Journal of the
Experimental Analysis of Behavior

A NEW VARIABLE INTERVAL SCHEDULE WITH CONSTANT HAZARD RATE AND FINITE TIME RANGE

Mehdi Bugallo[1], Armando Machado[1], and Marco Vasconcelos[2,3]

[1]UNIVERSITY OF MINHO, PORTUGAL
[2]UNIVERSITY OF AVEIRO, PORTUGAL
[3]CESAM, CENTRE FOR ENVIRONMENTAL AND MARINE STUDIES, PORTUGAL

We propose a new variable interval (VI) schedule that achieves constant probability of reinforcement in time while using a bounded range of intervals. By sampling each trial duration from a uniform distribution ranging from 0 to 2 T seconds, and then applying a reinforcement rule that depends linearly on trial duration, the schedule alternates reinforced and unreinforced trials, each less than 2 T seconds, while preserving a constant hazard function.

Key words: exponential distribution, hazard function, uniform distribution, variable interval schedule, temporal contingencies

---

In a Variable Interval (VI) schedule, a reinforcer becomes available after a variable interval elapses since the previous reinforcer or the onset of a stimulus. Ideally, a VI schedule would suppress all time-based reinforcement contingencies while using intervals from a finite range. We present a method to construct such schedule.

The interval of a VI T-s schedule is a random variable whose average defines the schedule parameter, T. Thus, in a VI 60-s schedule, for example, a reinforcer is set up, on average, 60 s after the previous reinforcer or the stimulus onset; the actual interval may be much shorter or much longer than 60 s. To implement a VI schedule, researchers need to construct the distribution of its intervals. The simplest method may be to sample the interval from a range—say, from 0 to 2 T seconds—such that each value within the range is equally likely to occur. The method yields a set of intervals uniformly distributed between 0 and 2 T, with average T; we refer to this schedule as VI_{uni} (for examples of their use see, e.g., Catania & Reynolds, 1968; Ferster & Skinner, 1957).

The main drawback of a VI_{uni} schedule is that it does not suppress temporal contingencies, as longer intervals without a reinforcer are more likely to end with a reinforcer than shorter intervals. To illustrate, if the reinforcer was not set up until the last second, surely (i.e., with probability 1) it will be set up in the last second, whereas if it was not set up for the first second, the probability it will be set up in the next second equals only $1/(2\text{ T}-1)$. More generally, in a VI_{uni} schedule the momentary probability of reinforcement increases with time, and the function relating these two variables, the hazard function, equals $h(t) = 1/(2\text{ T}- t)$.

Another method to construct a VI schedule samples the intervals from an exponential distribution with mean T. We refer to the resulting schedule as VI_{exp}.[1] In contrast with VI_{uni}, in a VI_{exp} the momentary probability of reinforcement does not change with time: If no reinforcer was set up for t seconds, the probability it will be set up in the next second remains constant. The hazard function does not vary with t, $h(t) = 1/\text{T}$. In fact, the

---

This study was conducted at the Psychology Research Centre (UID/PSI/01662/2013) of the University of Minho. The research was supported by the Portuguese Foundation for Science and Technology (FCT) and the Portuguese Ministry of Science, Technology, and Higher Education through national funds and cofinanced by FEDER through COMPETE2020 under the PT2020 Partnership Agreement (POCI-01-0145-FEDER-007653). Mehdi Bugallo was supported by FCT Doctoral Grant PD/BD/114370/2016.

We thank Léo Vandamme for his help with the geometric demonstration.

Correspondence concerning this article should be addressed to: Mehdi Bugallo, School of Psychology, University of Minho, Campus de Gualtar, 4710-057 Braga, Portugal. Email: mehdi.bugallo@gmail.com

doi: 10.1002/jeab.442

1. Millenson's (1963) algorithm uses the geometric distribution as the discrete analogue of the exponential distribution. In this case, the sampled intervals are multiples of a short interval Δt (e.g., 1 s). With the modern computer, we see no advantage of this method over VI_{exp} and therefore do not discuss it further. The same remark applies to Farmer's (1963) method based on the t-τ system.

exponential distribution is the *only* probability distribution with constant hazard function.

In a VI_{exp} schedule, the sampled interval has no upper limit. This feature may be problematic because, as several authors have noted (e.g., Bancroft & Bourret, 2008; Hantula, 1991), very large intervals may have adverse effects on the subject's behavior (e.g., extinction). One way to overcome the problem uses the Fleshler and Hoffman (1962) algorithm: The user enters T, the VI parameter, and N, the number of desired intervals, and the algorithm returns N different intervals. The longest interval lasts for T + T×*ln*(N) seconds, where *ln* stands for the natural logarithm. Shuffled, these intervals approximate a VI_{exp} T schedule. Although the algorithm solves the "longest interval problem", it introduces two new problems. First, because the reinforcers are set up only at N specific moments since the previous reinforcer, the hazard function is no longer constant, in fact it is not possible to limit the interreinforcement interval and preserve a constant hazard function. Second, and relatedly, when used repeatedly for a large number of sessions, the N intervals may occasion temporal discriminations, particularly if N is not large. Increasing N may reduce but does not eliminate the problems.

Here we present another method to construct VI schedules that solves some of the aforementioned problems and opens new avenues for research. The new method eliminates a common feature of traditional VI schedules, namely, that each interval ends with reinforcement or, equivalently, that the distribution of the schedule intervals is identical with the distribution of the reinforced intervals. In the new schedule, there are two types of intervals, randomly mixed, intervals that end with reinforcement (+) and intervals that end without reinforcement (-). Because of this mixture of intervals with and without a reinforcer, a distinct stimulus (e.g., dark chamber) must separate consecutive intervals. Hence, similar to a peak procedure (Catania, 1970; Roberts, 1981), the new schedule consists of a sequence of trials, each comprising an interval of variable length that may or may not be followed by a reinforcer, and separated from the next trial by a distinctly signaled intertrial interval (ITI). Figure 1 illustrates the details.

On each trial, the interval starts with the onset of a stimulus (e.g., the illumination of a key, the insertion of a lever), and ends with the offset of the stimulus (the key light switches off; the lever retracts). The schedule specifies (a) the distribution of the trial intervals and (b) the rule to decide whether to reinforce each interval.

Distribution of Trial Intervals

Let the random variable X represent the interval duration. We assume that X follows a uniform distribution from 0 to 2 T. Figure 2 shows its density function, the horizontal line at 1/(2 T). Geometrically, the density corresponds to the height of the large rectangle with base 2 T; its area equals 1. To determine the value of X for a specific trial, the experimenter would typically sample a random number from 0 to 1 and multiply it by 2 T. The resulting trial intervals will have mean T and standard deviation $T/\sqrt{(3)}$.

Reinforcement Rule

Having selected a trial interval, the experimenter must then decide whether to make a reinforcer available at the end of the interval. The decision is probabilistic and depends on the duration of the interval, *t*. Let $P(+|X = t)$ be the probability of reinforcement given an interval *t*-seconds long, with $0 \le t \le 2$ T. $P(+|X = t)$ must be defined in such a way that the probability of reinforcement in time—the hazard function—remains constant. To state the constraint formally, we define $P(X \approx t, +)$ as the probability that the trial interval ends in the very short period $(t, t + \Delta t)$ *and* the trial is a reinforced trial. The hazard function $h(t)$ is then equal to

$$
\begin{aligned}
h(t) &= P(X \approx t, + | X > t) \\
&= \frac{P(X \approx t, +)}{P(X > t)} \\
&= \frac{P(X \approx t)P(+|X = t)}{P(X > t)} \\
&= \frac{1/(2T)P(+|X = t)}{(2T - t)/(2T)} \\
&= \frac{P(+|X = t)}{2T - t}
\end{aligned}
$$

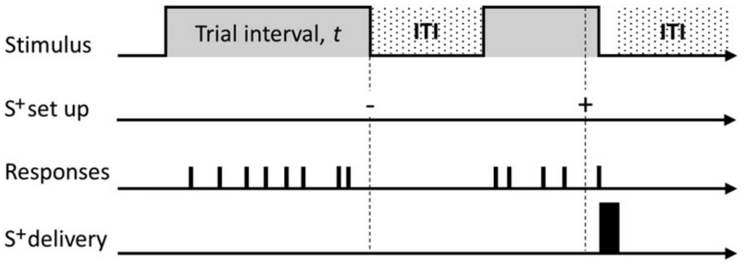

Fig. 1. New VI schedule. Trials start with the onset of a stimulus that remains on for a variable interval, *t*. Unreinforced trials end at *t*. Reinforced trials end with the reinforcer that follows the first response after *t*. An ITI separates consecutive trials.

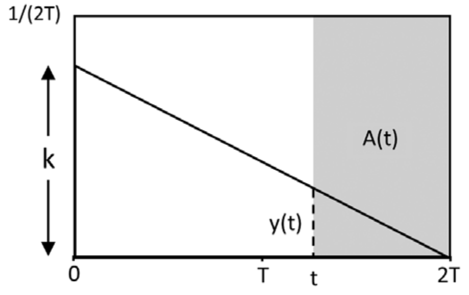

Fig. 2. Geometric representation of the new VI T schedule. Sampled intervals can range from $t = 0$ to $t = 2$ T and are uniformly distributed. For a given *t*, the value y(*t*) is the probability that a trial interval lasts *t* seconds and is reinforced. The area of the shaded rectangle, A(*t*), is the probability that a trial interval lasts more than *t* seconds. The ratio y(*t*)/A(*t*), the hazard function, equals k. The ratio of the areas of the large triangle to the large rectangle is the proportion of reinforced trials, *p*. Restrictions: $0 \le t \le 2$ T; $0 \le k \le 1/(2$ T); $0 \le p \le 0.5$.

The hazard function will be independent of *t* provided

$$P(+|X=t)=k(2T-t) \tag{1}$$

for some constant *k*. Because P(+|X = *t*) is a probability, *k* must be between 0 and 1/(2 T).

The constancy of the hazard function has a useful geometric demonstration. In Figure 2, the shaded rectangle with base (2 T− *t*) and height 1/(2 T) has area A(*t*) = (2 T− *t*)/(2 T), which is the denominator of the hazard function. The height of the small triangle equals y (*t*) = *k*(2 T− *t*)/(2 T), which is the numerator of the hazard function. Because the two right triangles are similar,

$$\frac{k}{2T}=\frac{y(t)}{2T-t},$$

from which it follows that

$$k=\frac{y(t)}{(2T-t)/(2T)} = \frac{y(t)}{A(t)};$$

that is, the ratio y(*t*)/A(*t*), the hazard function, is constant.

The large rectangle with unit area represents all trial intervals. The large triangle with area *k*T represents all reinforced trials. Hence, the ratio of their areas, *k*T, represents the (unconditional) probability of reinforcement per trial, or, more simply, the expected proportion of reinforced trials, *p*. When *k* equals its maximum value of 1/(2 T), the hypotenuse of the large triangle coincides geometrically with the diagonal of the large rectangle and the ratio of the area of the large triangle to the area of the large rectangle is maximal and equal to 0.5; hence the constraint $0 \le p \le 0.5$.

The equality $p = k$T allows us to define the reinforcement rule in a more meaningful way. Letting $k = p/$T, we obtain

$$P(+|X=t)=\frac{p}{T}(2T-t), \text{ with } \begin{cases} 0\le p\le 0.5 \\ 0\le t\le 2T. \end{cases} \tag{2}$$

Equation (2) relates the (conditional) probability of reinforcement at the end of a *t*-s interval to the two schedule parameters, the expected proportion of reinforced trials, *p*, and the average interval duration, T. The ratio *p*/T is the average rate of reinforcement in the presence of the trial stimulus.

Implementing the New VI Schedule in Standard Software Packages

To clarify how the new VI schedule works, we present next an algorithm that implements it. The algorithm requires only a source of random numbers ranging uniformly from 0 to 1 (e.g., the rnd() function in several computer languages).

1. Set the two schedule parameters, T and *p*:
   a. Set the mean trial duration, T (e.g., T = 60 seconds);
   b. Set the probability of reinforcement per trial, *p*, with $0 < p \leq 0.5$, (e.g., $p = 0.4$).
2. Repeat for each trial:
   a. Generate a random number *t* from the uniform distribution (0, 2 T):
      i. Take a random number *u* from 0 to 1 (e.g., $u = 0.1867$);
      ii. Multiply *u* by 2 T to obtain *t* (e.g., $t = 0.1867 \times 120 = 22.4$ s).
   b. Present the stimulus that signals the onset of the trial interval (e.g., turn on a key light);
   c. Run the trial interval for *t* seconds (in the example, wait 22.4 s);
   d. Decide whether or not to reinforce using Equation (2):
      i. Take a random number *r* from 0 to 1;
      ii. If $r < p(2\,T - t)/T$, reinforce; otherwise, do not. In the example, reinforce if $r < 0.4 \times (120\text{-}22.4)/60 \approx 0.651$;
   e. If the trial is not to be reinforced, turn the stimulus off and start the ITI;
   f. If the trial is to be reinforced, wait for a response, and when it occurs turn the stimulus off, deliver the reinforcer, and start the ITI;
   g. At the end of the ITI, check if the session is over. If it is, go to 3 below; if it is not, go to a) above.
3. End the session.

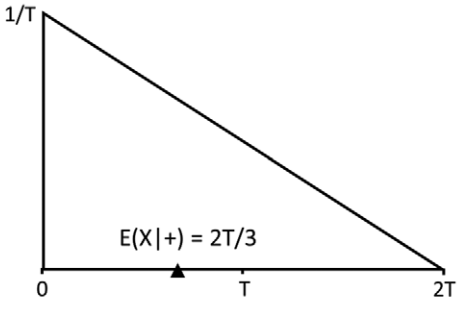

Fig. 3. Density function of reinforced trials, $P(X\approx t|+) = (2\,T - t)/(2T^2)$. The symbol at 2 T/3 shows the mean of the reinforced trials.

Additional Schedule Properties

The reinforced intervals range from 0 to 2 T, but instead of following the uniform distribution as the trial intervals do, they follow a triangular distribution. To characterize this distribution, return to Figure 2 where the hypotenuse of the large triangle shows $P(X\approx t,+)$. If we normalize $P(X\approx t,+)$ by the overall probability of reinforcement, $P(+) = p = k$T, we obtain $P(X = t|+)$, the density function of the reinforced intervals,

$$\begin{aligned} P(X=t|+) &= \frac{P(X\approx t,+)}{P(+)} \\ &= \frac{2T-t}{2T^2} \\ &= \frac{1}{T}\left(1-\frac{t}{2T}\right) \end{aligned} \qquad (3)$$

Figure 3 shows the new (conditional) density function. Straightforward integration shows that the mean of this density function equals 2 T/3 and its standard deviation $\sqrt{2}T/3$. None of these parameters—in fact, none of the moments of the distribution—depends on *p*.

From the facts that the mean of all intervals, E(X), equals T, the proportion of reinforced trials equals *p*, and the mean of the reinforced trials, E(X|+), equals 2 T/3, we derive the mean of the unreinforced trials, E(X|-):

$$\begin{aligned} E(X) &= pE(X|+) + (1-p)E(X|-) \\ T &= p\frac{2T}{3} + (1-p)E(X|-), \end{aligned}$$

from which we obtain

$$E(X|-) = (2T/3)\left(1+\frac{0.5}{1-p}\right) \qquad (4)$$

The mean of the unreinforced intervals ranges from T (when $p = 0$) to 4 T/3 (when

$p = 0.5$). To summarize, as p increases from 0 to 0.5, the mean of the reinforced intervals remains constant at $E(X|+) = 2\,T/3$, but the mean of the unreinforced intervals increases from $1.5 \times E(X|+)$ to $2.0 \times E(X|+)$.[2]

From the previous results, we can also determine the expected time until a reinforcer is set up. To be concrete, suppose we reset variable Y to 0 whenever a reinforcer ends and then add to Y the successive interval durations until we include a reinforced interval. For example, an animal, having just received a reinforcer, experiences three consecutive unreinforced trials with durations t_1, t_2, and t_3 followed by a reinforced trial with duration t_4. In this instance, $Y = t_1 + t_2 + t_3 + t_4$. In other instances, Y would include a different number of intervals. We are interested in the expected value of Y, the total time from the end of a reinforced trial until another reinforcer is set up, excluding the ITIs. To compute the mean of Y, E(Y), we condition on the outcome of the first interval:

$$E(Y) = E(Y|+)p + E(Y|-)(1-p)$$

where $E(Y|+)$ is the expected time to reinforcement if the first interval is reinforced, and $E(Y|-)$ is the expected time to reinforcement if the first interval is not reinforced. If the first interval is reinforced, then Y will equal the mean duration of a reinforced interval, $E(X|+)$. If the first interval is not reinforced, then Y will equal the mean duration of a nonreinforced interval, $E(X|-)$, plus the very mean of Y. The latter follows because, after a nonreinforced trial, the expected time to reinforcement remains equal to E(Y). Hence,

$$\begin{aligned}E(Y) &= E(X|+)p + [E(X|-) + E(Y)](1-p)\\ &= E(X|+)p + E(X|-)(1-p) + E(Y)(1-p)\\ &= E(X) + E(Y)(1-p)\\ &= T + E(Y)(1-p)\end{aligned}$$

which, after solving for E(Y), yields

$$E(Y) = \frac{T}{p} \tag{5}$$

Because $p \le 0.5$, $E(Y) \ge 2\,T$. Obviously, the expected time to a reinforcer is minimal (2 T) when p is maximal (0.5).[3]

In fact, Y follows an exponential distribution with mean T/p (we use moment generating functions to show it formally in the Appendix). The new schedule divides the exponential distribution into unreinforced and reinforced time pieces, all within the range 0 to 2 T, without compromising its constant hazard function.

Comparison with VI_{exp} Schedules

We specify a VI_{exp} schedule with a single parameter, T, which represents both the mean time to food and the mean duration of the schedule intervals. In contrast, to specify the new VI, we need two parameters, T and p, because not all schedule intervals are reinforced intervals.

We propose to use the ratio T/p, the mean time to food in the presence of the trial stimulus, as the nominal value of the new VI. This choice is motivated by two related reasons. First, it facilitates the comparison between the two types of schedules, for a VI_{exp} with parameter T/p may be seen as the equivalent of a new VI with parameters T and p. Second, if we set the ITI to 0 in the new VI, we obtain a VI_{exp} with parameter T/p.[4] Hence, the new VI may be conceived of as a generalization of the traditional VI_{exp}: with an ITI > 0, it is a discrete-trial VI schedule; with an ITI = 0, it is a free-operant VI_{exp} schedule.

The new VI has the same molar feedback function as its VI relatives. Assuming random responding (i.e., a Poisson process yielding x responses on average per second), the mean rate of reinforcement per second in the new VI, $R(x)$, equals.

2. Following the same approach that led to Equation (4) but using density functions (of reinforced, unreinforced, and all trials) rather than mean values, one derives the density of unreinforced intervals: $g(t) = (1/(qT))*((1/2)-p*(1-t/(2\,T)))$, where $q = 1-p$. Geometrically, this line defines a trapezoid with base 2 T, short side $g(0) = (q-p)/(2qT)$ and long side $g(2\,T) = 1/(2qT)$.

3. The expression for E(Y) could be derived also by noting that p/T, as mentioned before, is the reinforcement rate in the presence of the trial stimulus. Hence, its reciprocal T/p is the mean time in the presence of the trial stimulus between consecutive reinforcers.

4. This result is entailed by the fact that the exponential is the only continuous distribution with constant hazard function.

$$R(x) = \frac{1}{\frac{T}{p} + \frac{1}{x}} = \frac{px}{p + xT}$$

the denominator of the first equation shows the average time to collect a reinforcer, the sum of the average time to set up the reinforcer, T/p, and the average time to emit a response, $1/x$, once a reinforcer is set up. The equation defines a hyperbola that starts at 0 for $x = 0$ and asymptotes at p/T, the rate of setting up a reinforcer, as x increases.

As in other interval schedules, in the new schedule the probability a reinforcer is set up increases with time since the last response. Hence, the new VI also differentially reinforces long interresponse times (IRTs). However, because an IRT will never be reinforced during the unreinforced trials, the probability of reinforcement following an IRT t-seconds long cannot exceed the overall probability of reinforcement, p. In the usual VI_{exp}, that probability goes to 1 as t increases. Nonetheless, because VI schedules also reinforce runs of responses and thereby induce short IRTs (Catania & Reynolds, 1968; Appendix 1), it is unlikely that differences in the probability of reinforcement following long IRTs cause significant behavioral differences.

The new schedule combines reinforced and unreinforced intervals according to a rule that yields bounded trial intervals *and* constant probability of reinforcement in time. Its combination rule is not unique, though. An alternative rule combines reinforced intervals sampled from a truncated exponential distribution (i.e., a distribution ranging from $t = 0$ to $t = T_{max}$) with unreinforced intervals of duration T_{max}. The algorithm would be as follows: At trial onset, sample from an exponential distribution with mean time to food equal to T (as in the usual VI_{exp} T). Then a) if the sampled value is less than T_{max}, reinforce at the end of the interval, and b) if the sampled value is greater than T_{max}, end the interval at T_{max} without reinforcement. This "truncated-exponential" VI is similar to the VI proposed above in that it also dissociates reinforced from unreinforced intervals, requires two parameters to be implemented, T and T_{max}, and maintains a constant hazard function from $t = 0$ to $t = T_{max}$. Its main drawback may be the constant duration of all unreinforced intervals. We do not exhaust here the combination rules.

To summarize, the new schedule has similarities and differences with the traditional VI_{exp}. Concerning the similarities, it preserves a constant hazard function, has the same molar feedback function, a hyperbola, and reinforces differentially longer IRTs. Concerning the differences, it dissociates the two distributions "confounded" in VI_{exp}, the distribution of schedule intervals and the distribution of reinforced intervals. The former is rectangular and has mean T; the latter is triangular and has mean 2 T/3. With a 0-s ITI, the new schedule becomes a regular VI_{exp} schedule.

Potential Uses of the New VI Schedule

The fate of any new scientific tool depends on its empirical fruitfulness. We provide five examples of how the new VI schedule may be used to study in novel ways some enduring problems in behavioral analysis.

Resistance to Extinction in the Presence of a Stimulus

Random sampling can yield a long interval in a VI_{exp} schedule, and a large number of consecutive unreinforced intervals in the new VI schedule. In both cases, unbounded interreinforcement intervals (an inevitable consequence of a constant hazard function) pose the risk of extinction. But which schedule engenders greater resistance to extinction, a VI_{exp} with parameter T/p or a new VI schedule with the same nominal value? The former has the keylight (say) present throughout the interreinforcement interval; the latter has the keylight present intermittently during the same interval. In the same vein, of two instances of the new VI with the same nominal value, which one engenders greater resistance to extinction (*cf.*, T = 10 s, $p = 0.2$ versus T = 25 s, $p = 0.5$)? If resistance to extinction depends exclusively on the ratio T/p, the two schedules should not differ; if T and p exert separate effects—perhaps extinction in the presence of a stimulus depends on the maximum duration of the stimulus experienced by the animal—the results should differ.

The Value of Conditioned Reinforcers

The two schedules mentioned above could be associated with a left key (T = 10, $p = 0.2$)

and a right key (T = 25, $p = 0.5$), and then used in a multiple schedule, with the components alternating every 3 min. Once behavior stabilizes, the experimenter could introduce occasional choice trials, with both keys illuminated. If the value of a stimulus depends only on the reinforcement rate in its presence, then the pigeon will be indifferent between the two keys. However, if value depends on, say, the average duration of the reinforced intervals in the presence of a stimulus, (2 T/3 in the new schedule), then the pigeon should prefer the right key, the option with the smaller average. The same issue could be studied with a concurrent chain schedule having the two new VI instances as terminal links.

Contrasting Models of Timing

Scalar Expectancy Theory (e.g., Gibbon, 1977; Gibbon, Church, & Meck, 1984) assumes that, in temporal tasks, only the distribution of reinforced intervals determines behavior, with extinction intervals having no effect. In contrast, the Learning-to-Time model (Machado, 1997; Machado, Malheiro, & Erlhagen, 2009) assumes that both reinforcement and extinction intervals affect behavior. To test the model assumptions, the experimenter could fix T and vary p across conditions, thereby changing the distribution of unreinforced intervals while preserving the distribution of reinforced intervals.

Classical Conditioning

If the reinforcer is made response-independent, the new schedule becomes a variable time (VT) schedule. It may then be used in classical conditioning experiments in which a researcher needs to use a variable duration CS associated with a constant US probability in time. By manipulating T and p, the experimenter may shed further light on how the CR distributes across the CS (cf. Harris, Gharaei, & Pincham, 2011; Kirkpatrick & Church, 2003).

Schedule Performance

Finally, the new schedule may also be used to identify how the various properties of the distribution of reinforced intervals affect the distribution of responses in time (e.g., Catania & Reynolds, 1968; Church & Lacourse, 2001). The new schedule is particularly suitable to this end because none of the moments of the distribution of reinforced intervals changes with T.

Conclusion

The proposed VI schedule is a hybrid of the traditional uniform and exponential VI schedules. It combines the advantages of the two. Similar to a VI_{uni}, its intervals are random samples in the range from 0 to 2 T; similar to a VI_{exp}, its hazard function is constant. Contrary to the Fleshler and Hoffman (1962) algorithm, the new schedule samples a continuous distribution of intervals (it is not limited to N intervals) while eliminating any temporal contingency during the trial stimulus. As a new research tool, it may help researchers address some enduring behavioral issues in novel ways.

References

Bancroft, S. L., & Bourret, J. C. (2008). Generating variable and random schedules of reinforcement using Microsoft Excel macros. *Journal of Applied Behavior Analysis, 41*, 227-235. http://doi.org/10.1901/jaba.2008.41-227

Catania, A. C. (1970). Reinforcement schedules and the psychophysical judgments: a study of some temporal properties of behavior. In W. N. Schoenfeld (Ed.), *The theory of reinforcement schedules* (pp. 1-42). New York: Appleton-Century-Crofts.

Catania, A. C., & Reynolds, G. S. (1968). A quantitative analysis of the responding maintained by interval schedules of reinforcement. *Journal of the Experimental Analysis of Behavior, 11*, 327-383. http://doi.org/10.1901/jeab.1968.11-s327

Church, R. M., & Lacourse, D. M. (2001). Temporal memory of interfood interval distributions with the same mean and variance. *Learning and Motivation, 32*, 2-21. https://doi.org/10.1006/lmot.2000.1076

Farmer, J. (1963). Properties of behavior under random interval reinforcement schedules. *Journal of the Experimental Analysis of Behavior, 6*, 607-616. https://doi.org/10.1901/jeab.1963.6-607

Ferster, C. B., & Skinner, B. F. (1957). *Schedules of reinforcement.* New York: Appleton-Century-Crofts. https://doi.org/10.1037/10627-000

Fleshler, M., & Hoffman, H. S. (1962). A progression for generating variable-interval schedules. *Journal of the Experimental Analysis of Behavior, 5*, 529-530. https://doi.org/10.1901/jeab.1962.5-529

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

Gibbon, J., Church, R., & Meck, W. (1984). Scalar timing in memory. In L. Allan and J. Gibbon (Eds.), *Timing and time perception.* New York: Annals of the New York Academy of Sciences (pp 52–77). https://doi.org/10.1111/j.1749-6632.1984.tb23417.x

Hantula, D. A. (1991). A simple Basic program to generate values for variable-interval schedules of

reinforcement. *Journal of Applied Behavior Analysis, 24,* 799-801. https://doi.org/10.1901/jaba.1991.24-799
Harris, J. A., Gharaei, S., & Pincham, H. L. (2011). Response rates track the history of reinforcement times. *Journal of Experimental Psychology: Animal Behavior Processes, 37,* 277. https://doi.org/10.1037/a0023079
Kirkpatrick, K., & Church, R. M. (2003). Tracking of the expected time to reinforcement in temporal conditioning procedures. *Learning & Behavior, 31,* 3-21. https://doi.org/10.3758/BF03195967
Machado, A. (1997). Learning the temporal dynamics of behavior. *Psychological Review, 104,* 241-265. https://doi.org/10.1037/0033-295X.104.2.241
Machado, A., Malheiro, M. T., & Erlhagen, W. (2009). Learning to time: A perspective. *Journal of the Experimental Analysis of Behavior, 92,* 423-458. https://doi.org/10.1901/jeab.2009.92-423
Millenson, J. R. (1963). Random interval schedules of reinforcement. *Journal of the Experimental Analysis of Behavior, 6,* 437-443. https://doi.org/10.1901/jeab.1963.6-437
Roberts, S. (1981). Isolation of an internal clock. *Journal of Experimental Psychology: Animal Behavior Processes, 7,* 242-268. https://doi.org/10.1037/0097-7403.7.3.242
Ross, S. (2010, 8th Ed.). *A first course in probability.* London: Pearson.

Received: November 4, 2017
Final Acceptance: May 11, 2018

Appendix

Let Y stand for the interreinforcement interval, excluding the ITIs. The claim that Y follows an exponential distribution may be inferred from the general principle that the exponential distribution is the only distribution with constant hazard function. Given that during the trial stimulus of the new schedule the hazard function is constant, it follows from the general principle that if we add the trial intervals between reinforcers, the result (i.e., Y) must follow the exponential distribution.

Another proof that Y is exponential uses the method of moment generating functions (mgf).

The mgf of the reinforced intervals, X^+, equals

$$M_{X^+}(s) = E\left(e^{sX^+}\right) = \int_0^{2T} \frac{1}{T}\left(1-\frac{t}{2T}\right) e^{st}\,dt = \frac{e^{2sT}-2sT-1}{2s^2T^2}, \qquad \text{(A1)}$$

and the mgf of the unreinforced intervals, X^-, equals

$$M_{X^-}(s) = E\left(e^{sX^-}\right) = \int_0^{2T} \frac{1}{T}\left(\frac{1-2p}{2q}+\frac{p}{q}\frac{t}{2T}\right) e^{st}\,dt = \frac{e^{2sT}(sT-p)-sT(1-2p)+p}{2qs^2T^2}, \qquad \text{(A2)}$$

where $q = 1-p$.

In the new VI schedule, the interreinforcement interval (excluding the ITI) consists of a sum of a random number (possibly 0) of unreinforced intervals, X^-, and one reinforced interval, X^+. That is,

$$Y = X_1^- + X_2^- + \cdots + X_N^- + X^+ = \sum_{i=1}^{N}\left(X_i^-\right) + X^+ \qquad \text{(A3)}$$

where N is a random variable, independen
t of the sequence of X's, that takes the values 0, 1, 2,... according to the geometric distribution,

$$P_N(n) = pq^n, \quad (n = 0,1,\ldots). \qquad \text{(A4)}$$

To compute the moment generating function of Y, we use a well-known property concerning the expected values, namely, that E(Z), the expected value of the random variable Z, can be computed by conditioning on the value of the random variable H: E(Z) = E(E(Z|H)) (see, e.g., Ross, 2010). In the present case, this yields

$$E\left(e^{sY}\right) = E\left(E\left(e^{sY}\mid N\right)\right)$$

To compute the inner conditional expectation in the right hand side of the preceding equation, we note that all random variables included in Y (i.e., X^+ and X_i^-) and N are independent. Hence,

$$E\left(e^{sY}\mid N=n\right) = E\left(\exp\left\{sX^+ + s\sum_{i=1}^{n} X_i^-\right\}\mid N=n\right)$$

$$= E(\exp\{sX^+\})E\left(\exp\left\{s\sum_{i=1}^{n} X_i^-\right\}\right)$$
$$= M_{X^+}(s)[M_{X^-}(s)]^n.$$

More generally then,

$$E\left(e^{sY}\right) = E\left(E\left(e^{sY} \mid N\right)\right)$$
$$= E\left\{M_{X^+}(s)[M_{X^-}(s)]^N\right\}$$
$$= M_{X^+}(s)E\left([M_{X^-}(s)]^N\right). \qquad \text{(A5)}$$

The first factor in Equation (A5) is given directly by Equation (A1). To compute the second factor, involving an expectation over N, we use Equations (A2) and (A4) to obtain

$$E\left([M_{X^-}(s)]^N\right) = p\sum_{n=0}^{\infty} q^n \left(\frac{e^{2sT}(sT-p) - sT(1-2p) + p}{2qs^2T^2}\right)^n$$
$$= p\sum_{n=0}^{\infty} \left(\frac{e^{2sT}(sT-p) - sT(1-2p) + p}{2s^2T^2}\right)^n$$

For $0 < s < p/T$, the sum of the geometric series converges and therefore

$$E\left([M_{X^-}(s)]^N\right) = p\frac{2s^2T^2}{2s^2T^2 - e^{2sT}(sT-p) + sT(1-2p) - p}$$

By factoring the denominator, the preceding expression becomes

$$E\left([M_{X^-}(s)]^N\right) = p\frac{2s^2T^2}{(e^{2sT} - 2sT - 1)(p - sT)} \qquad \text{(A6)}$$

Finally, multiplying Equation A6 by the mgf of X^+ yields the mgf of Y,

$$M_Y(s) = M_{X^+}(s)E\left([M_{X^-}(s)]^N\right)$$
$$= \left(\frac{e^{2sT} - 2sT - 1}{2s^2T^2}\right)p\frac{2s^2T^2}{(e^{2sT} - 2sT - 1)(p - sT)}$$
$$= \frac{p}{p - sT}$$
$$= \frac{1}{1 - s(T/p)} \qquad \text{(A7)}$$

We recognize in Equation (A7) the mgf of an exponential distribution with mean T/p.