2014 for review). 4) There are no free parameters, hence no learning rates and no decaying stimulus traces. 5) Deciding between the two possible kinds of stochastic models mediates both cue competition (aka credit assignment) and the timing of conditioned responding. Thus, the theory tightly integrates what have been treated as separate aspects of associative learning. 6) The theory explains the parametric invariances in the acquisition and extinction of conditioned responses. These invariances constitute a serious explanatory challenge for associative theories of associative learning (Gallistel and Gibbon 2000, Balsam and Gallistel 2009, Gallistel 2012).

The theory is rooted in two information-theoretic principles that constrain the choice of stochastic forms ('models' in the statisticians' sense): the maximum-entropy principle (Jaynes 1957, Jaynes 2003) and the minimum-description-length principle (Rissanen 1978, Rissanen 1999, Grünwald, Myung et al. 2005). Both principles mediate the mathematical realization of Occam's razor, construed as "assume as little as possible."

The maximum entropy principle counsels us to choose the model form that has the maximum entropy (maximum residual uncertainty) given the parameters to be estimated from the data. When only the first moment (the mean) is estimated, this principle dictates the choice of the exponential distribution if the data are real valued, like inter-event intervals, and the Bernoulli if they are binary (happens/doesn't happen). When only the first two moments are estimated (the mean and standard deviation) and the data are real valued, the maximum entropy principle dictates the choice of the Gaussian distribution.

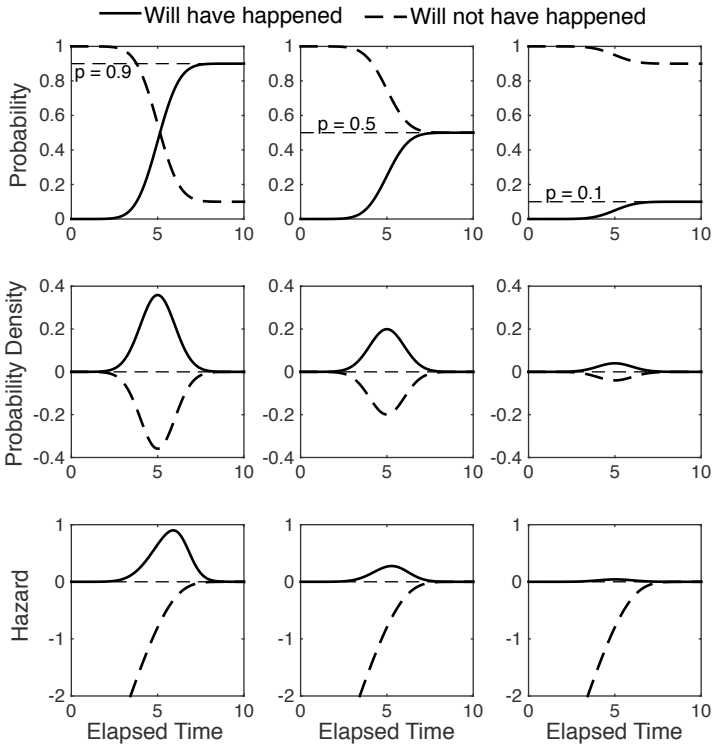


Figure 1. *BernoulliGauss functions ${}^{B}G(L, p,\mu,\sigma)$, for values of p ranging from 0.9 (left column) to .5 (middle) to .1 (right). Elapsed Time = L (for latency). Top row gives the cumulative distributions; second row the probability density functions; third row the hazard functions. The values of these functions are 2-element vectors (solid and dashed curves). See text for further explanation.*

In what follows, we will use the term BernoulliGauss to refer to an evitable Gaussian: that is, a Gaussian distribution whose event of interest may fail to occur. More concretely, a BernoulliGaussian ${}^{B}G(p,\mu,\sigma)$ is the distribution generated by a binary-outcome process (happens/doesn't happen) unfolding in time in accord with a Gaussian distribution on *t* (Figure 1). It describes the expected location in time of an event that occurs with probability, *p*, following a point cue at an expected latency, μ, with an expected temporal dispersion measured by σ.

The exponential distribution is the distribution of the intervals between events generated by a Poisson process (aka a random rate). The temporal locations of the events cannot be predicted, only their rate. Therefore, the exponential distribution can only be

signaled by states. State cues, by definition, have duration. The onset or offset of a state may serve as a point cue (a temporal landmark), but the state itself cannot so serve because it is spread out in time. The BernoulliGauss, by contrast, can only be signaled by a point cue, an event that marks a point in time, from which a latency can be measured, as in delay conditioning and trace conditioning protocols. Thus, we have two possible cue-model pairs: state-exponential and point-BernoulliGauss, hereafter denoted by State_E and Point_^{B}G.

The minimum-description-length principle is that the more a stochastic model enables lossless compression of data already seen, the better it predicts the data not yet seen. Thus, learning to better predict the future is the same as learning to better compress the data extracted from past experience (Barron, Rissanen et al. 1998, Grünwald, Myung et al. 2005)—provided that model complexity is properly taken into account. In our theory, this principle mediates the choice between the stochastic models permitted by the maximum-entropy-principle.

A striking feature of the theory is that in solving the cue competition problem, it also solves the response-timing problem The issues of cue competition and response timing are together evident when considering Figure 2, which shows the first two CS presentations in an experimental protocol like Rescorla's (1968) truly-random-control experiment. CS is short for a conditioned stimulus, a stimulus that may predict a motivationally significant event in conditioning protocol. The latter is called a US, short for unconditioned stimuli. In Figure 2, the potential CSs are the Background (aka the context), which is the chamber in which the protocol is run, and a transient tone, which comes on and goes off repeatedly while the subject is in the chamber. The constant presence of the background is indicated by the gray. The intermittent presence of the superimposed tone (conventionally, the CS) is indicated by the white boxes riding on the background gray. The US is indicated by the dot.

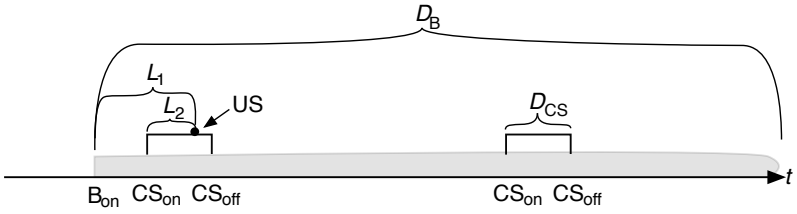


Figure 2. *Time line for the first two trials (CS presentations, white rectangles) in one of Rescorla's (1968) protocols. The subject is introduced to the box (gray background) at B_{on}. The US is a shock to the feet. Its latency is L_1 if measured from B_{on} or L_2 if measured from CS_{on}. The state cues are the Background (gray) and the CS (white). The point cues are the onsets and offsets of these states. The appropriate encoding and prediction model is ambiguous at time t (see text).*

At time *t* in Figure 2, there is massive ambiguity as to the best model for encoding the so-far experienced timing of US occurrences and predicting its recurrence. Four 1-cue stochastic models are consistent with the theory's principles:

Bck_Exp model. The shocks occur at random in the experimental chamber. That the first one happens while the tone CS is present is pure coincidence. This was

in fact the case in one of Rescorla's (1968) protocols, the so-called truly random control.

CS_ExpE model. USs occur randomly but only when the tone CS is present. This was the case in Rescorla's (1968) CS-US contingency protocol.

$Bck_{on}{}_{}^{B}G$ *model*. Placement in the box predicts a shock at Latency 1 (cf Fanselow 1990).

$CS_{on}{}_{}^{B}G$ *model*. Shocks occur with some probability less than 1 at Latency 2 following the onsets of the tone CS, as in a delay conditioning protocol with partial reinforcement.

Our computational realization of the minimum-description-length principle's prescription for stochastic model selection explains the results observed in the classic cue competition experiments (Rescorla 1968, Wagner, Logan et al. 1968, Kamin 1969), including the retroactive blocking and unblocking versions of those experiments (Matzel, Schachtman et al. 1985, Yin, Grahame et al. 1993, Blaisdell, Gunther et al. 1999, Urushihara and Miller 2010).

Its success rests on two fundamental results in information theory and statistics: i) Shannon's efficient coding theorem and ii) Rissanen's (1989) refined MDL solution to the problem of devising a language-independent measure of a stochastic model's complexity.

Shannon (1948) proved that a maximally efficient encoding of the data from a given source must encode the i^{th} possible datum with a code word whose length is proportionate to $\log_2(1/p_i)$, where p_i is the probability of that datum from that source. His theorem establishes a monotone inverse mapping between the objective probabilities of the possible events for a given source and the optimal relative lengths of the code words, that is, the physically realized symbols by which those events may be represented in a receiver, such as the brain; the higher the probability of a datum, the shorter the optimal symbol for it. Samuel Morse understood this principle intuitively when he chose to symbolize 'e'—the most frequent letter in English—with a dot, the shortest possible code in his system.

The to-be-encoded events of principal interest in our theory are the homogeneous and heterogeneous inter-event intervals. A homogeneous inter-event interval is the interval between two events of the same kind, for example, between two USs of the same kind or between two successive onsets of the same CS. A heterogeneous inter-event interval is the interval between two different events, for example between a CS onset and a US, or between a CS onset and its offset, or between the onset of one CS and the onset of a different CS.

By Shannon's theorem, a scheme for encoding data into memory is an implicit stochastic model for the data it encodes, because an encoding implicitly assigns a relative frequency (probability) to each possible datum when it uses a code of a given length to stand for that datum. In an inefficient code, the relative lengths of the symbols do not match the logarithms of the inverses of the relative frequencies of the events to which they refer. ASCII is an inefficient code for the typographic characters on a keyboard,

because it assigns an 8-bit word to every character. It implicitly assumes that every character is used with equal frequency, which is far from the case.

The more the probability distribution implicit in a coding scheme deviates from the optimal scheme, the less it compresses the encoded data and the more poorly the implicit stochastic model predicts future data from that source. The ASCII code is thus a terrible guide to the relative frequency with which different typographic elements are used. The Kullback-Leibler divergence gives, *in bits per datum encoded*, the cost of encoding data non-optimally, that is with a stochastic model that does not accurately represent the source statistics (Kullback 1987, Cover and Thomas 1991). The better a stochastic model is—that is, the more it captures non-random structure in the data—the higher the frequency-weighted average of the probabilities it assigns; hence the shorter the average code word.

A stochastic model, like the exponential distribution or the Gaussian distribution or a regression model, assigns probabilities to possibilities. If two different models assign two different sets of probabilities to the same set of possibilities, then the model that assigns the better set (the set in which the assigned probabilities better match the actually experienced relative frequencies), is the better model, *assuming the models are equally complex*. However, they may not be equally complex. In our case, a Point:^{B}G model is more complex than a State:E model. Rissanen's refined MDL method for measuring the cost of a stochastic model (Rissanen 1999, Grünwald, Myung et al. 2005) solves the problem of finding the happy medium in the trade-off between model complexity and model accuracy. Overly complex models find bogus structure in the data, which they use to achieve illusory data-coding efficiency, an efficiency that does not continue to be realized as more data come in. Moreover, because the structure they find is bogus, they badly predict the future.

Rissanen elaborated a principle that enables one to compute the language-independent complexity of a stochastic model in the same currency (bits or nats) as the cost of encoding the data already seen using the symbol sizes dictated by the probabilities assigned by that model. The cost of encoding the model itself is a function only of its so-called parametric complexity and the size of the data set. Thus, the cost of storing the stochastic model used to generate the code for the data appears as a surcharge on the amount of data encoded. The best model among those entertained is the one that achieves the lowest total cost, where total cost is the sum of the cost of encoding the data using word lengths dictated the stochastic model plus the cost of encoding the stochastic model itself. The models chosen in this way make the smallest demand on memory and best predict the future (Grünwald, Myung et al. 2005). We assume that in the evolution of the neurobiological machinery that encodes experience into a memory medium and uses that experience to anticipate future experience, the more efficient use of the memory medium and the accurate prediction of the future have both increased fitness.

Solving the Cue Competition Problem

If State_E models—models in which different state cues predicted different random rates—were the only possible form of stochastic model, then the problem of finding

the model that maximizes coding efficiency would be solved by the matrix-algebra computation at the core of Rate Estimation Theory (Gallistel 1990, Gallistel and Gibbon 2000):

$$\lambda_{\mathrm{C}} = T^{-1}\lambda_{\mathrm{R}},$$

where $\lambda_{\mathrm{C}} = \langle \hat{\lambda}_1, \hat{\lambda}_2, \ldots, \hat{\lambda}_n \rangle$ is the vector of corrected US rate assignments, one for each state cue in the protocol, $\lambda_{\mathrm{R}} = \langle N_1/T_1, N_2/T_2, \ldots, N_n/T_n \rangle$ is the vector of raw rates, and **T** is the temporal coefficient matrix:

$$\mathbf{T} = \begin{matrix} 1 & \frac{T_{1,2}}{T_1} & \cdots & \frac{T_{1,n}}{T_1} \\ \frac{T_{2,1}}{T_2} & \ddots & \cdots & \frac{T_{2,n}}{T_2} \\ \vdots & \cdots & \ddots & \vdots \\ \frac{T_{n,1}}{T_n} & \cdots & \cdots & 1 \end{matrix}$$

In this matrix computation, N_i is the cumulative count of USs that occur in the presence of the i^{th} state; T_i is the *cumulative presence* of the i^{th} state (the reading on a clock that runs only when the i^{th} state is present, and never resets); and $T_{i,j}$ is the cumulative presence of the superposed i^{th} and j^{th} states (the reading on a clock that runs only when those two state cues are superposed, and never resets).

The matrix computation identifies the assignment of US rates to state cues that maximizes the likelihood, that is, the probability of the data. In so doing, it identifies the assignment that maximizes coding efficiency under the assumption that only exponential models are appropriate.

An exponential model is *not* appropriate in delay and trace conditioning. In those protocols, the US repeatedly follows a point cue at a more or less fixed interval. In such cases, the BernoulliGauss distribution, which we denote by $^{\mathrm{B}}\mathrm{G}$, gives a more efficient encoding. The second stage of our computational model identifies the cases in which a Point: $^{\mathrm{B}}\mathrm{G}$ model gives a better encoding than a State:E model.

The cumulative $^{\mathrm{B}}\mathrm{G}$ is a 2-element-vector function of L, the elapsing time, as measured from a point cue (Figure 1, top row). The first element of the vector function specifies the probability of the event's having already happened; the second element specifies the complementary probability of its having not happened yet:

$$^{\mathrm{B}}\mathrm{G}(L, p, \mu, \sigma) = \langle p\Phi(L, \mu, \sigma) \quad (1-p)\big(1 - \Phi(L, \mu, \sigma)\big) \rangle,$$

where Φ is the cumulative normal distribution. The surprisal when the predicted event happens at latency L following the point predictor is $-\log p - \log\phi(L, \bar{L}, \sigma)$, where ϕ is the normal probability density function. The surprisal when it fails to happen at

approximately the expected time ($\bar{L}$) following the point predictor is –log(1-p). A failure is deemed to have occurred when the event does not occur within a latency of 1.5$\bar{L}$.

When encoding the inter-event intervals that terminate with a US, the second stage of the computation compares distributions of the homogeneous US-US intervals to the distributions of heterogeneous intervals terminating in a US. In doing so it asks whether a cue other than the US itself does a better job of predicting the wait time for the next US. Both the homogeneous and the heterogeneous intervals are measured on a clock that runs only when the relevant state is present.

In the case of delay conditioning, where the US occurs only during the CS and always at a fixed interval after CS onset, the first stage attributes a non-zero US rate only to the CS. Therefore, the second stage of computation—the model selection stage—considers only the models that could predict the USs that occur only in the presence of the CS. In Figure 2, the only heterogeneous interval that could predict the US in the presence of the CS is the CS_{on}–>US interval. Therefore, the CS_{on}:^{B}G(p,μ,σ) model is ^{B}G(p=0.5, μ=L_2, σ=wL_2), where w is the subject's Weber fraction for the representation of duration.

The Weber fraction, w, limits the scalar precision with which the brain represents a quantity. A consequence of this scalar limit on precision is that the σ parameter of the ^{B}G function is defined even when there is only one datum, as in Figure 2, where there is only one US, hence only one L_2. The competing state-cue model is CS:E(μ=L_2). The CS_{on}:^{B}G model assigns higher probability to the L_2 datum than does the CS:E model, first because its mode is at L_2, and, second because it concentrates the probability mass around that mode. However, it has greater parametric complexity than the exponential model. Thus, the question is whether the greater complexity of this stochastic model form is justified by reduced data encoding costs. The total cost of encoding data vector, **D**, using stochastic model, *M,* is:

$$C(\mathbf{D}^n|M) = -\log P\left(\mathbf{D}\middle|M,\hat{\theta}(\mathbf{D})\right) + COMP_n(M),$$

where n is the length of the data vector. The first term is the probability of the data given both the stochastic model and the maximum likelihood estimate of its parameter vector, $\hat{\theta}(\mathbf{D})$. The second term is the so-called parametric complexity of the stochastic model. For the Bernoulli model,

$$COMP_n(\mathrm{B}) = \log \sum_{i=0}^{i=n} \binom{n}{i} \left(\frac{i}{n}\right)^i \left(\frac{n-i}{n}\right)^{n-i}, \qquad (1)$$

which, for $n > 100$, is approximated by

$$COMP_n(\mathrm{B}) = \frac{1}{2}\log n + .2. \qquad (2)$$

We approximate the exponential model as a Bernoulli process by discretizing time into bins of width τ, where τ is sufficiently small so that the p in the geometric

approximation to the exponential (the probability of an event falling in any one bin) is always very small. Then, the *n* for any inter-event interval, *I*, is correspondingly large, because we have chosen τ such that $n(I) = I/\tau \gg 1$ for all actual *I*. We then use (2) to approximate, the complexity of the exponential model by:

$$COMP_{T/\tau}(\mathrm{E}) = \frac{1}{2}\log\frac{T}{\tau} + .2, \qquad (3)$$

where *T* is the cumulative presence of the predicting state.

The parametric complexity of a Gaussian model is approximated by

$$COMP_{n_{\mathrm{US}}}(\mathrm{G}) = 1 + \log\frac{2\frac{T}{\tau}\sqrt{n_{\mathrm{us}}}}{\sqrt{2\pi}}, \qquad (4)$$

In this formula, n_{us} is the number of USs, hence the number of Point–>US intervals in the encoded data set. The encoded data set is comprised of the number of USs the model imputes to a given point cue; *w* is the Weber fraction for the brain's representation of duration; and *T* is the maximum value that the latency could have at a given point in the subject's experience of that predictive state (Foster and Stine 2005).

The complexity of the $^{\mathrm{B}}$G model is given by either (1) + (4) or (2) + (4), depending on N_{CS}, which is the number of $\mathrm{CS}_{\mathrm{on}}$ or $\mathrm{CS}_{\mathrm{off}}$. These numbers give the number of occasions on which the onset or the offset of the CS *could* have predicted a US. In a partial reinforcement protocol, $N_{\mathrm{CS}} > N_{\mathrm{US}}$, because in such a protocol the CS often fails to predict a US, while every US is predicted by a CS.

With these formulae, we compute the total costs of the candidate encodings as a function of time and trials, using one or another model. The protocol shown in Figure 2 is ambiguous: depending on how it further unfolds, four different 1-cue models may prove to deliver the best encoding of it. In Figure 3, we plot the total costs as these 4 possible protocols lead to different stochastic models:

1. Subsequent USs (dots) are scattered at random without regard to the presence of absence of the CS. In that case, the B_E model prevails. This is Rescorla's (1968) truly random control, in which the subject eventually attributes the USs to the background.

2. Subsequent USs appear only when the CS is present but are scattered at random within the CS intervals. In that case, the CS_E model prevails. This is Rescorla's (1968) contingency protocol in which the subject eventually attributes the USs to the CS.

3. Further USs may occur but always at latency L_2 following a CS onset. In that case, the $\mathrm{CS}_{\mathrm{on}}$_$^{\mathrm{B}}$G model prevails. This is the delay conditioning protocol, the most common of protocols. The timing of the anticipatory

behavior (conditioned responses) produced by this protocol differs dramatically from the timing of the conditioned responses produced by Protocol 2—see following section on Response Timing.

4. There is no 2nd US so long as the session continues, but the US recurs at the same latency on each subsequent placement in the test chamber. In that case, the $B_{on}\_^{B}G$ model prevails, because its cost does not increase with further unreinforced CS presentations, whereas the cost of encoding the data with the $CS_{on}\_^{B}G$ model does increase, because the estimate of its *p* parameter gets smaller and smaller.

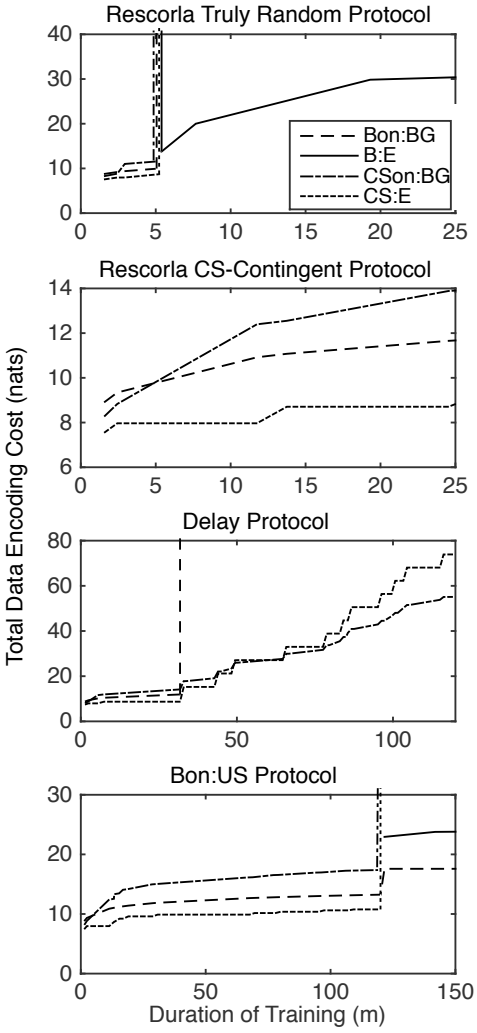


Figure 3. *The total encoding costs as a function of time elapsed in the protocol. For each of 4 protocols (top to bottom), 4 different models are considered (legend). The training time shown for a protocol is that required for the model with the lowest final cost to emerge as the best model. The final cost was the cost after 20 sessions (120 minutes/session). The models with the lowest final cost are those that accord with the literature on cue competition for these protocols: In the Truly Random protocol, the USs are attributed to the context (Background); in the Contingent protocol, they are attributed to the CS state; in the Delay protocol, they are attributed to CSon; in the Bon:US protocol, they are attributed to B_{on}. Vertical segments indicate when the cost for a given model goes to, or returns from, infinity. For example, in the top plot, the costs of the CSon:^{B}G, CS:E and Bon:^{B}G go to infinity at shortly after 5 minutes, when the first ITI US occurs (and the 2nd US in the first session), at which point, those three models become invalid because none of them can predict that US. The cost of the B:E model has been infinite up to this point, because it is ascribed 0 rate by the matrix computation. With the occurrence of an ITI US, the background is no longer ascribed 0 rate and so the B:E model now predicts both the first and second US.*

Response Timing

A stochastic distribution provides a suite of four parameterized functions: a cumulative probability function, *F*; a probability mass or probability density

function, F'; the survivor function, 1- F, and the hazard function, $F'/(1-F)$. Roughly speaking, the hazard function gives the probability that the event will happen in the next moment, *given that it has not yet happened*. Thus, the hazard function is a natural basis for the timing of anticipatory behavior.

Associative theories of associative learning do not, generally speaking, attempt to explain the timing of conditioned responses. When they do attempt it, they make assumptions about the kinetics of different neurons activated by the CS. These assumptions introduce a large number of free parameters (Fiala, Grossberg et al. 1996, Meck 2003, Mauk and Buonomano 2004). The analytic theory explains cue competition and response timing using the assumptions already specified in the explanation of cue competition, with no free parameters.

The exponential hazard function is flat; the probability that a US (aka a reinforcer) occurs in the next moment is independent of how long it has been since: 1) the last US; 2) observation began; 3) the Poisson process began generating USs. Thus, when the exponential model is used to encode the subject's experience of US latencies and to anticipate their recurrence, one expects to see steady responding, as, in fact, one does. The best known example is the steady responding seen in variable interval schedules of reinforcement (Ferster and Skinner 1957, Gallistel, Craig et al. 2013), where the distribution of inter-reinforcement intervals is approximately exponential.

There are also examples in the Pavlovian fear conditioning literature. Libby and Church (1975) used 4 protocols with foot shock as the US. The US occurred only during CS presentations. In two protocols, CS duration varied exponentially with an expectation of 1 m. In the other two, it was fixed at 1 m. The rate of US occurrence in the presence of the CS was 1/minute in all 4 protocols.

In one of the two protocols where CS duration varied exponentially, the fixed rate of US occurrence was achieved by having the US occur at a random rate of 1/minute whenever the CS was on. In the other, it was achieved by having a US occur at the termination of each CS. In either case, the US hazard stepped up abruptly when the CS came on and remained at that same higher level so long as the CS stayed on. As may be seen in Figure 4A, these protocols produced similar patterns of responding; the fear index went up abruptly at CS onset then subsided to a more or less flat level within 20 seconds of CS onset. There was no statistically significant decline in the fear ratio after 30 s into the CSs in either of these protocols. The abrupt overshoot in fear at CS onset is reminiscent of the overshoot in brightness at points were luminance abruptly increases. In other words, it is a contrast effect. Here, the transient overshoot occurs in response to an abrupt increase in the imminence of shock. The flat fear response later in a CS is predicted by the CS:E model's flat hazard function.

In two protocols where CS duration was fixed at 1 minute, the pattern of responding differed dramatically depending on whether the CS occurred at a random rate during the CSs or at the termination of each CS (Figure 4B). In the first case, the US hazard decreased as the interval since CS onset increased, because the hazard of CS termination increased. In the second case, the US hazard was low at CS onset and increased

throughout the CS. In the first case, the fear index decreased throughout the CS; in the second case, it increased. Thus, in all 4 protocols, the pattern of responding reflects the US hazard, which is signaled by the hazard function for the stochastic model that TATAL predicts for that protocol.

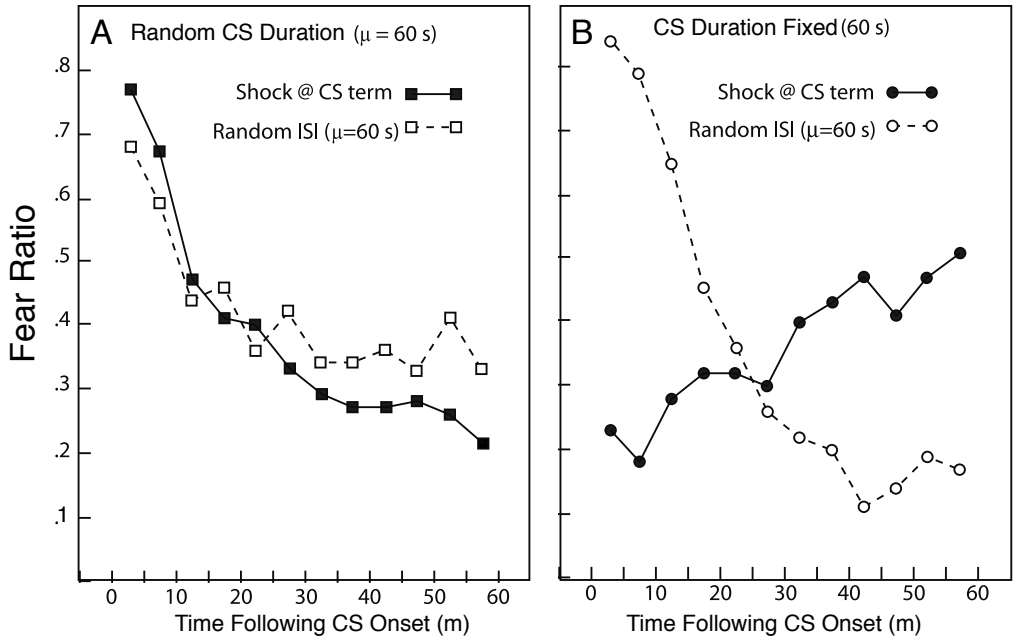


Figure 4. *Mean fear ratio as a function of time since CS onset for 4 different fear-conditioning protocols, in each of which shocks occurred at a rate of 1/m in the presence of the CS.* **A.** *CS duration random with 1m expectation. USs occurred either at CS termination or randomly during CS.* **B.** *CS duration fixed at 1m. USs occurred either randomly during CS or at its termination. Retraced from Figures 2 and 3 in Libby and Church (1975) with permission of the author(s) and the publisher.*

A well-known protocol for demonstrating the fundamental role of remembered reinforcement latency in associative learning is the peak protocol. In this protocol, CS onset signals that a US (a small amount of food) will be delivered with some probability (usually 0.5) in response to the first response made after a fixed interval has elapsed (the fixed reinforcement latency). On trials where the reinforcement (US) is not delivered (called peak trials), the CS persists for 3 or 4 times the expected reinforcement latency.

Subjects trained on the peak protocol begin responding abruptly when some proportion of the reinforcement latency has elapsed. On peak trials (unreinforced trials), they cease responding equally abruptly when the elapsed time in the CS exceeds by some proportion the expected reinforcement latency (Cheng and Westwood 1993, Church, Meck et al. 1994, Gallistel, King et al. 2004). The proportions at which responding begins and ends vary from peak trial to peak trial. Averaging responding over the peak trials and normalizing by the maximum of the average gives the probability that the subject is responding as a function of the time elapsed since CS onset (Figure 5, solid curve). The dashed line in Figure 5 is the ^{B}G hazard function when $p = 0.5$ and $w = .33$.

The ^{B}G hazard function provides an appropriate time signal on which to base the decisions to start and stop responding. It does not, of course, predict the observed probability-of-responding curve exactly. One would not expect it to do so, because the abrupt onsets and offsets of responding are determined by the location of, and variability in, start and stop decision thresholds (Church, Meck et al. 1994). That is why there are marked between-subject differences in the exact forms of the peak functions. However, like the ^{B}G hazard function, empirical peak functions always rise to a peak at the expected latency of reinforcement and then fall. That is, it provides the timing signal needed for an appropriately timed response.

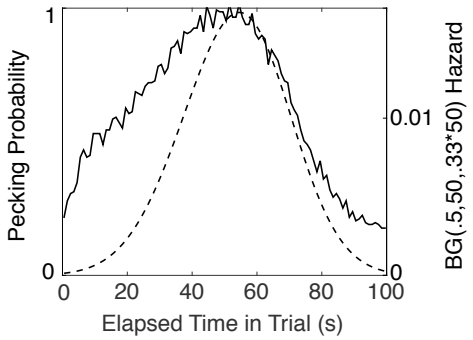


Figure 5. *Probability of a conditioned responding (pecking at the illuminated key) as a function of time elapsed in peak trials for one pigeon in a protocol with a 50s feed latency following key illumination (solid curve, plotted against left axis) and the ^{B}G hazard function for that latency with p = .5 and a Weber fraction of .33 (dashed curve, plotted against right axis). The behavioral data are from the laboratory of John Gibbon (deceased).*

In delay conditioning, the ^{B}G hazard function drops to essentially 0 immediately after CS onset. Thus, although the onset of a CS shortens the anticipated wait for the next US, its immediate effect is to lower the US hazard. Fairhurst, Gallistel and Gibbon (2003) showed that this effect can be made behaviorally manifest using the peak protocol. They ran a peak protocol with two different CSs; one that predicted reinforcement at a 60 s latency; the other predicted reinforcement at a15 s latency. The two CSs occurred singly on some trials and together on others. When they occurred together, the onset of the CS that predicted reinforcement at a 60s latency was followed after 45 seconds by the onset of the CS that predicted reinforcement at a 15 s latency. Reinforcement occurred at the jointly predicted latency on 50% of all trials. On the 2-CS peak trials, the rise in the probability of responding in response to the 60s predictor was exactly the same as when the 60-s predictor occurred without the 15-s predictor, as it should be. However, at the onset of the 15-s predictor, the probability of responding dropped abruptly, only to rise again much more steeply as the 15 s latency elapsed. It is difficult to see how theories that explain anticipatory responding by postulating selective associations to neurons whose innate kinetics cause the activities stimulated by a CS to peak at different latencies could explain this result. In such a theory, the onset of the second CS should immediately increase the probability of a conditioned response, not reduce it.

Inhibitory Conditioning

Inhibitory *state* cues predict a reduction in the US *rate*. If, for example, USs occur only in the absence of a tone, then the tone's presence signals that the rate predicted by the Background or Context is suppressed.

Inhibitory *point* cues predict the *failure* of a US to occur at a *time* when it is otherwise expected. The protocols that convincingly show timed suppression employ three CSs, two of which, the excitatory CSs, predict a US coincident with their termination at the end of a fixed delay. During training, the third CS, the inhibitory CS, is presented together with one of the excitatory CSs on some trials. On those trials, the US fails to occur. As training proceeds, the anticipatory (conditioned) response to the excitatory CS is suppressed on the trials when it occurs in compound with the inhibitory CS. On test trials, the inhibitory CS is paired with the other excitatory CS, the one not used in the inhibitory training. It suppresses the anticipatory responses to that CS as well, but only if that excitatory CS predicts the US with the same latency as the one used in training. In other words, the inhibitory effect is localized in time, just as is the excitatory effect (Denniston, Cole et al. 1998, Burger, Denniston et al. 2001, Denniston, Blaisdell et al. 2004).

The computation of the inhibitory effects is the same as for the excitatory effects. The inhibitory CS may function as a state cue or as a point cue. In either case, the matrix computation gives a positive rate for the excitatory cue and a negative rate for the inhibitory cue. The second computational stage, the computation of model complexity computation, then determines which is the better inhibitory model, the CS:E model (State:Exponential) or the CS_{on}:^{B}G model (Point:BernoulliGauss). The CS_{on}:^{B}G model for an inhibitory CS prevails over the CS:E model when the onset of the excitatory cue predicts the time at which the US is expected and the inhibitory CS cancels or reduces the probability that the US occurs *at that time*.

Recall that the ^{B}G function outputs a 2-element vector (Figure 1). For the cumulative ^{B}G distribution, one element, p_e, gives the probability as a function of time that the US will have happened (Figure 1, top, solid curves). This probability starts at 0 and rises to asymptote at p, the probability of a US on any given trial. The second element, $p_i = 1 - p_e$ gives the probability that the US will not have happened *yet*; it starts at 1 and decreases to asymptote at 1–p (Figure 1, top, dashed curves). The decrease in the second element reflects the increase in the first, because the two elements must sum to 1 at every moment.

The excitatory and inhibitory probability densities are the derivatives of the cumulative functions. In the typical delay inhibitory conditioning protocol, where the US always occurs when only the excitatory CS comes on and always fails to occur when the excitatory and the inhibitory CS both come on, p_e, the probability density for the excitatory effect of the excitatory CS, and p_i, probability density for the inhibitory effect of the inhibitory CS, are equal in magnitude and opposite in sign at all values of the latency. Their sum, therefore, is everywhere 0, which means that the hazard function is flat at 0 throughout the trial.

The ^{B}G function formalizes the notion of the No-US, which has long been an essential, but conceptually problematic feature of associative theories of associative learning (Gleitman, Nachmias et al. 1954, Gallistel 2012). The ^{B}G formalization removes the conceptual difficulties by localizing the No-US in time, by making it specific to the

failure of an explicitly encoded US, and by distinguishing between inhibitory contexts (state cues) and inhibitory point cues (CS onsets and offsets).

The Parametric Invariances in Acquisition and Extinction

Trials-to-acquisition is the number of CS presentations required before the (usually abrupt, see Gallistel, Balsam et al. 2004) appearance of an anticipatory response to the CS (the conditioned response). It varies by as much as two orders of magnitude from subject to subject (Gallistel, Balsam et al. 2004, Papachristos and Gallistel 2006). It also depends on which conditioned response one chooses to record (Gallistel, Craig et al. 2013 Figures 5 and 6). The same is true for trials to extinction; the number of successive unreinforced CS presentations required to suppress a previously established conditioned response. Clearly, predicting trials to acquisition and trials to acquisition requires free parameters that are adjusted subject-by-subject and that differ for differing choices of the conditioned response that is to be extinguished.

We are interested in what a theory of associative learning can explain without resort to free parameters. We focus therefore on three robust parametric invariances observed in the acquisition and extinction of conditioned responses in classical conditioning protocols:

i. *Time-scale invariance*: Median trials to acquisition is proportionate to the CS duty cycle, that is, to the ratio of T_{CS} (cumulative time on the clock that runs whenever the CS is present) and T_B (cumulative time on the clock that runs whenever the subject is in the conditioning context). The smaller this ratio is, the sooner the conditioned response appears. The duration of the CS has no effect on trials to acquisition in and of itself; varying CS duration varies trials to acquisition only if it varies the duty cycle (Gibbon and Balsam 1981, Gallistel and Gibbon 2000, Ward, Gallistel et al. 2012, Ward, Gallistel et al. 2013). The number of CS-US pairings (reinforced trials) delivered in a given amount of training time is also irrelevant to the progress of conditioning (Gottlieb 2008). These two aspects of the time-scale invariance of acquisition pose serious explanatory challenges to associative theories of associative learning. In those theories, the absolute interval between predictive CS and US and the number of CS-US pairings are the protocol parameters that principally determine the rate at which an association strengthens.

ii. The number of reinforced trials to acquisition—trials on which the US predicted by the CS actually occurs—is invariant under partial reinforcement. Interpolating, for example, an average of 9 unreinforced trials between each reinforced trial does not increase the number of reinforced trials required for the conditioned response to appear (Gibbon, Farrell et al. 1980, Williams 1981, Gottlieb 2005, Harris 2011). Because the interpolation of unreinforced CS presentations weakens net excitatory strength in associative theories of associative learning, this invariance also constitutes a strong explanatory challenge for such theories.

iii. The number of expected reinforcements that must be omitted to extinguish a conditioned response is invariant under partial reinforcement; increasing the number of unreinforced CS presentations interpolated randomly between reinforced CS presentations by a factor of 10 during training increases trials to extinction by that same factor. Therefore, the number of expected reinforcements that have failed to occur when the conditioned response no longer occurs is unaltered by the partial reinforcement during training (Gibbon, Farrell et al. 1980). In its qualitative form, this is known as the partial reinforcement extinction effect; partial reinforcement *increases* trials to extinction. This fact has constituted a strong explanatory challenge for associative theories of associative learning for more than half a century (Kimble 1961, pp. 318-329).

We can understand the parametric invariances by focusing on the information-theoretic divergence *of* one stochastic distribution *from* another. The divergence, which is asymmetric (hence the emphasis on 'of' and 'from'), is measured by the Kullback-Leibler divergence *of* one distribution, Q, *from* another, P:

$$D_{KL}(P||Q) = \sum_i P(i)\log\frac{P(i)}{Q(i)}.$$

When the stochastic model, Q, that currently dictates the encoding of the data diverges from a model, P, that would encode the data more economically, the Kullback-Leibler divergence measures the additional coding cost per datum encoded. The excess cost of using the poorer code to encode n data is:

$$nD_{KL}(P||Q). \qquad (5)$$

Experiment shows that the stochastic differences that determine trials-to-acquisition depend only on the differences in the raw rate of US occurrence in a given context and the raw rate of US occurrence during CSs. The divergence of a normal distribution (assumed to be supported only on $\mathbb{R}^+$) from an exponential may be factored into two terms (Balsam and Gallistel 2009), one dependent on the difference between the first moments (the expectations) and one dependent on the square root of the second moment of the normal distribution (σ). Experiment shows that median trials to acquisition depends only on the difference in the first moments (Ward, Gallistel et al. 2012). What matters in acquisition is the factor by which CS onset alters the expected wait time. This factor is ratio of the US rate in the presence of the conditioned stimulus (${}^{\mathrm{CS}}_{\mathrm{US}}N/T_{\mathrm{CS}}$) to the overall background rate, the rate when subject is in the test environment (${}^{\mathrm{B}}_{\mathrm{US}}N/T_{\mathrm{B}}$).

In excitatory acquisition, the CS_Exp model, which attributes the USs to the CS state, competes with the Bck_Exp model, which attributes them to the background state (aka the context). The Kullback-Leibler divergence of a lower random rate, λ_C, (for context rate) from a higher random rate, λ_T, (for trial rate, i.e., rate during a CS) is $log(\lambda_T/\lambda_C) = log(C/T)$, where C denotes the contextual US cycle period. This period is $1/\lambda_C$. It is the average wait time for USs in the experimental context (without regard to whatever transient CSs may or may not be present from time to time). T denotes the average wait time during a delay trial, that is, $1/\lambda_T$. From this and (5), we see that the difference in encoding cost between the CS_Exp model and the Bck_Exp model in excitatory CS

conditioning grows as a scalar function of the *C*/*T* ratio. The duty cycle is the inverse of this ratio. That is why trials to acquisition is proportionate to the duty cycle, as was first shown by Gibbon and Balsam (1981). In other words, this well-established quantitative result is predicted by simple information-theoretic considerations.

Reinforcing only on average every n^{th} CS presentation increases the expected wait time in the CS state by a factor of *n.* However, it also increases the background wait time (the US cycle period) by that same factor. Thus, partial reinforcement does not alter the CS duty cycle, the factor by which CS onset reduces the expected wait time. Therefore, the fact that partial reinforcement does not increase the number of reinforced CS presentations required for the conditioned response to appear is another manifestation of the time-scale invariance of acquisition.

An often-advanced intuitive explanation for the partial reinforcement extinction effect (Mowrer and Jones 1945, Kimble 1961, Baum 2012) is that partial reinforcement during training makes it harder to distinguish the training state (when CSs or responses are occasionally reinforced) from extinction state (when CSs or responses are never reinforced). The Kullback-Leibler divergence quantifies this intuition. When the experimenter stops delivering USs at the start of extinction, the subject continues for some while to encode its experiences using the $CS_{on}\_^{B}G$ model, with the estimate for the Bernoulli *p* parameter developed during training. In extinction, that parameter estimate is no longer valid. For the Bernoulli distribution, which is supported on [0,1], the divergence *of Q*, the data-encoding distribution appropriate during training, *from P* which is now, in extinction, the better encoding distribution, is

$$P(1)\log\frac{P(1)}{Q(1)} + P(0)\log\frac{P(0)}{Q(0)}. \qquad (6)$$

That is, in extinction, the poorer stochastic model, *Q*, is the Bernoulli distribution in which $p = p_{tr}$, the probability of reinforcement during acquisition (initial training), while the better model is Bernoulli for which $p = 0$. Substituting 0 for $P(1)$ and 1 for $P(0)$ in (6), yields after some algebra

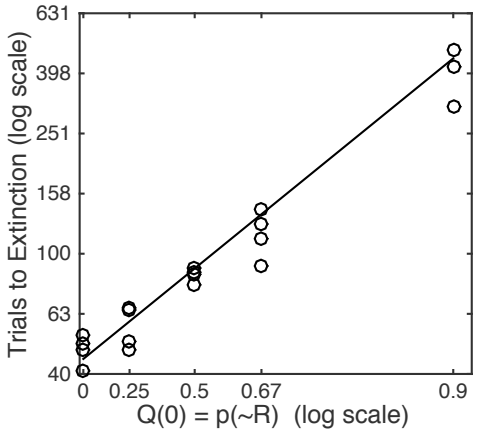
$$-\log\big(Q(0)\big) = \log\frac{1}{Q(0)},$$

from which we see, in accord with intuition, that the discriminability of the extinction state from the training state decreases as the probability of non-reinforcement during training increases.

The effect on trials to extinction of the probability of non-reinforcement during training has been assessed over the range $.9 \geq Q(0) \geq 0$, that is, from reinforcement of every CS presentation to reinforcement of as few as a randomly chosen 1 in 10 (Gibbon, Farrell et al. 1980). Over that range, the log of trials to extinction is a linear function of $-\log(Q(0))$ with a slope of 1; see Figure 6. Thus, the higher the proportion of unreinforced trials during training, the more slowly evidence of the change accumulates. As may be seen in Figure 6, quantitative results on the partial reinforcement extinction effect—trials to extinction increases in proportion to the probability of non-reinforcement during training (Gibbon,

Farrell et al. 1980)—are explained by TATAL. The explanation requires no new assumptions, and it depends on no free parameters. Like trials to acquisition, it is quantitatively predicted by simple, analytic information-theoretic considerations.

Figure 6. *Trials to extinction vs the probability of an unreinforced CS presentation, p(~R), during training. Logarithmic scales. The data points derive from the regression lines in the top 4 panels of Gibbon, Farrell et al (1980). The solid line with slope 1 was placed by eye. Its good agreement with the data implies that the number of trials to extinction is proportionate to trials/reinforcement during training, as predicted by Kullback-Leibler divergence of the training distribution from the extinction distribution*.



Assessing Model Viability

There are theoretical and empirical reasons to believe that subjects do not update their stochastic models observation by observation. On our theory, updating the model entails re-encoding the data in memory. The cost of doing so cannot be negligible. Therefore, from a theoretical perspective, model revision should not be undertaken unless the current model is no longer viable. This consideration is in the basic spirit of the theory, which is to keep the experience-encoding model as simple as possible, but no simpler.

On the empirical side, the conditioning literature contains several examples of a paradoxical combination of behavioral stasis and abrupt change. For example, human subjects when estimating outcome-by-outcome the hidden parameter of a Bernoulli process do not change their estimate trial by trial. Despite the local variations in relative frequency inherent in a random sequence, they keep a constant estimate, sometimes for hundreds of successive trials—an example of behavioral stasis. However, when the hidden parameter does change, subjects respond quickly and abruptly (Gallistel, Krishan et al. 2014).

There are similar results for non-human subjects: Given concurrent variable interval schedules, where rewards are delivered at two different locations at different random rates, animal and human subjects match the ratio of the average duration of their visits to the two locations to the ratio of the incomes flowing from those two sources. When computed reward-by-reward, the income flows are extremely noisy, because the intervals between the rewards at each location are distributed exponentially. The parameters of rats' visits to these two locations are unaffected by these large, purely random reward-by-reward fluctuations in the income flows—another example of behavioral stasis. However, when the rate parameters for the two exponentials change frequently, subjects (rats and mice) respond to each change in the income flows by changing the parameters of their visits quickly and abruptly (Mark and Gallistel 1994). These results and others of a

similar nature imply that stochastic models are updated only when the experiences they encode require an update.

The stasis-vs-abruptness paradox led Gallistel et al (2014) to formulate a principle informally termed "if it ain't broke, don't fix it," where 'it' refers to the current stochastic model. The reasoning behind the principle is that it is easier to determine whether the current stochastic model does justice to the data so far accumulated than to determine why it has failed and what needs to be done to fix or replace it. More often than not, the current model continues to be acceptable after an additional observation. Therefore, an efficient system should update its model (make new parameter estimates) only on those rare occasions when the latest experienced renders the current model no longer tenable. Implementing this principle requires a measure of the viability of the current model.

The model-viability test proposed Gallistel et al (2014) relies on the surprisingly simple distributional properties of an easily computed information-theoretic statistic, $n\mathrm{D}_{\mathrm{KL}}(P||Q)$, the product of the number of data encoded (n) and a Kullback-Leibler divergence, $\mathrm{D}_{\mathrm{KL}}(P||Q)$. The distributions whose divergence is measured (the P and Q distributions) both take the form of the current model for the encoded data vector **D**. They differ only as regards their estimates of values for its parameters. The P distribution uses estimates based on the entire vector. Therefore, its parameter estimates change, at least a little, almost every time a new datum is incorporated into **D**. The Q distribution uses the estimates that determine the current encoding of **D,** the estimates that will continue to determine the encoding as long as the current model continues viable. These estimates, the parameter estimates for the Q distribution, do not change as the data vector grows, *unless* the current encoding scheme is shown by the $n\mathrm{D}_{\mathrm{KL}}(P||Q)$ statistic to be inefficient.

For the Bernoulli model form,

$$\mathrm{D}_{\mathrm{KL}}(P||Q) = \sum P_i \log\frac{P_i}{Q_i} + \sum (1-P_i)\log\frac{1-P_i}{1-Q_i}$$

For the Gaussian model form,

$$\mathrm{D}_{\mathrm{KL}}(P||Q) = \frac{\frac{\sigma_P^2 + \left(\mu_Q - \mu_P\right)^2}{\sigma_Q^2} - 1 + \log\frac{\sigma_Q^2}{\sigma_P^2}}{2}$$

When the current model is correct, the distribution of the $n\mathrm{D}_{\mathrm{KL}}(P||Q)$ statistic is gamma(.5,1) regardless of the sample size, n. This is proved to be the case for the Bernoulli distribution in Gallistel et al (2014). Because the exponential when discretized becomes the Bernoulli, this result also holds for the exponential. We have determined by simulation that this result also holds for the Gaussian. Therefore, a simple decision threshold on the $n\mathrm{D}_{\mathrm{KL}}$ statistic, can mediate the datum-by-datum decision whether or not to change the current model. This greatly reduces the computational burden on the system. It needs to evaluate alternatives to the current model only when the current model is no longer viable, which is rarely the case.

Finding Principles for Determining the On-Deck Models

In the analytic theory, the minimum-description-length principle decides which is the best model from among the stochastic models that the system can generate. In practice, however, our system can generate a potentially infinite number of models. It can do so in response to parametric non-stationarities, that is, changes in the parameters of an otherwise viable model form, such as occur during extinction and reacquisition. Or, it can do so in response to changes in predictive cues. Or, it can do so in response to changes in the form of the encoding model.

Non-stationarity. Both real life and experimental associative learning protocols have non-stationarities. Often these are simply changes in parameter values, as, for example, in extinction and reacquisition protocols, where the probability of reinforcement goes to 0 during extinction then returns to its training value during reacquisition. Each such change requires a new and more complex stochastic model. The new model has to specify the parameters before the change, the parameters after the change, and the change point. The non-stationarity of model parameters generates the simplest kind of compound model—concatenations of the same primitive model form with the same predictor cues, but with intermittent changes in its parameters.

Changes in predictive cues. Early in conditioning, there are often changes in model cues and/or model form. When the first US occurs during a transient CS, as in Figure 2, the simplest model is the CS_Exp model. However, the occurrence of the first US during a CS may be what is colloquially called "just a coincidence." Subsequent occurrences may reveal no contingency between that CS and the US. In that case, a model in which the background is the predictive state will replace the model in which the CS is the predictive state.

Changes in stochastic form. On the other hand, subsequent occurrences may reveal a fixed latency from CS onset to the US, in which case a $CS_{on}\_^{B}G$ model will replace the CS_Exp model. Because these replacement models are not excluded by early experience, they should be kept on deck, so to speak. The fact that they are not much more complex than the current model and still viable should place them among the alternative models that are the first to be evaluated when the initial model proves unsatisfactory. We call these likely replacements the "on-deck" models.

An infinity of more complex models may remain viable more or less indefinitely. Thus, there is a question how many such alternative models should be on deck. While the minimum-description length principle will forestall any one of the much more complex models becoming the preferred model until simpler models have failed, it is a waste of computational effort to evaluate a host of alternatives every time the current model fails. It is also a waste of effort to evaluate those relatively simple models (ones that would otherwise be on deck) that may already be seen to be inadequate by the nD_{KL} statistic. This statistic can be used not only to assess the continuing validity of the current model but also to prescreen more complex models.

Still more complex models. Returning now to the infinity of more complex models: Subsequent experience may reveal that there is a CS-US rate contingency only in that context (that test box). In which case, the stochastic model will require ANDing that context and that CS. This situation is called occasion setting in the conditioning literature (Holland 1992). CSs can function as both negative and positive occasion setters. Negative occasion setting requires NANDing the occasion-setting cue and the cue for which it sets the occasion.

Configural conditioning is also well established empirically (Rescorla 1973, Rescorla, Grau et al. 1985): In positive figural conditioning, CS1 and CS2 do not predict the US when presented alone but do predict it when presented "in compound." Positive configural conditioning requires a model in which the cues are ANDed. In negative configural conditioning, CS1 and CS2 each predict the US when presented alone but not when presented "in compound." This requires a model in which the cues are XORed.

Subsequent experience may also reveal that a mixture distribution is required. For example, in the double-standard time-left protocol, pigeons learn to switch appropriately between a choice in which, on the one key (the time-left key), the initially very long latency to reward declines linearly as the trial progresses, while on the other key (the standard key), the latency will be either very short (15s) or very long (230s). These reinforcement latencies are measured from a randomly chosen moment of commitment. When that moment arrives, the pigeon is stuck with the latency programmed for the key it is pecking at that moment. The two latencies on the standard key require for their representation a mixture of two Gaussians, and that is what in fact the pigeon learns (Bruner, Gibbon et al 1994).

In short, arbitrarily many more complex models can be built from the two primitive distributions in our theory. We are still exploring the issue of how best to order the models whose ability to represent the data is to be tested when the current model fails, that is, the on-deck models. We want to be confident that the principles for placing alternative models on deck are such as to guarantee that a better model does not fail to be discovered as soon as the data establish it as the best model because it does not make it on deck. Simply counting the parameters to be estimated together with the connectives to be used provides a good preliminary ordering of the possible models. A rule that puts on deck only those models with a count no greater than 2 higher than the count for the current model and that are not rejected by the $n\mathrm{D}_{\mathrm{KL}}$ statistic may prove acceptable.

Discussion

The basic idea behind TATAL is simple: the brain encodes its experiences with the aid of the simplest adequate stochastic model. That same model enables it to anticipate future occurrences of the encoded events. The brain encodes the durations of experienced intervals (latencies, wait times) using two maximally simple, innately specified stochastic distributions: the exponential and the BernoulliGauss. The first of these applies when the encoded events are randomly distributed in time; the second applies when the events occur with a predictable probability at predictable times.

The brain's most basic preliminary assumption is that events are randomly distributed in time. Thus, it always computes the rates associated with the different experienced states. This computation enables it to anticipate the increases and decreases in expected wait times concomitant with changes in the mix of superposed state cues. Because rates are additive, the computation is simple; it requires only basic matrix algebra (the solving of simultaneous equations) to determine which rates must be attributed to which states.

It tests the viability of its basic model (the exponential) using the $n\mathrm{D}_{\mathrm{KL}}(P||Q)$ statistic. If the events are clustered in time within a state rather than distributed randomly, that statistic will reveal the inadequacy of the exponential model. The brain will then redo its encoding of the experienced intervals using the ${}^{\mathrm{B}}\mathrm{G}$ function. This function is a Bernoulli distribution that unfolds as a Gaussian function of the time elapsed since a point cue. A point cue is a temporal landmark that makes it possible to predict the time and probability of events that follow at a predictable latencies.

The computation that enables the brain to decide on the appropriate model takes account of the differing complexities of the competing models; the exponential has only one parameter to be estimated from experience; the ${}^{\mathrm{B}}\mathrm{G}$ has three. An insight of fundamental importance in the theory of probabilistic inference is that a process for choosing between competing stochastic models that fails to take account of model complexity overfits the data. Overfitting data already seen with a needlessly complex model worsens the accuracy with which the data not yet seen (future events) are predicted. In other words, our intuitive sense that simpler models are better provided they are not *too* simple—Occam's razor—is a mathematically provable truth about stochastic models. We suggest that Occam's razor appeals to our intuition because it is implicit in the brain's machinery for stochastic model selection and stochastic model selection is at the core of the brain's capacity to represent its experience. The analytic theory identifies the selective pressures that would drive the evolution of machinery that embodies Occam's razor; machinery that minimizes memory load and maximizes predictive accuracy.

The theory is more powerful than the other computational theories of associative learning known to us. The computational realization of these simple ideas—which is achieved entirely with closed-form computations—explains a wider range of well-established results in the vast associative-conditioning literatures than any other theory known to us. The same assumptions explain cue competition, conditioned inhibition, the timing of conditioned responses, and the parametric invariances. The theory is naturally extendable to configural conditioning, occasion setting and protocols involving mixture distributions. None of the explanations depends on assumptions about the values of free parameters, such as learning rates, the widths of windows of associability, and/or the rates at which stimulus traces decay. The theory is fully specified; it does not depend on an ad hoc, pre-theoretical parsing of the continuous flow of experience into discrete trials. Most associative models of associative learning are not fully specified; their use requires an ad hoc parsing of the continuous flow of experience into discrete trials.

Insofar as one takes greater power, greater simplicity and full specification as indications that one theory approaches closer to the underlying reality than another, the success of TATAL has implications for our understanding of the neurobiology of cognition. TATAL brings to the center of attention the coding question: how are the facts the brain derives from a subject's experience of its environment encoded in the brain's memory medium (Gallistel and Matzel 2013)? TATAL's most basic assumption is that the brain encodes the durations of experienced intervals into an information-preserving medium with the aid of simple stochastic models.

Most experimental psychologists, computational neuroscientists and neurobiologists interested in the neurobiology of learning and memory assume that experience-produced alterations in synaptic conductances are the medium of enduring memories. This conviction drives most research and theorizing in this centrally important area of cognitive neuroscience. Yet, there is no attempt to state how in theory—or to discover how in fact—changes in synaptic conductances encode the elementary quantitative facts of experience. The question is rarely addressed with anything more specific than allusions to distributed codes in vast networks of synaptic connections (Martin and Morris 2002). How facts such as the durations of several hundred inter-event intervals, demarcated by different events, and spread out in experience over weeks or months could be so encoded and then read out into the computations required for stochastic model selection remains an undiscussed mystery.

From an information theoretic perspective, the durations of intervals are numbers pure and simple. Any medium for encoding information should transparently be able to encode them. The only biological entities whose structure transparently enables the encoding of numbers are the polynucleotides that swarm within every neuron. We know how to encode numbers in polynucleotides, because we do it when we use bar-coding to elucidate neuronal connectivity and intracellular in vivo neurochemistry (Rosenthal 2001, Peikon, Gizatullina et al. 2014).

There are no suggestions about how to encode numbers (durations, distances, directions, probabilities, intensities, etc) in altered synaptic conductances. None of the neurally oriented theories of interval timing known to us proposes a scheme for encoding durations in patterns of altered synaptic conductance. In all of them, the information about the experienced duration resides not in the altered synaptic conductances themselves but rather in the innate dynamics of the presynaptic neurons or in the ongoing activity state of a recurrent circuit or sometimes in a tacitly assumed memory whose physical realization is unspecified. They all assume, as associative theories always have assumed (Hull 1930), that a complex circuit involving many neurons has been rewired by experience so that it behaves *as if* the *average* duration of a repeatedly experienced interval were encoded, but that average interval is not encoded in the transparent way in which it would be if the data were stored in computer memory or in a polynucleotide sequence. Information about facts as simple as the durations of intervals is said to be distributed throughout a neural network in such an opaque way that we must abandon hope of ever putting our experimental fingers on it. It is in there only implicitly, not explicitly, because the nets are sub-symbolic computing machines (Smolensky 1986).

Recent results on the neurobiology of the timing of the conditioned eye blink by cerebellar Purkinje cells (Johansson, Jirenhed et al. 2014, Wetmore, Jirenhed et al. 2014, Johansson, Carlsson et al. 2015) open a radically different perspective on the neurobiology of memory and computation, a perspective consistent with the implications of the analytic theory. These recent results show that the memory for the duration of the CS_{on}–>US interval is stored inside the Purkinje cell itself, not in the altered conductances of the synapses that the parallel fibers make on the Purkinje cell, nor in the innate dynamics of the granule cell responses to the CS, nor in the innate dynamics of the Purkinje cell itself. The same Purkinje same cell can learn different intervals. The cell's read-out of the interval it has most recently learned is independent of the firing pattern in the parallel fibers that carry the CS-generated signal to it.

Other recent results from very different learning protocols using different species (both vertebrate and invertebrate) suggest the same conclusion, although not so directly as the results just cited. Results from fear conditioning in mice (Ryan, Roy et al. 2015) and gill-withdrawal conditioning in Aplysia (Chen, Cai et al. 2014) suggest that the acquired information in fear conditioning is not encoded in the altered synaptic conductances themselves but rather within the postsynaptic neurons. Results from olfactory fear conditioning in mice show that the learned predictive power of one odor versus other non-predictive odors is manifest in differential transmitter release in the olfactory glomeruli from first-order olfactory neurons (Kass, Rosenthal et al. 2013).

When we look for the medium of memory inside neurons rather than in synapses or circuits, we look for a molecular medium. The medium most obviously suited to this purpose is the polynucleotide; it can store 2 bits per nucleotide. Moreover, we know that every cell contains elaborate machinery for reading and editing and rearranging the information in polynucleotides. There are, however, many other possibilities, so we do not here champion the polynucleotide. We champion only the more cautious hypothesis that there is a molecular medium within neurons, which stores information in much the way a computer does, and which makes that information accessible to computation, in much the way that DNA makes inherited information accessible to computation. To store information in the way a computer does, which is also the way DNA does (Gallistel 2016, in press), one needs addressable banks of thermodynamically stable switches. Many different molecules may plausibly be imagined to serve in this capacity.

The savings in space and energy expenditure from storing information at the molecular level rather than at the circuit level are measured in orders of magnitude (Gallistel 2016, in press). If the information on which computations operate is stored in molecular media inside neurons, then much of the computational machinery that operates on that information is likely to be implemented there, too, rather than at the circuit level. In other words, many computations now thought to be mediated by circuit level structures requiring spatially extensive and energetically expensive signaling may in fact be implemented by intracellular neurochemistry. Here again, the savings to be realized in the energy required to effect the computations and the space required for the requisite machinery are measured in orders of magnitude. In this regard, Sterling and Laughlin write (Sterling and Laughlin 2015, p.124), "These advantages—compactness, energy efficiency, and ability to adapt and match—all suggest the principle compute with

chemistry. It is cheaper." To that we would add the further principle, Store information in molecular media; it's energetically cheaper, it takes up vastly less space, and it enables molecular level computation.

Acknowledgements

We are grateful to Peter Balsam, Timothy Shahan, David Freestone, Fredrik Johansson, and Greg Olsen for helpful comments on earlier drafts, and to Jacob Feldman for suggesting to CRG some years back that the minimum description length approach to model simplicity was of potential interest for theory development in cognitive science.

References

Arcediano, F., M. Escobar and R. R. Miller (2003). "Temporal integration and temporal backward associations in humans and nonhuman subjects." Learning and Behavior **31**: 242-256.

Baker, A. G. and P. Mercier (1989). Attention, retrospective processing and cognitive representations. Contemporary learning theories:Pavlovian conditioning and the status of traditional learning theory. S. B. Klein and R. R. Mowrer. Hillsdale, NJ, Lawrence Erlbaum Associates**:** 85-116.

Balci, F., D. Freestone and C. R. Gallistel (2009). "Risk assessment in man and mouse." Proceedings of the National Academy of Science U S A **106**(7): 2459-2463.

Balsam, P. and C. R. Gallistel (2009). "Temporal maps and informativeness in associative learning." Trends in Neurosciences **32**(2): 73-78.

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

Barnet, R. C., R. P. Cole and R. R. Miller (1997). "Temporal integration in second-order conditioning and sensory preconditioning." Animal Learning and Behavior **25**(2): 221-233.

Barnet, R. C. and R. R. Miller (1996). "Second order excitation mediated by a backward conditioned inhibitor." Journal of Experimental Psychology: Animal Behavior Processes **22**(3): 279-296.

Barron, A., J. Rissanen and B. Yu (1998). "The minimum description length principle in coding and modeling." IEEE Transactions on Information Theory **44**(6): 2743-2760.

Baum, W. (2012). "Extinction as discrimination: The molar view." Behavioural Processes **90**: 101-110.

Blaisdell, A. P., L. M. Gunther and R. R. Miller (1999). "Recovery from blocking achieved by extinguishing the blocking CS." Animal Learning and Behavior **27**: 63-76.

Brunner, D., Gibbon, J., & Fairhurst, S. (1994). "Choice between fixed and variable delays with different reward amounts." Journal of Experimental Psychology: Animal Behavior Processes, 20, 331-346.

Burger, D., C., J. Denniston, C. and R. R. Miller (2001). "Temporal Coding in Conditioned Inhibition: Retardation Tests." Animal Learning & Behavior **29**(3): 281-290.

Chen, S., D. Cai, K. Pearce, P. Y.-W. Sun, A. C. Roberts and D. L. Glanzman (2014). "Reinstatement of long-term memory following erasure of its behavioral and synaptic expression in Aplysia." eLife.

Cheng, K. and R. Westwood (1993). "Analysis of single trials in pigeon's timing performance." Journal of Experimental Psychology: Animal Behavior Processes **19**: 56-67.

Church, R. M., W. H. Meck and J. Gibbon (1994). "Application of scalar timing theory to individual trials." Journal of Experimental Psychology: Animal Behavior Processes **20**(2): 135-155.

Cover, T. M. and J. A. Thomas (1991). Information theory. New York, Wiley.

Denniston, J. C., A. P. Blaisdell and R. R. Miller (2004). "Temporal Coding in Conditioned Inhibition: Analysis of Associative Structure of Inhibition." Journal of Experimental Psychology: Animal Behavior Processes **30**: 190-202.

Denniston, J. C., R. P. Cole and R. R. Miller (1998). "The role of temporal relationships in the transfer of conditioned inhibition." Journal of Experimental Psychology: Animal Behavior Processes **24**(2): 200-214.

Dickinson, A. (2001). "Causal learning: An associative analysis. ." Quarterly Journal of Experimental Psychology **54B**: 3-25.

Elman, J. L. (1990). "Finding structure in time." Cognitive Science **14**: 179-211.

Fairhurst, S., C. R. Gallistel and J. Gibbon (2003). "Temporal landmarks: Proximity prevails." Animal Cognition **6**(2): 113-120.

Fanselow, M. S. (1990). "Factors governing one-trial contextual conditioning." Animal Learning and Behavior **18**: 264-270.

Ferster, C. B. and B. F. Skinner (1957). Schedules of reinforcement. East Norwalk, CT, Appleton-Century-Crofts.

Fiala, J. C., S. Grossberg and D. Bullock (1996). "Metabotropic glutamate receptor activation in cerebellar Purkinje cells as substrate for adaptive timing of the classically conditioned eye-blink response." Journal of Neuroscience **16**(11): 3760-3774.

Foster, D. P. and R. A. Stine (2005). The contribution of parameters to stochastic complexity. Advances in Minimum Description Length: Theory and Applications. P. D. Grünwald, I. J. Myung and M. A. Pitt. Cambridge, MA, MIT Press.

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

Gallistel, C. R. (2012). "Extinction from a rationalist perspective." Behavioural Processes **90**: 66-88.

Gallistel, C. R. (2016, in press). The neurobiological bases for the computational theory of mind. In Perspectives on the Work of Jerry Fodor. L. Gleitman and R. G. d. Almeida, Editors. New York, Oxford University Press.

Gallistel, C. R., P. D. Balsam and S. Fairhurst (2004). "The learning curve: Implications of a quantitative analysis." Proceedings of the National Academy of Sciences **101**(36): 13124-13131.

Gallistel, C. R., A. R. Craig and T. A. Shahan (2013). "Temporal contingency." Behavioural Processes **101**: 89-96.

Gallistel, C. R. and J. Gibbon (2000). "Time, rate, and conditioning." Psychological Review **107**(2): 289-344.

Gallistel, C. R., A. King and R. McDonald (2004). "Sources of Variability and Systematic Error in Mouse Timing Behavior." Journal of Experimental Psychology: Animal Behavior Processes **30**(1): 3-16.

Gallistel, C. R. and A. P. King (2010). Memory and the computational brain: Why cognitive science will transform neuroscience. New York, Wiley/Blackwell.

Gallistel, C. R., M. Krishan, Y. Liu, R. R. Miller and P. E. Latham (2014). "The perception of probability." Psychological Review **121**: 96-123.

Gallistel, C. R. and L. D. Matzel (2013). "The neuroscience of learning: Beyond the Hebbian Synapse." Annual Review of Psychology **64**: 169-200.

Gershman, S. J., A. A. Moustafa and E. A. Ludvig (2014). "Time representation in reinforcement learning models of the basal ganglia." Frontiers in Computational Neuroscience **7**: 8.

Gibbon, J., L. Farrell, C. M. Locurto, H. J. Duncan and H. S. Terrace (1980). "Partial reinforcement in autoshaping with pigeons." Animal Learning and Behavior **8**: 45-59.

Gibbon, J., M. D. Baldock, C. Locurto, L. Gold and H. S. Terrace (1977). "Trial and intertrial durations in autoshaping." Journal of Experimental Psychology: Animal Behavior Processes. **3**: 264-284.

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

Gibbon, J., L. Farrell, C. M. Locurto, H. J. Duncan and H. S. Terrace (1980). "Partial reinforcement in autoshaping with pigeons." Animal Learning & Behavior **8**: 45-59.

Gleitman, H., J. Nachmias and U. Neisser (1954). "The S-R reinforcement theory of extinction." Psychological Review **61**(1): 23-33.

Gluck, M. A. and R. F. Thompson (1987). "Modeling the neural substrates of associative learning and memory: a computational approach." Psychological Review **94**(2): 176-191.

Gottlieb, D. A. (2005). "Acquisition with partial and continuous reinforcement in rat magazine approach." Journal of Experimental Psychology: Animal Behavior Processes **31**(3): 319-333.

Gottlieb, D. A. (2008). "Is the number of trials a primary determinant of conditioned responding?" Journal of Experimental Psychology: Animal Behavior Processes **34**(2): 185–201.

Grossberg, S. and N. A. Schmajuk (1991). Neural dynamics of adaptive timing and temporal discrimination during associative learning. Pattern recognition by self-organizing neural networks. G. A. Carpenter and S. Grossberg. Cambridge, MA, MIT Press**:** 637-674.

Grünwald, P. D., I. J. Myung and M. A. Pitt (2005). Advances in minimum description length: theory and applications, (2005). Cambridge, MA, Bradford Books: MIT Press.

Harris, J. A. (2011). "The acquisition of conditioned responding." Journal of Experimental Psychology: Animal Behavior Processes **37**: 151-164.

Hawkins, R. D. and E. R. Kandel (1984). "Is there a cell-biological alphabet for simple forms of learning?" Psychological Review, **91**: 375-391.

Holland, P. C. (1992). Occasion setting in Pavlovian conditioning. Psychology of Learning and Motivation: Advances in Research and Theory. D. L. Medin. New York, Academic Press. 28.

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

Jaynes, E. T. (1957). "Information theory and statistical mechanics." Physical Review **106**(4): 620-630.

Jaynes, E. T. (2003). Probability theory: The logic of science. New York, Cambridge University Press.

Johansson, F., H.A. E., Carlsson, A. Rasmussen, C.H. Yeo and G. Hesslow (2015). "Activation of a Temporal Memory in Purkinje Cells by the mGluR7 Receptor." Cell Reports.

Johansson, F., D.-A. Jirenhed, A. Rasmussen, R. Zucc and G. Hesslow (2014). "Memory trace and timing mechanism localized to cerebellar Purkinje cells." Proceedings of the National Academy of Science **111**: 14930–14934.

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

Kass, M. D., M. C. Rosenthal, J. Pottackal and J. P. McGann (2013). "Fear learning enhances neural responses to threat-predictive sensory stimuli." Science **342**: 1389-1392.

Kheifets, A. and C. R. Gallistel (2012). "Mice take *calculated* risks." Proceedings of the National Academy of Science **109**: 8776-8779.

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

Kullback, S. (1987). "The Kullback-Leibler distance." American Statistician **41**: 340-341.

LeCun, Y., Y. Bengio and G. Hinton (2015). "Deep learning." Nature **521**: 436-444.

Libby, M. E. and R. M. Church (1975). "Fear gradients as a function of the temporal interval between signal and aversive event in the rat." Journal of Comparative and Physiological Psychology **88**(2): 911-916.

Mackintosh, N. J. (1975). "A theory of attention: Variations in the associability of stimuli with reinforcement." Psychol. Rev. **82**: 276-298.

Mark, T. A. and C. R. Gallistel (1994). "Kinetics of matching." Journal of Experimental Psychology: Animal Behavior Processes 20(1): 79-95.

Martin, S. J. and R. G. M. Morris (2002). "New life in an old idea: The synaptic plasticity and memory hypothesis revisited." Hippocampus **12**: 609-636.

Matzel, L. D., T. R. Schachtman and R. R. Miller (1985). "Recovery of overshadowed association achieved by extinction of the overshadowing stimulus." Lng. Motiv. **16**: 398-412.

Mauk, M. and D. Buonomano (2004). "The neural basis of temporal processing." Annual Review Neuroscience **27**: 307–340.

McClaren, I. P. L., H. Kaye and N. J. Mackintosh (1989). An associative theory of the representation of stimuli: Applications to perceptual learning and latent inhibition. Parallel distributed processing: Implications for psychology and neurobiology. R. G. M. Morris. Oxford, Clarendon Press.

Meck, W. H., Ed. (2003). Functional and neural mechanisms of interval timing. New York, CRC Press.

Meck, W. H., Ed. (2003). Functional and neural mechanisms of interval timing. New York, CRC Press LLC.

Menzel, R., J. Fuchs, A. Kirbach, K. Lehmann and U. Greggers (2011). Navigation and Communication in Honey Bees. Honeybee Neurobiology and Behavior. A Tribute to Randolf Menzel. G. C. Galizia, D. Eisenhardt and M. Giurfa. New York, Springer Verlag**:** 103 - 116.

Moser, E. I., E. Kripff and M.-B. Moser (2008). "Place cells, grid cells, and the brain's spatial representationa system." Annual Review of Neuroscience **31**: 69-89.

Mowrer, O. H. and H. M. Jones (1945). "Habit strength as a function of the pattern of reinforcement." Journal of Experimental Psychology **35**: 293-311.

Papachristos, E. B. and C. R. Gallistel (2006). "Autoshaped Head Poking in the Mouse: A Quantitative Analysis of the Learning Curve." Journal of the Experimental Analysis of Behavior **85**: 293-308.

Peikon, I., D. Gizatullina and A. Zador (2014). "In vivo generation of DNA sequence diversity for cellular barcoding." Nucleic Acids Research.

Piccinini, G., & Bahar, S. (2013). “Neural computation and the computational theory of cognition.” Cognitive Science, 37(3), 453-488.

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.

Rescorla, R. A. (1973). "Evidence for "unique stimulus" account of configural conditioning." Journal of Comparative and Physiological Psychology 85(2): 331-338.

Rescorla, R. A., J. W. Grau and P. J. Durlach (1985). "Analysis of the unique cue in configural discriminations." Journal of Experimental Psychology: Animal Behavior Processes 11: 356-366.

Rescorla, R. A. and A. R. Wagner (1972). A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. Classical conditioning II. A. H. Black and W. F. Prokasy. New York, Appleton-Century-Crofts**:** 64-99.

Rissanen, J. (1978). "Modeling by shortest data description." Automatica **14**(5): 465-471.

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

Rissanen, J. (1999). "Hypothesis selection and testing by the MDL principle." The Computer Journal **42**: 260–269.

Rosenthal, S. J. (2001). "Bar-coding biomolecules with fluorescent nanocrystals." Nature Biotechnology **19**: 621-622.

Ryan, T., D. Roy, M. Pignatelle, A. Arons and S. Tonegawa (2015). "Engram cells retain memory under retrograde amnesia." Science **348**: 1007-1013.

Savastano, H. I. and R. R. Miller (1998). "Time as content in Pavlovian conditioning." Behavioural Processes **44**(2): 147-162.

Shannon, C. E. (1948). "A mathematical theory of communicatioin." Bell Systems Technical Journal **27**: 379-423, 623-656.

Smolensky, P. (1986). Information processing in dynamical systems: foundations of harmony theory. Parallel distributed processing: foundations. D. E. Rumelhart and J. L. McClelland. Cambridge, MA, MIT Press. **1:** 194-281.

Sterling, P. and S. B. Laughlin (2015). Principles of Neural Design. Cambridge, MA, MIT Press.

Urushihara, K. and R. Miller (2010). "Backward blocking in first-order conditioning." Journal of Experimental Psychology Animal Behavior Processes **36**: 281-295.

Wagner, A. R. (1981). SOP: A model of automatic memory processing in animal behavior. Information processing in animals: memory mechanisms. N. E. Spear and R. R. Miller. Hillsdale, NJ, Lawrence Erlbaum**:** 5-47.

Wagner, A. R., F. A. Logan, K. Haberlandt and T. Price (1968). "Stimulus selection in animal discrimination learning." Journal of Experimental Psychology **76**(2): 171-180.

Ward, R. D., C. R. Gallistel and P. D. Balsam (2013). "It's the information!" Behavioural Processes **95**: 3-7.

Ward, R. D., C. R. Gallistel, G. Jensen, V. L. Richards, S. Fairhurst and P. D. Balsam (2012). "Conditional stimulus informativeness governs conditioned stimulus—unconditioned stimulus associability." Journal of Experimental Psychology: Animal Behavior Processes **38**(1): 217-232.

Wetmore, D. Z., D.-A. Jirenhed, R. Anders, F. Johansson, M. J. Schnitzer and G. Hesslow (2014). "Bidirectional Plasticity of Purkinje Cells Matches Temporal Features of Learning." The Journal Neuroscience **34**: 1731-1737.

Williams, B. A. (1981). "Invariance in reinforcements to acquisition, with implications for the theory of inhibition." Behaviour Analysis Letters **1**: 73-80.

Yin, H., N. J. Grahame and R. R. MIller (1993). "Extinction of comparator stimuli during and after acquisition: Differential effects on Pavlovian responding." Learning and Motivation **24**: 219-241.

Biographies

Jason Wilkes holds an M.S. in mathematical physics and an M.A. in psychology. His interests include the cognitive psychology of deductive and inductive reasoning, and the computations underlying various forms of inference and decision under uncertainty. Recent work has focused on the representation of scalar magnitude and the computational nature of conditioning phenomena. His first book, Burn Math Class, was recently published by Basic Books. He is currently a graduate student in psychology at the University of California, Santa Barbara.

C. R. Gallistel is Distinguished Professor Emeritus at Rutgers University. He obtained his PhD in Behavioral Neuroscience from Yale University in 1963. He joined the Psychology Faculty at the University of Pennsylvania, where he rose through the ranks to Professor and Chair. Subsequently, he was Distinguished Professor at UCLA and then at Rutgers, where he co-chaired the Center for Cognitive Science. His current research develops highly automated systems for screening genetically altered mice for alterations in basic cognitive functions.