Psychological Review

Contingency, Contiguity, and Causality in Conditioning: Applying Information Theory and Weber's Law to the Assignment of Credit Problem

C. R. Gallistel, Andrew R. Craig, and Timothy A. Shahan

Online First Publication, August 29, 2019. http://dx.doi.org/10.1037/rev0000163

CITATION

Gallistel, C. R., Craig, A. R., & Shahan, T. A. (2019, August 29). Contingency, Contiguity, and Causality in Conditioning: Applying Information Theory and Weber's Law to the Assignment of Credit Problem. *Psychological Review*. Advance online publication. http://dx.doi.org/10.1037/rev0000163

American Psychological Association

Psychological Review

© 2019 American Psychological Association
ISSN: 0033-295X

http://dx.doi.org/10.1037/rev0000163

THEORETICAL NOTE

Contingency, Contiguity, and Causality in Conditioning: Applying Information Theory and Weber's Law to the Assignment of Credit Problem

C. R. Gallistel
Rutgers University

Andrew R. Craig
Utah State University and SUNY Upstate Medical University

Timothy A. Shahan
Utah State University

Contingency is a critical concept for theories of associative learning and the assignment of credit problem in reinforcement learning. Measuring and manipulating it has, however, been problematic. The information-theoretic definition of contingency—normalized mutual information—makes it a readily computed property of the relation between reinforcing events, the stimuli that predict them and the responses that produce them. When necessary, the dynamic range of the required temporal representation divided by the Weber fraction gives a psychologically realistic plug-in estimates of the entropies. There is no measurable prospective contingency between a peck and reinforcement when pigeons peck on a variable interval schedule of reinforcement. There is, however, a perfect retrospective contingency between reinforcement and the immediately preceding peck. Degrading the retrospective contingency by gratis reinforcement reveals a critical value (.25), below which performance declines rapidly. Contingency is time scale invariant, whereas the perception of proximate causality depends—we assume—on there being a short, fixed psychologically negligible critical interval between cause and effect. Increasing the interval between a response and reinforcement that it triggers degrades the retrograde contingency, leading to a decline in performance that restores it to at or above its critical value. Thus, there is no critical interval in the retrospective effect of reinforcement. We conclude with a short review of the broad explanatory scope of information-theoretic contingencies when regarded as causal variables in conditioning. We suggest that the computation of contingencies may supplant the computation of the sum of all future rewards in models of reinforcement learning.

Keywords: reinforcement learning, Pavlovian conditioning, operant conditioning, cue competition, time scale invariance

Supplemental materials: http://dx.doi.org/10.1037/rev0000163.supp

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

Instrumental/operant conditioning is often said to depend on a contingency between the operant response and reinforcement (Balleine & Dickinson, 1998; Schultz, 2015; B. F. Skinner, 1981). That poses the question, what is contingency? It has often been taken to mean temporal pairing. Skinner (1948, p. 168) wrote, "To say that a reinforcement is contingent upon a response may mean nothing more than that it follows the response. It may follow because of some mechanical connection or because of the mediation of another organism; but conditioning takes place presumably because of the temporal relation only, expressed in terms of the order and proximity of response and reinforcement." This formulation shifts the explanatory burden from the concept of contingency to the

C. R. Gallistel, Department of Psychology, Rutgers University; Andrew R. Craig, Department of Psychology, Utah State University, and Department of Pediatrics, SUNY Upstate Medical University; Timothy A. Shahan, Department of Psychology, Utah State University.

Timothy A. Shahan and Andrew R. Craig designed the experiments that provided the data for our theoretical analyses, and Andrew R. Craig ran them. C. R. Gallistel did the analyses and wrote a first draft of the manuscript. All three authors collaborated on the extensive revisions of that manuscript.

The code that implements the analyses is posted on GitHub CRGalliste/Contingency together with the raw data. An analysis of Pavlovian conditioning in information theoretic terms was first given by P. D. Balsam, S. Fairhurst, and C. R. Gallistel, Pavlovian contingencies and temporal information. *Journal of Experimental Psychology: Animal Behavior Processes, 32*, 284–294. The analysis showing that there is no measurable prospective contingency was published in Gallistel et al. (2014).

Correspondence concerning this article should be addressed to C. R. Gallistel, Department of Psychology, Rutgers University, New Brunswick, 252 7th Avenue, 10D, New York, NY 10001. E-mail: galliste@ruccs.rutgers.edu

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

concepts of order and proximity. It does not, however, specify which order is critical nor what constitutes proximity. Is it the prospective order, the order when one looks forward from the response to the ensuing reinforcement? Or is it the retrospective order, the order when one looks backward from reinforcement to the preceding response(s)? And what about proximity? Is it defined in absolute terms or in relative terms? And what constitutes the response: Is it individual instances of the operant behavior (e.g., a pigeon's key peck)? Or is it the rate of the behavior? These are centrally important question in the analysis of operant behavior—and, more generally, in our understanding of associative learning processes and reinforcement learning.

The question of how to make contingency computable is intimately connected to the assignment of credit problem, a problem of central interest in conditioning theory (Staddon & Zhang, 2016), in the development of reinforcement learning algorithms in machine learning (Sutton, 1984), and in the understanding of the neurobiology of the frontal cortices (Chau et al., 2015; Noonan, Chau, Rushworth, & Fellows, 2017). In Pavlovian conditioning, the assignment of credit problem is often called the problem of cue competition (Denniston, Savastano, Blaisdell, & Miller, 2003) because the various cues that precede a reinforcement are conceived of as competing for associative strength. In machine learning, the challenge is to develop algorithms that discover key steps in the paths leading to desired outcomes.

The question of what constitutes contingency gained urgency in conditioning theory when Rescorla (1968) degraded the contingency between the presence of a Pavlovian conditioned stimulus (CS; a tone) and an unconditioned stimulus (US; also called a reinforcement; in this case, a shock to the rat's feet) while keeping their temporal pairing constant. When he eliminated the contingency by delivering shocks at the same rate in the presence and absence of the CS, rats did not develop a conditioned response to the conditional stimulus (CS). He concluded that the contingency between the CS and reinforcement drove the associative process, not their temporal pairing. Further results establishing an important role for contingency have been obtained by many other researchers, using both Pavlovian and instrumental paradigms (Balleine & Dickinson, 1998; Colwill & Rescorla, 1986; Dickinson, Nicholas, & Adams, 1982; Dickinson & Charnock, 1985; Hallam, Grahame, & Miller, 1992; Hammond, 1980; Murphy & Baker, 2004). Although skeptics remain (Papini & Bitterman, 1990), the conclusion that contingency is important now guides research on the neurobiology of associative learning (Jackson, Horst, Pears, Robbins, & Roberts, 2016; Schultz, 2015).

The question of how to measure contingency remains. A central problem is that the measures most often suggested require the construction of a contingency table, one or more of the entries in which are counts of event failures. For example, Schultz (2015) assumes that contingency is measured by Δp. In Pavlovian protocols, this is $p(\mathrm{US}\,|\,\mathrm{CS}) - p(\mathrm{US}\,|\,{\sim}\mathrm{CS})$ and in operant protocols $p(\mathrm{R}\,|\,\mathrm{r}) - p(\mathrm{R}\,|\,{\sim}\mathrm{r})$, where R denotes reinforcement and r denotes response. This measure requires a count of the number of times the CS (or the response) failed to occur. Pearson's coefficient of correlation, which is often suggested as a better metric (Gibbon, 1981; Hallam et al., 1992; J. Gibbon, Berryman, & Thompson, 1974; Scott & Platt, 1985), requires counts of the instances of ${\sim}\mathrm{R}\,|\,\mathrm{r}$, $\mathrm{R}\,|\,{\sim}\mathrm{r}$ and ${\sim}\mathrm{R}\,|\,{\sim}\mathrm{r}$. The application of these metrics has repeatedly forced experimenters and theorists to impose what they have themselves recognized as arbitrary discretizations of time into pseudotrials to obtain the counts of events that did not occur. The choice of a duration for these pseudotrials has a large impact on the resulting measures (Hallam et al., 1992, p. 243f). For discussions of the conceptual and practical difficulties of measuring contingency using contingency tables, see (Gallistel, Craig, & Shahan, 2014; Gibbon, 1981; Gibbon et al., 1974; Hallam et al., 1992; Hammond & Paynter, 1983; Hammond, 1980; Scott et al., 1985). Clearly, before we can treat contingency as causal in conditioning, we need to show that it is an objectively measurable property of conditioning protocols, a measure that is not dependent on an arbitrary discretization of time.

Information-Theoretic Contingency

Here we review the results of applying the information-theoretic definition of contingency to the analysis of operant conditioning protocols. This measure does not require counting events that did not happen or considering their impact on net associative strength. It requires only estimating the parameters of analytic probability distributions (e.g., the exponential distribution) or, when that is not possible, obtaining probabilities from normalized histograms made from the experimental samples. The entropies of the analytic distributions (or of the normalized histograms) yield estimates of the mutual information between the variables of interest. The entropy of a distribution is a parameter just like, say, the variance.

Intuitively, the value of a second variable is contingent on the value of a first variable to the extent that knowledge of the value of the first provides information about the value of the second. In accord with this intuition, an event in Shannon's theory of information communicates information about the timing of another event insofar as knowledge of the time of occurrence of the first event reduces an observer's uncertainty about the timing of the second.

Uncertainty, denoted H, is measured by Shannon's formula: $H = \sum_1^n p_i \log(1/p_i)$, where p_i is the probability of the ith possibility in a set of n mutually exclusive and exhaustive possible values for a variable. The uncertainty about the current value of a variable is the entropy of the probability distribution that variable supports. This quantity is also called the available information because it constitutes the upper limit on how much information can be obtained about the current value of the variable.

The amount of information about a random variable, Y, that can be acquired from knowledge of the value of another random variable, X, is their mutual information, which is denoted by $I(X, Y)$. The information-theoretic contingency of Y on X is their mutual information divided by the source information:

$$C(X, Y) = I(X, Y)/H_{\mathrm{Y}},$$

where H_{Y} denotes the entropy of the Y distribution. It is the fraction of the available information about the value of Y gained by learning the value of X.

Computing the entropy of a distribution is straightforward when dealing with discrete variables for which n is finite. It requires more care when applied to continuous variables. For a continuously distributed variable, such as the exponential distribution of wait times (interevent intervals) generated by a Poisson process, the entropy increases without limit as one partitions the range of wait times more and more finely. This dependence on the granu-

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

larity of the representation of time is a fact of central importance when we consider the role of mutual information in operant conditioning. It forces us to consider what is a psychologically realistic granularity.

We suggest that a psychologically appropriate discretization of durations may be derived from consideration of the dynamic range of the required representation together with knowledge of the Weber fraction. Psychologically and neurobiologically realized representations of empirical quantities have a limited dynamic range (Bavard, Lebreton, Khamassi, Coricelli, & Palminteri, 2018; Herrmann, Schlichting, & Obleser, 2014), which we denote by r, and a limited precision. We denote the inverse of precision by w (for Weber fraction), and we note that $n_b = r/w$. In words, the number of bins in a psychologically realistic partition—the number of psychologically distinguishable values of a continuous variable like wait time—is the dynamic range of the required representation divided by the Weber fraction. We apply this principle to understanding the effects of delay of reinforcement in operant conditioning. The results enable us to distinguish between retrospective contingency and causality as explanations for the effect of reinforcement in operant conditioning.

We make use of psychologically realistic discretizations whenever the entropies of empirical distributions cannot be obtained by analytic formulae. Purely empirical distributions are normalized histograms. They are obtained by partitioning the range of durations in the data into equal-width bins, counting the number of durations in each bin, and dividing the bin counts by the sum of the counts to obtain a discrete probability distribution—to which Shannon's formula is then applied.

Pavlovian Contingency

Reliance on empirical distributions is often unnecessary for the drawing of theoretical conclusions. In simple cases, such Rescorla's (1968) experiments, we can use analytic formulae to measure the information gained from an event such as CS onset that predicts a change in the parameter of a distribution. For example, the entropy (in nats) of an exponential distribution with $\lambda < 1$ is $1 - \ln\lambda$, where $\lambda = 1/\mu$ is the rate parameter of the Poisson process generating an exponential distribution of wait times with expectation μ.

In Rescorla's (1968) experiments, reinforcing events were scheduled by Poisson processes with rate parameters that could differ during periods when the CS was and was not present. We denote the context in which both the CS and the reinforcement occur by C. In Rescorla's experiments, C denotes the experimental chamber (sometimes called the background). The mutual information between CS onsets (denoted $CS\uparrow$) and the waits for reinforcement (denoted R) is the amount by which $CS\uparrow$ reduces the uncertainty in the distribution on R. In the absence of the CS, the subject's uncertainty is the entropy of the Poisson distribution with rate parameter $\lambda_R|\mathrm{C} = 1/\bar{R}|\mathrm{C}$, where $\bar{R}|\mathrm{C}$ is the average wait for reinforcement in that context, that is, in that experimental chamber whether or not the CS is present. After CS onset, it is $\lambda_R|CS\uparrow = 1/\bar{R}|\mathrm{C\&CS}$. The mutual information is the contextual or unconditional uncertainty about the wait times for reinforcement in the experimental chamber minus the conditional uncertainty:

$$\begin{aligned} I(\mathrm{CS}\uparrow,\mathrm{R}) &= H(\lambda_R|\mathrm{C}) - H(\lambda_R|\mathrm{C\&CS}\uparrow) \\ &= (1 - \ln\lambda_R|\mathrm{C}) - (1 - \ln\lambda_R|\mathrm{C\&CS}\uparrow) \\ &= \ln\lambda_R|\mathrm{C\&CS}\uparrow - \ln\lambda_R|\mathrm{C} = \ln\frac{\lambda_R|\mathrm{C\&CS}\uparrow}{\lambda_R|\mathrm{C}}. \end{aligned} \tag{1}$$

We see in Equation 1 that the mutual information is the log of what Balsam, Fairhurst, and Gallistel (2006) term the informativeness of the CS, that is, the ratio between the background rate and the rate following CS onset and before CS offset.) In this case, we do not need to normalize the mutual information because the time unit of the rate parameters appears in the numerator and in the denominator; hence, it cancels out. Computing the mutual information immediately explains Rescorla's essential result: When $(\lambda_R|\mathrm{CS}\uparrow) = (\lambda_R|\mathrm{C})$, that is, when the rate during a CS is the same as the contextual rate, a conditioned response to the CS does not develop because the rate ratio is 1 and $\log(1) = 0$. When the mutual information is 0, the contingency is 0.

We begin with this example of the explanatory power of information-theoretic contingency because it illustrates three important points: First, its computation does not require counting failures of responses or reinforcements to occur; second, attributing causal power to the information conveyed by CS onsets does not require the assumption that an event that did not occur—and whose time of occurrence could not have been anticipated—nonetheless causes something to happen in the brain; third, there are no free parameters in this computation. By contrast, the explanation of this result in the Rescorla and Wagner (1972) theory and its successors (e.g., McLaren & Mackintosh, 2000; Pearce & Hall, 1980) depends on theorist-selected values for three free parameters: the associability of the CS and the US, the associability of the CS and the ~US, and the duration of the pseudotrials into which session time is partitioned. The assumed duration of the pseudotrials, determines the count of failures—how many times reinforcement fails to occur in a given stretch of time. And, finally, reinforcements that do not occur alter net associative strengths in these theories. (For theories with additional free parameters, see Kutlu & Schmajuk, 2012; Luzardo, Alonso, & Mondragón, 2017; Vogel, Ponce, & Wagner, 2019). For the application of contingency-table measures to versions of Rescorla's experiment, see (Hallam et al., 1992.) For a recent review of the representation of time in models of associative learning and in temporal difference models, see Luzardo et al. (2017).

Prospective Operant Contingencies

Work subsequent to Balsam et al. (2006) computed the (information-theoretic) contingency between pecks and reinforcement when pigeons pecked a key on a variable interval (VI) schedule of reinforcement (Gallistel et al., 2014). A VI scheduling program arms the response key at variable intervals following the delivery of the last reinforcement. Whenever the key is armed, the next peck delivers reinforcement.

In Gallistel et al. (2014), the arming intervals were chosen from an approximation to an exponential distribution. The X variable was the distribution of intervals from pecks to the next reinforcement (without regard to intervening pecks). The Y variable was the distribution of intervals to reinforcement from the same number of randomly chosen moments in time. The empirical peck-conditional distribution (the distribution of X) superimposed on the empirical unconditional distribution (the distribution of Y). Thus, the two

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

distributions had the same entropy. The pigeon's uncertainty about when to expect the next reinforcement given it had just made a peck was as great as its unconditional uncertainty, so the prospective contingency between pecks and reinforcement was unmeasurably small.

The same result may be obtained analytically, given the following two considerations. First, the unconditional uncertainty (the source information about wait times for reinforcement) is the entropy of the distribution of interreinforcement intervals (IRIs). In practice, the average interval between pecks is always much shorter than the expectation of the VI. In that case, the distribution of IRIs is the exponential with the expectation of the VI schedule parameter. And second, because pecking is scheduled by a process in the bird that is for practical purposes statistically independent of the process that schedules reinforcements, the pecks constitute randomly chosen moments in time. The memorylessness of the Poisson process (its Markov property) means that the distribution of wait times from randomly chosen moments in time is the same as the distribution of IRIs. This analytic explanation for the lack of a prospective contingency between pecks and reinforcements poses even more strongly the question, why do subjects respond steadily on VI schedules? It cannot be because responses predict rewards (cf. Niv, Daw, Joel, & Dayan, 2007; Niv, Daw, & Dayan, 2005; Schultz, Dayan, & Montague, 1997) unless one can specify in what objectively measurable way they do so.

We next asked whether the pecking was driven by a contingency between the rate of pecking and the interreinforcement interval. (Perhaps from the pigeon's perspective its response is not making a peck, it is pecking at some rate.) To address this question, we measured (information-theoretically) the contingency of IRI on the average interpeck interval (IPI) during that IRI. To make this measure, we fed Shannon's formula probabilities derived from three normalized histograms, one for the distribution of the IPIs observed during the IRIs intervals (the marginal distribution of IPIs), one for the IRIs (the marginal distribution of the IRIs), and one for the joint distribution (the scatter plot of *IRI* vs. *IPI*). The mutual information between an IRI and the corresponding IPI is the entropy of the IRI distribution plus the entropy of the IPI distribution minus the entropy of their joint distribution:

$$I(IPI, IRI) = H_{\mathrm{IPI}} + H_{\mathrm{IRI}} - H_{\mathrm{IPI,IRI}}.$$

In this case, the temporal unit (e.g., minutes, seconds, milliseconds) is an arbitrary scale factor in the measure of mutual information—the smaller the temporal unit, the greater the mutual information. Normalization, dividing by the source information, is required to make contingency independent of the arbitrary choice of a unit of time.

The two-dimensional bins (boxes) within which the points on a scatter plot are counted to obtain the joint distribution are obtained by projecting upward from the IPI axis the bin boundaries for that axis and projecting rightward the bin boundaries for the IRI axis. These intersecting projections of bin boundaries form a lattice superimposed on the scatter plot. The number of bins in this lattice is $m \times n$, where m is the number of bins for the x-axis and n the number for the y-axis. For sample sizes numbering in the hundreds, as ours did, the number of bins in the joint distribution must not become so large that a considerable number of bins have fewer than five counts, because then the plug-in estimates for even the larger p values in the joint distribution become highly susceptible to small sample error (the undersampling problem). We wanted to test whether the choice of the number of bins for our marginal distribution (the values for m and n; hence, for $m \times n$, the number of bins in the joint histogram) had an impact on our estimates of the entropies, hence of the contingency. Therefore, we tested for sensitivity to a choice of partitions by using a crude partition (four IPI bins, five IRI bins) and a finer partition (eight bins on both axes).

A plug-in estimate for a parameter of a distribution is made on the assumption that the values computed from samples (in this case, the entropies computed from the probabilities supplied by the marginal and joint normalized histograms) approximate the values of those parameters for the distributions from which the sample was drawn. Empirical multidimensional distributions are rarely well described by analytic formulae. Thus, their parameters (e.g., their means, their variances, their entropies) are commonly estimated by computing the parameters of sample distributions on the assumption that the values for the parameters of a sample approximate those of the parameters of the distribution from which the sample was drawn. The undersampling problem arises when the samples are so small as to make this assumption doubtful. The number of bins in a joint distribution goes up exponentially with the dimensionality of the joint distribution. Therefore, as a rule of thumb, sample sizes must increase exponentially with the dimensionality of a joint distribution to avoid the undersampling problem in the estimate of its entropy. This technicality is conceptually important when it comes to considering psychologically realistic entropy estimates, as we explain when we come to consider the implications of Weber's Law.

The contingency values from both partitions (4×5 and 8×8) were very similar and very low. The correlations between IRI and IPI were also low. Moreover, they were often negative, that is, higher peck rates (shorter IPIs) were associated with longer interreinforcement intervals. Therefore, we concluded that the steady pecking observed on a VI schedule is probably not driven by the contingency between rate of pecking and the interreinforcement interval (Gallistel et al., 2014). This does not mean that the prospective contingency between rate of pecking and the interreinforcement interval is not important in other contexts. We point out in the concluding section that this contingency explains the well-known difference between rates of responding on variable interval schedules and variable ratio schedules with identical overall rates of reinforcement (Catania, Matthews, Silverman, & Yohalem, 1977; Nevin, Grace, Holland, & McLean, 2001).

Retrospective Contingency

We concluded our previous work on contingency in operant conditioning by noting that the contingencies considered so far are prospective, like the contingencies in Pavlovian conditioning (Gallistel et al., 2014). It has long been believed that operant or instrumental conditioning differs from Pavlovian conditioning in that in operant conditioning reinforcing events must in some sense act backward in time (see Timberlake, 1993 for an extensive discussion of this issue). Retrospective contingencies are an objective feature of the data generated by an operant protocol, which we can measure information theoretically, without counting failures to occur. In animals whose brains can look back in an episodic memory from reinforcing events to the most recent response and

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

stimulus events, retrospective contingencies may be subjectively real as well. We suggest that the computation of this retrospective contingency, or a good proxy for it, drives many aspects of conditioned behavior.

The retrospective contingency between reinforcement and pecks in a VI schedule is 1 because every reinforcement is immediately preceded at a fixed, very short interval (.01 s in our data) by the peck that triggered it. This retrospective reinforcement-conditional distribution has 0 entropy. The entropy of the unconditional distribution is the entropy of the distribution of interpeck intervals because: This distribution is approximately exponential, and the reinforcements constitute randomly chosen points in time. The Markov property of the exponential implies that the distribution looking back from randomly chosen points in time (in this case, looking back from reinforcements) to exponentially distributed events (in this case, pecks) is the same as the distribution looking forward in time. Both distributions are the same as the distribution of IPIs.

We have confirmed this analytic deduction empirically by in fact looking back from randomly chosen points in time to the most recent peck. Thus, the retrospective contingency is the entropy of the unconditional distribution of pecks minus 0 (the entropy of the conditional distribution) divided by the entropy of the unconditional distribution.

The retrospective contingency between reinforcements and responses is 1 in every operant schedule of reinforcement in which responses trigger reinforcements immediately (i.e., at fixed intervals much shorter than the average interval between responses), which is to say almost all schedules that have been experimentally investigated. Thus, it has broadly applicable explanatory power in operant conditioning.

Retrospective contingency propagates backward through a sequence of instrumental actions: There is a retrospective contingency between reward and entering a goal box, and a retrospective contingency between entering the goal box and making the correct choice at the preceding choice point, and so on backward through arbitrarily many choice sequences. Moreover, the dopamine signal has been shown to encode the duration of the retrospective interval from a cue that follows reinforcement at a fixed interval, looking back to the reinforcement that it follows (Fonzi, Lefner, Phillips, & Wanat, 2017). Thus, there is reason to believe that retrospective contingency is an important causal variable in operant/instrumental associative learning paradigms.

The effect of gratis reinforcements on retrospective contingency. There are at least two often-investigated ways to attenuate the perfect retrospective contingency between reinforcement and the response that triggers it in an operant paradigm. The first way is to provide some percentage of the reinforcements gratis. To examine the effects of gratis reinforcements on the pecking rate of pigeons responding on VI schedules, we ran four blocks of sessions. During the first (control) block, 0% of the reinforcements were delivered response independently. This block was followed by a sequence of three blocks of 40 sessions each, in which 67%, 80%, and 100% of the reinforcements were delivered independently of the subject's behavior.

The scheduling algorithm that implements gratis reinforcement is described in the Methods section of the Guide to online supplemental material. The online supplemental material includes the other experimental details relevant to specialists. It also includes a Matlab structure (.mat file) containing the raw data (395 sessions for each of eight birds) and a Matlab script that derived our results and created our figures. This structure may also be downloaded from GitHub along with the TSsystem Toolbox. The GitHub repository (http://www.github.com/CRGallistel/Contingency) contains a 140-page illustrated manual explaining how to use the TSsystem Toolbox.

In the two conditions in which only a portion of the reinforcements were provided gratis, the distributions of intervals from random moments in time back to the preceding peck were well described by exponentials—reflecting the fact that the distributions of interpeck intervals are approximately exponential. The distributions of the intervals from reinforcements back to the preceding pecks were mixtures of the Dirac delta distribution with delay parameter .01 (the trigger interval) and the approximately exponential distribution of interpeck intervals. The mixture parameter was the fraction of the reinforcements presented gratis.

Because both the unconditional distribution of retrospective intervals and the reinforcement-conditional distributions were well described by analytic functions, we used those functions to assign probability densities to our bins rather than relying on the plug-in estimates from normalized histograms. The probabilities assigned to the bins were the bin widths times the corresponding probability densities. These probabilities were fed to Shannon's entropy formula. We used a bin width .011 s because, with that choice of bin width, the probability assigned to the first bin was simply the proportion of response-triggered intervals. The choice of a bin width affects the entropies, but the effect is scalar. The scalar appears in the numerator and denominator of the contingency ratio, so it has no effect on the computed contingency, provided it is as narrow or narrower than the trigger interval.

Figure 1 plots the mean normalized peck rates against the retrospective contingency of each of the eight birds in each of the four conditions (32 points in all). To obtain the dependent variable (normalized mean peck rate), we computed for each bird its mean pecks per minute in each of the four conditions. We removed the pronounced between-subjects variations in these mean pecking rates by dividing them by the pecking rate in the condition with 0% gratis reinforcement, the condition that produced the highest rate. Thus, the normalized mean peck rates in Figure 1 are the mean peck rates in each condition divided by the peck rate in the condition with 0% gratis reinforcement.

It may be seen in Figure 1 that reducing the retrospective contingency between reinforcement and the preceding peck from 1.0 to 0.25 has little effect on the rate of responding, whereas the rate of responding falls off steeply with reductions below this critical degree of contingency. Hallam et al. (1992) also concluded that the effect of contingency on responding in Pavlovian protocols was not linear.

Retrospective contingency and/or causality? Retrospective contingency may also be attenuated by lengthening the interval between the peck that triggers reinforcement and the delivery of that reinforcement. We call this the hang-fire interval (a metaphor drawn from the days of flint-lock musketry). Lengthening the hang-fire interval allows other pecks to occur between the trigger peck and the reinforcement.

Increasing the hang-fire interval raises the question of whether what we term the retrospective contingency between reinforcement and the response that triggers it is not better thought of as the

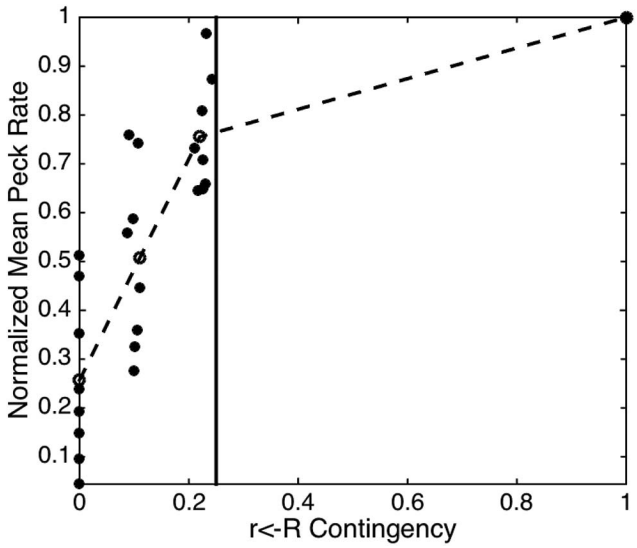

Figure 1. Normalized mean peck rate plotted against the retrospective contingency between reinforcement (R) and response (r) when the contingency is manipulated by gratis reinforcements. The dashed line connects the means (small open circles). The solid vertical line is at .25, at which point the normalized mean peck rate has attained more than 76% of its asymptotic value (range 65–98%). Plotted here are the individual values from the eight birds in the each of the four gratis conditions (32 data points).

perception of causality. The question whether the perception of causality differs from the perception of invariant temporal sequence and/or temporal contingency has roots in philosophy (Aristotle, 350BC/1853; Hume, 1739/2000). It continues in contemporary psychology (Cheng & Novick, 1990; Dickinson, 2001; Michotte, 1963; Miller & Matute, 1996; Shanks & Dickinson, 1987; Vadillo, Miller, & Matute, 2005; Watson, 1997). So long as contingency and causality remain purely intuitive concepts, there may be no way to settle it. In specifying how to compute contingency information theoretically, we have removed it from the realm of the purely intuitive. We see no prospect of settling the question of contingency versus causality unless quantitative constraints on the perception of causality are also specified.

In his classic work on proximate causation, Michotte (1963) gives reason to think that the perception of proximate causality involves a critical interval of fixed duration. In his experiments, one object (or image) collided with another, launching it into a departing motion. Michotte concluded that a psychologically negligible delay between contact and launching is essential to the percept. In one experiment, he determined the critical interval to be 100 ms in one observer and 130 ms in another. If this short interval is in fact fixed (not time scale invariant), then its existence would enable a distinction between the perception of retrospective contingency and the perception of proximate causation.

Looking for a critical, psychologically negligible short interval between response and reinforcement enables us to ask whether the contiguity of response and reinforcement plays a role in addition to contingency, as asserted by Schultz (2015), or in lieu of contingency (Papini & Bitterman, 1990). As Schultz (2015) makes clear, temporal contiguity is another formulation of temporal pairing, the long-standing idea that two stimuli become associated only if they occur adjacently in time, that is, with a psychologically negligible separation. Thus, the contiguity question and the causality question are intimately connected; they both depend on a notion of temporal adjacency between events, which may be taken to mean separation by less than a fixed, psychologically negligible interval.

To assess whether conditioning might depend on a psychologically negligible fixed interval between the reinforcement-triggering response (r) and the triggered reinforcement (R), we varied the trigger latency, the interval during which the reinforcement hangs fire. Over one block of sessions, it was 0.1 s; over another block, 0.4 s; over another block; 1.6 s; and over yet another block, 6.4 s. Each of these four blocks of sessions was repeated with three different VI schedules: VI 15 s, VI 60 s, and VI 240 s.

We denote intervals looking backward from Reinforcement to the preceding response by r ← R. Examples of the distribution these intervals are shown in Figure 2. The histograms in Figure 2 are 10-bin histograms on a linear scale. The bins in these histograms span only the range of the observed intervals. The hang-fire delay sets the upper limit on that range. The counts in the rightmost bins in the histograms in Figure 2—that is, the counts of the longest intervals—come from those instances in which no other peck fell within the hang-fire interval. The counts in the other bins come from instances in which other pecks occurred during that

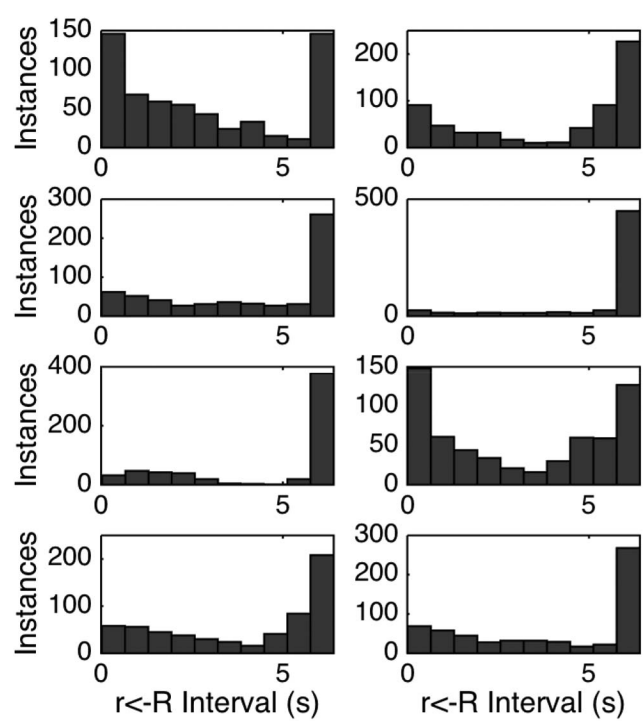

Figure 2. Histograms of the r ← R intervals from the 6.4-s hang-fire interval condition of the VI 240-s sessions, one panel for each subject. Each bin is 6.4/10 = 0.64 s wide. The rightmost bin contains the count for the cases in which there was no peck intervening between the reinforcement-triggering peck and the reinforcement that it triggered. The other bins contain counts from instances in which one or more pecks intervened between the reinforcement-triggering peck and reinforcement delivery.

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

interval. There are pronounced between-subjects differences in the proportion of the total count that falls in the rightmost bins because that proportion depends on the overall rate of pecking and on the extent to which pecks are delivered in bursts. Both parameters (rate and burstiness of pecking) vary markedly between subjects. The higher the proportion in the rightmost bin is, the lower the entropy of the distribution is, hence the greater the retrospective contingency is.

The distributions of r ← R intervals in Figure 2 cannot be described by analytically specified distributions, so we must use the probabilities obtained from normalized histograms. This raises the question of what bin widths we should use in making psychologically realistic histograms. In both neurobiology and technology, information-transmitting mechanisms maximize information transmission by scaling their dynamic range to the range spanned by the values they are currently transmitting (Fairhall, Lewen, Bialek, & de Ruyter van Steveninck, 2001a, 2001b; Maravall, Petersen, Fairhall, Arabzadeh, & Diamond, 2007; Wark, Lundstrom, & Fairhall, 2007). In sensory neurobiology, this automatic scaling is called adaptation (Brenner, Bialek, & de Ruyter van Steveninck, 2000). When confronted with a wider range of stimulus magnitudes, sensory channels reduce their gain so as to compass the wider range. The number of distinguishable signals a channel can transmit over a short interval does not increase as the gain is reduced. Therefore, the spread of stimulus magnitudes that produce the same signal level must increase when the gain in the channel is reduced.

We assume that the neural media that signal experiential magnitudes, including the abstract magnitudes like distance, duration, and numerosity, have a limited dynamic range and a limited resolution and likewise for the memory media that carry previously experienced magnitudes forward in time for use in later computations. We assume that the Weber fraction, w, measures the resolution. In experimental paradigms in which subjects must judge whether a currently experienced magnitude (e.g., the duration of the wait for the most recent reinforcement) falls short of or exceeds some standard, the dynamic range must be adjusted so that the standard falls somewhere within it. It seems most plausible that the dynamic range is adjusted so that the standard falls near the middle, thereby maximizing the ability to measure departures from the standard in either direction. On the assumption that the standard is approximately in the middle of the dynamic range, the number of distinguishable magnitudes is approximately $2/w$ because there will be $1/w$ distinguishable magnitudes below the standard and $1/w$ above it. The Weber fraction for the representation of remembered intervals ranges from about 0.125 to about 0.25 (Church, Meck, & Gibbon, 1994; Gallistel, King, & McDonald, 2004; Gibbon, Malapani, Dale, & Gallistel, 1997), implying eight to 16 distinguishable values within a given range.

A second question raised by the fact that Weber's Law applies to the psychological representation of duration is whether the time scale on which the bins are delimited should be linear or logarithmic. A still popular school of thought, dating back to Fechner (1860), is that Weber's Law implies a logarithmic representation of experiential quantities (Dehaene, 2003; Portugal & Svaiter, 2010).

Weber's Law implies a logarithmic mapping from an objective quantity to its subjective representation only if just noticeable differences in quantity are subjectively equal. In human subjects, just noticeable differences on sensory dimensions are not (Stevens, 1946, 1957). That the same is true for abstract dimensions like duration, distance, and numerosity seems obvious from introspection: For the average subject, durations of 1 and 1.125 s are just discriminable, as are durations of 1 and 1.125 hr. However, no one judges that a wait of one eighth of a second is the same as a wait of one eighth of an hour. The results of experiments designed to test for the consequences of a logarithmic versus a linear scale for subjective duration in pigeons and rats favor the linear hypothesis (Gibbon & Church, 1981). Thus, we believe that the underlying representation of magnitudes is a linear (indeed, approximately scalar) function of objective magnitude because it is in fixed-point computer memory, which memory nonetheless obeys Weber's Law (Gallistel, 2017). However, because controversy continues to surround this issue, we computed contingencies using both eight and 16 equal-width bins on both linear and logarithmic scales of retrospective interval duration.

We divided the range of intervals spanned by the union of the sets of r ← R intervals and the intervals from randomly chosen points in time back to the preceding peck (denoted r←rndT intervals) into eight or 16 equal-width bins (3- or 4-bit precision, respectively), on either a linear or a logarithmic scale, giving four different computations of the contingencies. The formula for obtaining the contingencies was the ratio between $H_{\mathrm{r}\leftarrow\mathrm{rndT}} - H_{\mathrm{r}\leftarrow\mathrm{R}}$ (the mutual information) and $H_{\mathrm{r}\leftarrow\mathrm{rndT}}$. What varied between the computations was the partitioning of the marginal distributions, hence also the partitioning of the joint distributions.

In Figure 3, we plot both the pecking rates (left ordinate) and the linear 3-bit retrospective contingencies (right ordinate) as functions of the hang-fire interval (abscissa). Linear 3-bit means that these contingences were computed from normalized histograms with $2^3 = 8$ equal width bins on a linear time scale. (The abscissa is logarithmic in Figure 3 only to space the experiment values for the hang-fire interval roughly equally.) A contingency of 0.25, which appeared to be a critical value in the experiments with gratis reinforcements, is marked with a horizontal dotted line. It is apparent that at the longer hang-fire intervals (1.6 and 6.4 s between the trigger peck and reinforcement delivery), subjects reduced their pecking rates to values that kept the contingency thus measured at or well above the critical interval.

On the other hand, when the contingencies were computed from normalized histograms with eight bins of equal width on a logarithmic time scale, they were not maintained at or above this critical level; indeed, they were often close to 0 (see Figure 4). The same patterns and the same contrast between the results from the use of a linear time scale and the results from the use of a logarithmic time scale were observed when we assumed 4-bit temporal resolution (computing the contingencies from 16-bin normalized histograms).

Delaying reinforcement degrades the retrospective contingency between reinforcement and the responses that trigger it only temporarily, whereas giving reinforcements gratis degrades it enduringly. The effect of delaying reinforcement is temporary because the following feedback process comes into play: When the contingency falls below the critical value, the response rate declines, in accord with the well-established fact that response rate is an approximately hyperbolic function of delay (Mazur, 1987). The decline in response rate spreads out the distribution of interresponse intervals. The spreading out of that distribution restores the

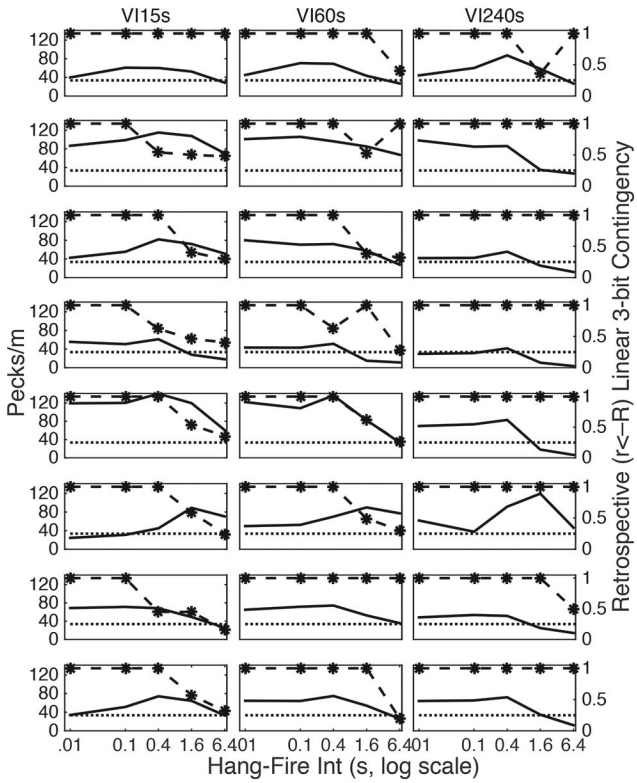

Figure 3. Mean pecking rates are plotted against the left ordinate (solid lines without data point markers) and the linear 3-bit (eight-bin) contingencies are plotted against the right ordinate (dashed lines with asterisk data point markers) as functions of the hang-fire interval. The dotted horizontal line at a contingency of 0.25 indicates the critical value observed in the experiments with gratis reinforcements. Each row is a subject. Each column is for a different VI. The subjects adjusted their response rates to keep the contingency at or above the critical value in all but two cases (in which it fell only slightly below this value).

contingency because it expands the basal time scale, the range spanned by the r←rndT intervals. Thus, feedback from the lowered response rate restores supracritical levels of contingency, as shown in Figure 5.

In the top row of Figure 5, the scale of the time axis spans only the range of r ← R intervals. At this scale, the histogram reveals the fine structure in the distribution of these short intervals. The r ← R intervals are all relatively short because the delay of reinforcement sets the upper limit on them. Note that the time scale in the upper left panel has a maximum of 0.4 s, whereas that in the upper right panel has a maximum of 6.4 s.

In computing contingencies, the scale on which intervals are represented must be the same for the intervals in the numerator (the conditional entropy) as for those in the denominator (the unconditional or basal entropy). The second row of Figure 5 shows the r ← R data, as in the top row, but now on the scale established by the r←rndT intervals from the corresponding conditions. The r←rndT intervals span a much greater range, as shown in the bottom row of Figure 5. In accord with Weber's Law, the precision of the representation of the distributions in Figure 5 is the same at all scales, that is, the same number of bins is used in each histogram. When the time scale, hence the dynamic range of the required representation, is extended to compass the r←rndT intervals, all of the r ← R intervals end up in the lowest bin (middle row of Figure 5). The increased scale of the representation obliterates the fine structure in the distribution of these intervals. The psychological contingency is similar in the two hang-fire delay conditions because the reduction in the pecking rate produced by the increased delay of reinforcement changes the time scale, thereby restoring the contingency to approximately the same level in the long-delay condition that it had been in the short-delay condition.

When the r ← R contingencies are computed after partitioning the experienced intervals into eight or 16 bins that were of equal width on a logarithmic subjective scale of duration, there is no critical contingency. High rates of responding are often observed even when the contingency thus computed is close to 0 (see Figure 4). This is because logarithmic partitioning yields several narrow bins at the very short durations. These are the bins into which the r ← R intervals are counted, so these relatively very short intervals are counted into several different bins. Logarithmic partitioning also widens the bins into which the long intervals are counted, so the longer intervals are counted into fewer bins. That is why the slowing of response rates does not reestablish the computed contingency when logarithmic partitioning is used. We take this failure to reveal the effect of lowered response rates on contingency as further evidence against the Fechnerian conclusion that subjective duration grows as the logarithm of objective duration.

We conclude from these results that what drives responding on VI schedules is the time-scale invariant retrospective contingency between reinforcement and the preceding peck. If the perception of

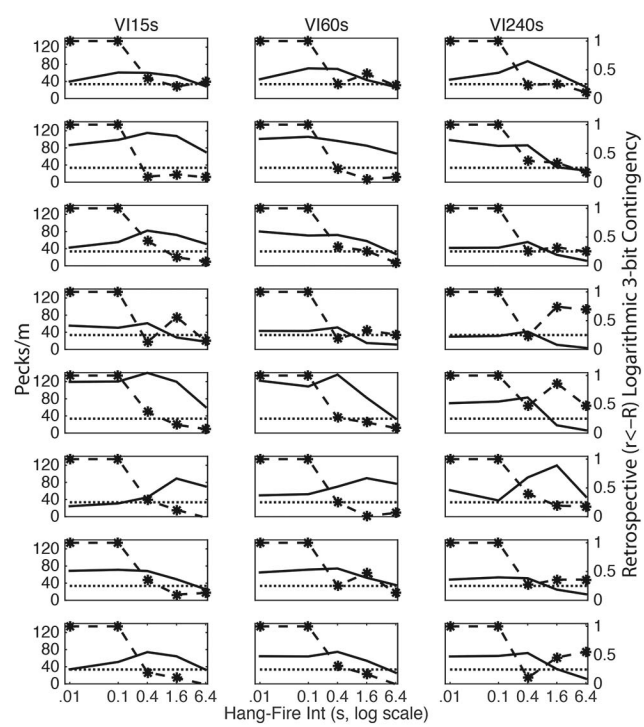

Figure 4. Same plots as in Figure 3, but the contingencies were computed from normalized histograms with eight bins of equal width on a logarithmic time scale.

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

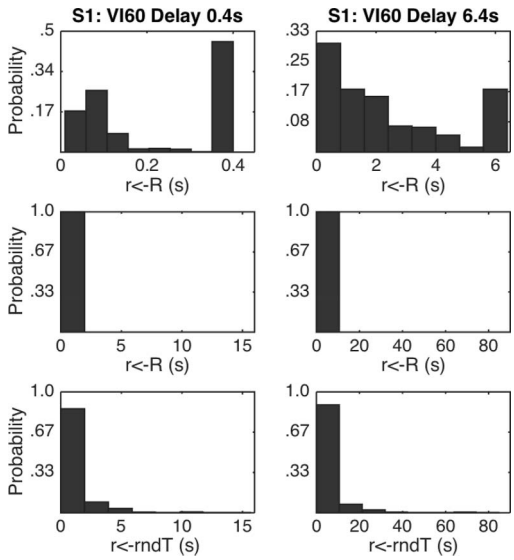

Figure 5. Illustrative probability distributions for the r ← R intervals (top two rows) and the r←rndT intervals (bottom row) from two different delay-of-reinforcement conditions (0.4 and 6.4 s) on a VI 60-s schedule. The middle row portrays the same data as the top row but on the scale established by the data in the bottom row. Note the varying time scales. The precisions are constant because the same number of bins is used in every panel.

proximate causality is equated with the perception of retrospective contingency, then this conclusion may equally well be phrased in terms of the perception of proximate causality. If, however, the perception of proximate causality is assumed to occur only if the event and its cause are separated by an interval less than some fixed small value, a value that does not depend on the time scale established by the pecking rate, then these results favor time-scale invariant retrospective contingency.

The conclusion that operant responding depends on the perception of a time-scale invariant retrospective contingency between reinforcement and response explains why operant responding may be acquired and maintained with delays of reinforcement much longer than those we have employed (Dickinson, Watt, & Griffiths, 1992; Dickinson, Watt, & Varga, 1996; Gleeson & Lattal, 1987; Lattal, Freeman, & Critchfield, 1989; Lattal & Gleeson, 1990; Sizemore & Lattal, 1977; Williams, 1976). Even with a 30-s resetting delay schedule—a condition in which pecks made in the dead interval between the trigger peck and the reinforcement restart the delay—pigeons and rats rapidly acquire sustained key pecking or lever pressing (Lattal & Gleeson, 1990). By contrast, yoked controls that receive the same reinforcements at the same times as the birds or rats to which they are yoked make very few conditioned responses. This important control proves that it is not the provision of reinforcement per se that elicits the operant behavior; it is its provision contingent on a response.

We conclude that the role of time-scale invariant retrospective contingency in operant/instrumental conditioning holds out to the limits that have so far been experimentally explored. These limits substantially exceed what may intuitively be regarded as immediate or contiguous reinforcement. Thus, our results call into question the assumption that contiguity is important in conditioning.

The explanatory scope of contingency. Luzardo et al. (2017) stress the importance of devising models of conditioning that unify the many well-established results in the timing literature with the similarly rich results from the response-acquisition and cue-competition literatures (see also Gallistel & Wilkes, 2016; Wilkes & Gallistel, 2016, 2017). Information-theoretic contingency is rooted in the timing literature because it assumes a metric temporal representation of the intervals between remembered events (Balsam & Gallistel, 2009; J. Gibbon, 1977). We conclude with a brief survey of important role that information-theoretically measured prospective and retrospective contingencies appear to play in acquisition, rate of responding, pattern of responding, and credit assignment.

Ubiquity of retrospective contingency in operant/instrumental conditioning. Retrospective reinforcement-response contingency is perfect in all classic operant schedules. It also emerges in delay of reinforcement (whether the delay is resetting or nonresetting) and differential reinforcement of low rate schedules at well above what we have here shown to be the critical contingency value of 0.25. All of these schedules reduce the rate of responding to the point at which a substantial retrospective contingency emerges. Retrospective contingency may be thought of as the basis for the perception of a causal connection between a response and an outcome but only if this perception is understood to be time scale invariant (not dependent on a fixed, psychologically negligible interval between cause and effect).

Differences in prospective contingency. With variable ratio schedules, there is a scalar relation between rate of responding and rate of reward at all rates of response. Thus, rate of reward is prospectively contingent on rate of responding up to the motoric limit. With variable interval schedules, the contingency between rate of responding and rate of reward is present only so long as the average interval between responses is less than the arming parameter of the VI schedule. It becomes more and more attenuated as the average interresponse interval becomes a shorter and shorter fraction of the arming parameter. This prospective contingency explains why all of our subjects adopted response rates that produced the maximum attainable rate of reinforcement, which is the reciprocal of the arming parameter. The difference in the properties of this prospective contingency under variable interval conditions and variable ratio conditions explains why a variable ratio schedule with the same rate of reinforcement as a variable interval schedule produces a much higher rate of responding (Catania et al., 1977; Nevin et al., 2001). (see Niv et al., 2007 for a model of reinforcement learning that explains this result but with several free parameters. Our explanation is parameter free.)

Patterns of performance in fixed interval and fixed ratio schedules. With a fixed interval schedule, subjects learn that reinforcement is contingent on the elapse of a fixed interval since the last reinforcement, which is why they pause after each reinforcement for an interval that is on average a fixed proportion of the fixed interval parameter (Church et al., 1994; Gibbon, 1977). With a fixed ratio schedule and a fixed number chain schedule with terminal choice links, they learn that reinforcement is contingent on making a fixed number of responses, which is why they respond in runs of more or less continuous very rapid responding, with well-marked run onsets (Berkay, Çavdaroğlu, & Balcı, 2016; Light et al., 2019; Rilling, 1967).

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

Assignment of credit. Contingency also explains the cue-competition/assignment-of-credit results, which led to the models of Rescorla and Wagner (1972) and others (Luzardo et al., 2017; Mackintosh, 1975; Pearce & Hall, 1980; Wagner, 1981). In all of these phenomena, a cue occurs in a context of other cues, within which context the rate predicted by the outcompeted cue is already predicted by that context. Contingency explains them all in the same way it explains Rescorla's (1968) result: When the conditional, that is, cue-signaled, rate of reinforcement is the same as the rate signaled by the context, the contingency is 0 because the mutual information is 0. (For reviews and computational details, see Gallistel, 1990; Gallistel & Gibbon, 2000; Wilkes & Gallistel, 2016.)

Role of the US-US interval in acquisition. Simply computing the mutual information explains the fact discovered by Gibbon and Balsam (1981) that the slope of log-log plot of trials to acquisition versus the log of the ratio between the US-US interval and the CS-US interval in Pavlovian delay conditioning is not significantly different from -1 (Balsam et al., 2006; Gallistel & Gibbon, 2000; Gibbon et al., 1981).

No effect of partial reinforcement on reinforcements to acquisition. It also explains why deleting seven of every eight trials in a Pavlovian delay conditioning paradigm, while preserving the separation between the undeleted trials, has no effect on the progress of conditioning (Gallistel, 2009; Gottlieb, 2008). It also explains the fact that partial reinforcement—interleaving large numbers of unreinforced trials among the reinforced trials—has no effect on the number of reinforced trials required for the appearance of the conditioned response (Gallistel, 2012; Gibbon, Farrell, Locurto, Duncan, & Terrace, 1980).

Role of temporal information in reinforcement learning. Our analyses provide an answer to a question posed by an influential paper on the application of reinforcement learning models to the understanding of instrumental behavior and its neural substrates: ". . . Experiments have yet to suggest where or how the temporal information is constructed and used by the brain. It is not yet clear how far into the future such predictions can be made. . . ." (Schultz et al., 1997, p. 1597): Given our theoretical results, we suggest that: (a) the temporal information in reinforcement learning takes the form of probability distributions defined over wait times, (b) retrodiction is as important as prediction, and (c) both prediction and retrodiction are time scale invariant.

Negative contingency? Information-theoretic contingency is never negative because information itself (the entropy of a distribution) is like the variance of a distribution in that it is always positive or 0. This technical fact accords with the intuition that there is no such thing as negative uncertainty. The brain's uncertainty about the value of a variable may be 0, but it cannot be less than 0. Thus, it is interesting to consider how information-theoretic contingency may be applied to what is termed a negative contingency in a Pavlovian protocol. The simplest such protocol is the inverse of the Rescorla (1968) protocol with which we began: The Poisson process that schedules a random rate of reinforcement runs only when the CS is not present (Kalmbach, Chun, Taylor, Gallistel, & Balsam, 2019; Rescorla, 1969). In this case, it is CS offset that signals the increase in rate—bearing in mind that the contextual rate is $\lambda_R|C = 1/\bar{R}|C$, where $\bar{R}|C$ is the average wait for reinforcement in the experimental chamber (taking no account of whether the CS is or is not present). In this case, CS offset signals the increase in rate, so the relevant measure of contingency would seem to be $\ln\frac{\lambda_R|C\&CS\downarrow}{\lambda_R|C}$. (See Equation (1) for derivation, replacing the up arrow denoting onset with the down arrow denoting offset.) The empirical adequacy of this suggestion remains to be fully tested.

No free parameters. The above itemized results follow directly from computing the mutual information—and normalizing it by the available information when it includes temporal units. There are no free parameters. The number of bins into which a continuous variable like the rate of responding is partitioned in computing plug-in estimates of entropy is a parameter that appears in some computations. Its value, however, is narrowly constrained to between eight and 16 bins by empirical results on the Weber fraction for duration.

No causation by No-US events. Models that explain the above itemized results by appeal to information-theoretic contingency do not attribute causal effects on brain processes to No-USs, events that do not occur and whose occurrence cannot have been anticipated at any specifiable time (Gallistel, 2011, 2012; Gleitman, Nachmias, & Neisser, 1954; Mackintosh, 1975; Pearce & Hall, 1980; Rescorla et al., 1972; Wagner, 1981).

Given its lack of dependence on free parameters and the wide and diverse range of experimental results that it explains, mutual information—or a good proxy for it—seems to be a good candidate for a quantity that is computed by animal brains and that plays a broadly important causal role in both Pavlovian and operant conditioning. It will be interesting to see to what extent it can supplant the computation of the expected value of the sum of future rewards, which is the quantity that plays the central role in temporal difference models of reinforcement learning (McClelland, 2014; Nakahara & Kaveri, 2010; Schultz et al., 1997; Sutton & Barto, 1981; Sutton, 1988).

References

Aristotle. (350BC/1853). *Organon (English translation: The organon, or logical treatises of Aristotle, with the introduction of Porphyry)*. Retrieved from https://archive.org/details/organonorlogica00porpgoog

Balleine, B. W., & Dickinson, A. (1998). Goal-directed instrumental action: Contingency and incentive learning and their cortical substrates. *Neuropharmacology, 37,* 407–419. http://dx.doi.org/10.1016/S0028-3908(98)00033-1

Balsam, P. D., Fairhurst, S., & Gallistel, C. R. (2006). Pavlovian contingencies and temporal information. *Journal of Experimental Psychology: Animal Behavior Processes, 32,* 284–294. http://dx.doi.org/10.1037/0097-7403.32.3.284

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

Bavard, S., Lebreton, M., Khamassi, M., Coricelli, G., & Palminteri, S. (2018). Reference-point centering and range-adaptation enhance human reinforcement learning at the cost of irrational preferences. *Nature Communications, 9,* 4503. http://dx.doi.org/10.1038/s41467-018-06781-2

Berkay, D., Çavdaroğlu, B., & Balcı, F. (2016). Probabilistic numerical discrimination in mice. *Animal Cognition, 19,* 351–365. http://dx.doi.org/10.1007/s10071-015-0938-1

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

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

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

Chau, B. K. H., Sallet, J., Papageorgiou, G. K., Noonan, M. P., Bell, A. H., Walton, M. E., & Rushworth, M. F. (2015). Contrasting roles for orbitofrontal cortex and amygdala in credit assignment and learning in macaques. *Neuron, 87,* 1106–1118. http://dx.doi.org/10.1016/j.neuron.2015.08.018

Cheng, P. W., & Novick, L. R. (1990). A probabilistic contrast model of causal induction. *Journal of Personality and Social Psychology, 58,* 545–567. http://dx.doi.org/10.1037/0022-3514.58.4.545

Church, R. M., Meck, W. H., & Gibbon, J. (1994). Application of scalar timing theory to individual trials. *Journal of Experimental Psychology: Animal Behavior Processes, 20,* 135–155. http://dx.doi.org/10.1037/0097-7403.20.2.135

Colwill, R., & Rescorla, R. (1986). Associative structures in instrumental learning. In G. H. Bower (Ed.), *Psychology of learning and motivation* (Vol. 20, pp. 55–104). New York, NY: Academic.

Dehaene, S. (2003). The neural basis of the Weber-Fechner law: A logarithmic mental number line. *Trends in Cognitive Sciences, 7,* 145–147. http://dx.doi.org/10.1016/S1364-6613(03)00055-X

Denniston, J. C., Savastano, H. I., Blaisdell, A. P., & Miller, R. R. (2003). Cue competition as a retrieval deficit. *Learning and Motivation, 34,* 1–31. http://dx.doi.org/10.1016/S0023-9690(02)00505-2

Dickinson, A. (2001). The 28th Bartlett Memorial Lecture. Causal learning: An associative analysis. *Quarterly Journal of Experimental Psychology, 54,* 3–25. http://dx.doi.org/10.1080/02724990042000010

Dickinson, A., & Charnock, D. J. (1985). Contingency effects with maintained instrumental reinforcement. *Quarterly Journal of Experimental Psychology: Comparative & Physiological Psychology, 37B,* 397–416. http://dx.doi.org/10.1080/14640748508401177

Dickinson, A., Nicholas, D. J., & Adams, C. D. (1982). The effect of the instrumental training contingency on susceptibility to reinforcer devaluation. *Quarterly Journal of Experimental Psychology, 35B,* 35–51.

Dickinson, A., Watt, A., & Griffitths, W. J. H. (1992). Free-operant acquisition with delayed reinforcement. *Quarterly Journal of Experimental Psychology, 45B,* 241–258.

Dickinson, A., Watt, A., & Varga, Z. I. (1996). Context conditioning and free-operant acquisition under delayed reinforcement. *Quarterly Journal of Experimental Psychology, 49B,* 97–110.

Fairhall, A. L., Lewen, G. D., Bialek, W., & de Ruyter van Steveninck, R. (2001a). Adaptation in the fly H1 motion neuron. In T. K. Leen, T. G. Dietterich, & V. Tresp (Eds.), Vol. 13, pp. 124–130). *Advances in neural information processing systems*. Cambridge, MA: MIT Press.

Fairhall, A. L., Lewen, G. D., Bialek, W., & de Ruyter van Steveninck, R. R. (2001b). Efficiency and ambiguity in an adaptive neural code. *Nature, 412,* 787–792. http://dx.doi.org/10.1038/35090500

Fechner, G. (1860). *Elemente der Psychophysik*. Leipzig, Germany: Breitkopf & Härtel.

Fonzi, K. M., Lefner, M. J., Phillips, P. E. M., & Wanat, M. J. (2017). Dopamine encodes retrospective temporal information in a context-independent manner. *Cell Reports, 20,* 1765–1774. http://dx.doi.org/10.1016/j.celrep.2017.07.076

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

Gallistel, C. R. (2009). The importance of proving the null. *Psychological Review, 116,* 439–453. http://dx.doi.org/10.1037/a0015251

Gallistel, C. R. (2011). Contingency in learning. In N. M. Seel (Ed.), *Encyclopedia of the sciences of learning*. New York, NY: Springer.

Gallistel, C. R. (2012). Extinction from a rationalist perspective. *Behavioural Processes, 90,* 66–80. http://dx.doi.org/10.1016/j.beproc.2012.02.008

Gallistel, C. R. (2017). Finding numbers in the brain. *Proceedings of the Royal Society: Series B, 373*(1740), 20170119. http://dx.doi.org/10.1098/rstb.2017.0119

Gallistel, C. R., Craig, A. R., & Shahan, T. A. (2014). Temporal contingency. *Behavioural Processes, 101,* 89–96. http://dx.doi.org/10.1016/j.beproc.2013.08.012

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

Gallistel, C. R., King, A., & McDonald, R. (2004). Sources of variability and systematic error in mouse timing behavior. *Journal of Experimental Psychology: Animal Behavior Processes, 30,* 3–16. http://dx.doi.org/10.1037/0097-7403.30.1.3

Gallistel, C. R., & Wilkes, J. T. (2016). Minimum description length model selection in associative learning. *Current Opinion in Behavioral Sciences, 11,* 8–13. http://dx.doi.org/10.1016/j.cobeha.2016.02.025

Gibbon, J. (1977). Scalar expectancy theory and Weber's Law in animal timing. *Psychological Review, 84,* 279–335. http://dx.doi.org/10.1037/0033-295X.84.3.279

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

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

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

Gibbon, J., & Church, R. M. (1981). Time left: Linear versus logarithmic subjective time. *Journal of Experimental Psychology: Animal Behavior Processes, 7,* 87–107. http://dx.doi.org/10.1037/0097-7403.7.2.87

Gibbon, J., Farrell, L., Locurto, C. M., Duncan, H. J., & Terrace, H. S. (1980). Partial reinforcement in autoshaping with pigeons. *Animal Learning & Behavior, 8,* 45–59. http://dx.doi.org/10.3758/BF03209729

Gibbon, J., Malapani, C., Dale, C. L., & Gallistel, C. (1997). Toward a neurobiology of temporal cognition: Advances and challenges. *Current Opinion in Neurobiology, 7,* 170–184. http://dx.doi.org/10.1016/S0959-4388(97)80005-0

Gleeson, S., & Lattal, K. A. (1987). Response-reinforcer relations and the maintenance of behavior. *Journal of the Experimental Analysis of Behavior, 48,* 383–393. http://dx.doi.org/10.1901/jeab.1987.48-383

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

Gottlieb, D. A. (2008). Is the number of trials a primary determinant of conditioned responding? *Journal of Experimental Psychology: Animal Behavior Processes, 34,* 185–201. http://dx.doi.org/10.1037/0097-7403.34.2.185

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

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

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

Herrmann, B., Schlichting, N., & Obleser, J. (2014). Dynamic range adaptation to spectral stimulus statistics in human auditory cortex. *Journal of Neuroscience: The Official Journal of the Society for Neurosci-*

ence, 34, 327–331. http://dx.doi.org/10.1523/JNEUROSCI.3974-13.2014

Hume, D. (2000) (Original work published 1739). *Of the understanding. Book 1 of A treatise of human nature* (Vol. 1). Oxford, NY: Oxford University Press.

Jackson, S. A. W., Horst, N. K., Pears, A., Robbins, T. W., & Roberts, A. C. (2016). Role of the perigenual anterior cingulate and orbitofrontal cortex in contingency learning in the marmoset. *Cerebral Cortex (New York, N. Y.), 26,* 3273–3284. http://dx.doi.org/10.1093/cercor/bhw067

Kalmbach, A., Chun, E., Taylor, K., Gallistel, C. R., & Balsam, P. D. (2019). Time-scale-invariant information-theoretic contingencies in discrimination learning. *Journal of Experimental Psychology: Animal Learning and Cognition, 45,* 280–289. http://dx.doi.org/10.1037/xan0000205

Kutlu, M. G., & Schmajuk, N. A. (2012). Solving Pavlov's puzzle: Attentional, associative, and flexible configural mechanisms in classical conditioning. *Learning & Behavior, 40,* 269–291. http://dx.doi.org/10.3758/s13420-012-0083-5

Lattal, K. A., Freeman, T. J., & Critchfield, T. S. (1989). Response-reinforcer dependency location in interval schedules of reinforcement. *Journal of the Experimental Analysis of Behavior, 51,* 101–117. http://dx.doi.org/10.1901/jeab.1989.51-101

Lattal, K. A., & Gleeson, S. (1990). Response acquisition with delayed reinforcement. *Journal of Experimental Psychology: Animal Behavior Processes, 16,* 27–39. http://dx.doi.org/10.1037/0097-7403.16.1.27

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

Luzardo, A., Alonso, E., & Mondragón, E. (2017). A Rescorla-Wagner drift-diffusion model of conditioning and timing. *PLoS Computational Biology, 13,* e1005796. http://dx.doi.org/10.1371/journal.pcbi.1005796

Mackintosh, N. J. (1975). A theory of attention: Variations in the associability of stimuli with reinforcement. *Psychological Review, 82,* 276–298. http://dx.doi.org/10.1037/h0076778

Maravall, M., Petersen, R., Fairhall, A., Arabzadeh, E., & Diamond, M. (2007). Shifts in coding properties and information transmission during adaptation. *PLoS Biology, 5,* e19. http://dx.doi.org/10.1371/journal.pbio.0050019

Mazur, J. E. (1987). An adjusting procedure for studying delayed reinforcement. In M. L. Commons, J. E. Mazur, J. A. Nevion, & H. Rachlin (Eds.), *Quantitative analyses of behavior: Vol. 5. The effect of delay and of intervening events on reinforcement* (pp. 55–73). Hillsdale, NJ: Erlbaum.

McClelland, J. L. (2014). Temporal Difference Learnng. In J. L. McClelland (Ed.), *Explorations in parallel distributed processing: A handbook of models, programs, and exercises*. Retrieved from https://web.stanford.edu/group/pdplab/pdphandbook/handbookch10.html

McLaren, I. P. L., & Mackintosh, N. J. (2000). An elemental model of associative learning: I. Latent inhibition and perceptual learning. *Animal Learning & Behavior, 26,* 211–246. http://dx.doi.org/10.3758/BF03200258

Michotte, A. (1963). *The perception of causality*. New York, NY: Basic Books.

Miller, R. R., & Matute, H. (1996). Biological significance in forward and backward blocking: Resolution of a discrepancy between animal conditioning and human causal judgment. *Journal of Experimental Psychology: General, 125,* 370–386. http://dx.doi.org/10.1037/0096-3445.125.4.370

Murphy, R. A., & Baker, A. G. (2004). A role for CS-US contingency in Pavlovian conditioning. *Journal of Experimental Psychology: Animal Behavior Processes, 30,* 229–239. http://dx.doi.org/10.1037/0097-7403.30.3.229

Nakahara, H., & Kaveri, S. (2010). Internal-time temporal difference model for neural value-based decision making. *Neural Computation, 22,* 3062–3106. http://dx.doi.org/10.1162/NECO_a_00049

Nevin, J. A., Grace, R. C., Holland, S., & McLean, A. P. (2001). Variable-ratio versus variable-interval schedules: Response rate, resistance to change, and preference. *Journal of the Experimental Analysis of Behavior, 76,* 43–74. http://dx.doi.org/10.1901/jeab.2001.76-43

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

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

Noonan, M. P., Chau, B., Rushworth, M. F., & Fellows, L. K. (2017). Contrasting effects of medial and lateral orbitofrontal cortex lesions on credit assignment and decision making in humans. *The Journal of Neuroscience*, *37,* 7023–7035. http://dx.doi.org/10.1523/JNEUROSCI.0692-17.2017

Papini, M. R., & Bitterman, M. E. (1990). The role of contingency in classical conditioning. *Psychological Review, 97,* 396–403. http://dx.doi.org/10.1037/0033-295X.97.3.396

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

Portugal, R. D., & Svaiter, B. G. (2010). Weber-Fechner law and the optimality of the logarithmic scale. *Minds and Machines, 21,* 73–81.

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

Rescorla, R. A. (1969). Conditioned inhibition of fear resulting from negative CS-US contingencies. *Journal of Comparative and Physiological Psychology, 67,* 504–509. http://dx.doi.org/10.1037/h0027313

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

Rilling, M. (1967). Number of responses as a stimulus in fixed interval and fixed ratio schedules. *Journal of Comparative and Physiological Psychology, 63,* 60–65. http://dx.doi.org/10.1037/h0024164

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

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

Scott, G. K., & Platt, J. R. (1985). Model of response-reinforcer contingency. *Journal of Experimental Psychology: Animal Behavior Processes, 11,* 152–171. http://dx.doi.org/10.1037/0097-7403.11.2.152

Shanks, D. R., & Dickinson, A. (1987). Associative accounts of causality judgment. In H. B. Gordon (Ed.), *The psychology of learning and motivation: Advances in research and theory* (Vol. 21, pp. 229–261). San Diego, CA: Academic Press, Inc.

Sizemore, O. J., & Lattal, K. A. (1977). Dependency, temporal contiguity, and response-independent reinforcement. *Journal of the Experimental Analysis of Behavior, 27,* 119–125.

Skinner, B. F. (1948). Superstition in the pigeon. *Journal of Experimental Psychology, 38,* 168–172. http://dx.doi.org/10.1037/h0055873

Skinner, B. F. (1981). Selection by consequences. *Science, 213,* 501–504. http://dx.doi.org/10.1126/science.7244649

Staddon, J. E., & Zhang, Y. (2016). On the assignment-of-credit problem

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

in operant learning. In M. L. Commons, S. Grossberg, & J. E. Staddon (Eds.), *Quantitative analyses of behavior series. Neural network models of conditioning and action*. Hillsdale, NJ: Erlbaum.

Stevens, S. S. (1946). On the theory of scales of measurement. *Science, 103,* 677–680. http://dx.doi.org/10.1126/science.103.2684.677

Stevens, S. S. (1957). On the psychophysical law. *Psychological Review, 64,* 153–181. http://dx.doi.org/10.1037/h0046162

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

Sutton, R. S. (1988). Learning to predict by the methods of temporal differences. *Machine Learning, 3,* 9–44. http://dx.doi.org/10.1007/BF00115009

Sutton, R. S., & Barto, A. G. (1981). Toward a modern theory of adaptive networks: Expectation and prediction. *Psychological Review, 88,* 135–170. http://dx.doi.org/10.1037/0033-295X.88.2.135

Timberlake, W. (1993). Behavior systems and reinforcement: An integrative approach. *Journal of the Experimental Analysis of Behavior, 60,* 105–128. http://dx.doi.org/10.1901/jeab.1993.60-105

Vadillo, M. A., Miller, R. R., & Matute, H. (2005). Causal and predictive-value judgments, but not predictions, are based on cue-outcome contingency. *Learning & Behavior, 33,* 172–183. http://dx.doi.org/10.3758/BF03196061

Vogel, E. H., Ponce, F. P., & Wagner, A. R. (2019). The development and present status of the SOP model of associative learning. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology, 76,* 346–374. http://dx.doi.org/10.1177/1747021818777074

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

Wark, B., Lundstrom, B. N., & Fairhall, A. (2007). Sensory adaptation. *Current Opinion in Neuroscience, 17,* 423–429. http://dx.doi.org/10.1016/j.conb.2007.07.001

Watson, J. S. (1997). Contingency and its two indices within conditional probability analysis. *Behavior Analyst, 20,* 129–140. http://dx.doi.org/10.1007/BF03392770

Wilkes, J. T., & Gallistel, C. R. (2016). *Information theory, memory, prediction, and timing in associative learning* (original long version). Retrieved from https://osf.io/y9sg2

Wilkes, J. T., & Gallistel, C. R. (2017). Information theory, memory, prediction, and timing in associative learning. In A. Moustafa (Ed.), *Computational Models of Brain and Behavior* (pp. 481–492). New York, NY: Wiley/Blackwell. http://dx.doi.org/10.1002/9781119159193.ch35

Williams, B. A. (1976). The effects of unsignalled delayed reinforcement. *Journal of the Experimental Analysis of Behavior, 26,* 441–449. http://dx.doi.org/10.1901/jeab.1976.26-441

Received March 7, 2019
Revision received July 8, 2019
Accepted July 15, 2019 ■

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.