

MOLECULAR ANALYSES OF THE PRINCIPAL
COMPONENTS OF RESPONSE STRENGTHPETER R. KILLEEN, SCOTT S. HALL,
MARK P. REILLY, AND LAUREN C. KETTLE

ARIZONA STATE UNIVERSITY

Killeen and Hall (2001) showed that a common factor called *strength* underlies the key dependent variables of response probability, latency, and rate, and that overall response rate is a good predictor of strength. In a search for the mechanisms that underlie those correlations, this article shows that (a) the probability of responding on a trial is a two-state Markov process; (b) latency and rate of responding can be described in terms of the probability and period of stochastic machines called *clocked Bernoulli modules*; and (c) one such machine, the refractory Poisson process, provides a functional relation between the probability of observing a response during any epoch and the rate of responding. This relation is one of proportionality at low rates and curvilinearity at higher rates.

Key words: IRT distributions, latency, models, probability, rate, pigeons, rats

Skinner often asserted that rate was a measure of the probability of responding, and that “the task of an experimental analysis is to discover all the variables of which probability of response is a function” (Skinner, 1988, p. 214). This article takes up that task. Its strategy is not to seek new independent variables but to understand the relation between probability of response and the other dependent variables in our field: local rates, global rates, and latencies. It continues the analyses of response strength initiated by Killeen and Hall (2001). In their experiments, the probability, latency, and run rates of pigeons’ pecks were measured in trials procedures. Principal components analyses identified a factor called *strength* that was common to the three dependent variables. Overall response rate, a composite of these three components, was highly correlated with the common factor. By what mechanism might a common state give rise to correlations among these dependent variables? That is the question pursued in this paper. The tactic is to develop appropriate statistical models and machines that exemplify those models. Machines are constructed that will generate the probability of responding during a trial (a two-state Markov model), the latency of the

first response (various series-latency mechanisms), and response rate (a periodic process with noise). This is done because the machines epitomize the aspects of behavior captured by the equations, and it is easier to understand the machines than the equations by themselves. The flow of argument is data → probability models → machines that exemplify those models → probability as a fundamental variable → response rate as its measure. The unifying construct is probability.

Johnson and Morris (1987) reviewed the use of *probability* in the experimental analysis of behavior, and agreed with Dews that “Despite numerous attempts, rate of responding has yet to be converted in a general and rigorous way into a probability” (Dews, 1981, p. 116). Johnson and Morris suggested replacing *probability* with the more general term *propensity*. Propensity is essentially the way that Skinner, and we, use the term *strength*: a state of the organism that is revealed in certain measures (e.g., rate, probability, latency, and force of responding) given certain contexts (e.g., deprivation levels, experimental context, history of reinforcement). *Propensity* was an appropriately vague term given our state of knowledge 20 years ago. The goal of this paper is to strengthen the analyses of that key variable rather than weaken its name: Rate of responding will be converted into probability in a general and rigorous way. Johnson and Morris’ *propensity* still has a role—not as a kind of fuzzy probability, but as a generic reference to a moderate level of response strength. In its turn, strength is seen as a state

This research was supported by NSF Grant IBN 9408022 and NIMH Grant K05 MH01293. Scott Hall is now at the School of Psychology, University of Birmingham, England.

Reprints may be obtained from Peter Killeen, Department of Psychology, Box 1104, Arizona State University, Tempe, Arizona 85287-1104 (e-mail: killeen@asu.edu).

variable that gives rise to the key dependent variables through the mechanisms described here.

EXPERIMENTS FROM KILLEEN AND HALL (2001)

The methods involved in these experiments are summarized; see the original for additional details.

General Method

Four adult homing pigeons (*Columba livia*) were maintained at 80% of their free-feeding weights. A standard BRS/LVE operant chamber had a centrally located Gerbrands response key, background illumination from a houselight, and ambient noise from a speaker.

Sessions ended after 200 trials. Prior to each trial the chamber was dark for 9 s, followed by a 1-s warning stimulus (either illumination of red side-keylights or flickering of houselight). Trials began with the illumination of the response key and ended after 10 s, or after reinforcement (3.2-s access to milo grain) if that occurred first.

Variable-interval (VI) schedules and extinction. After pretraining, reinforcers were available according to one of four constant-probability VI schedules: VI 120 s, VI 240 s, VI 480 s, and VI 960 s. Blocks of 6 to 14 VI sessions alternated with blocks of four to six sessions of extinction. In extinction the hopper was empty. This was reported as Phase 2 of Experiment 1 in Killeen and Hall (2001).

Variable-ratio (VR) schedules, satiation, and extinction. Experimental sessions were conducted on alternate days, allowing the pigeons to satiate during the course of the session and return to running weights of approximately 84% ad libitum for the next session. Sessions lasted for 2 hr, and consisted of up to 360 trials. Trials were scheduled to last for 10 s, during which responses were reinforced on a VR schedule with a probability of 1/20, but were terminated after reinforcement. This condition lasted for 10 sessions, followed by one session of extinction. These 11 sessions were repeated three times and reported as Experiment 2 in Killeen and Hall (2001).

Fixed-interval (FI) schedules. Sessions lasted for 2 hr and were conducted on alternate days. The pigeons were trained on an FI 20-

s schedule that included a limited hold of 10 s: The first response to the left key after 20 s had elapsed from trial onset was reinforced; if no response had occurred by 30 s from trial onset, the trial ended. This condition lasted for 15 sessions, followed by a session of extinction. The pigeons were retrained on FI 20 s for 10 sessions. In this condition trials continued until a reinforcer was collected, or the session terminated after 120 min. A final extinction session was conducted. This was reported as Experiment 3. In all cases an inter-trial interval of 10 s followed trial termination.

Analyses. All analyses were based on the last five sessions of each condition. The probability of responding on a trial (p) was calculated by dividing the number of trials during which at least one response occurred by the total number of trials in a session. The mean latency (L) of the first response was calculated by summing the total time that elapsed before the first response and dividing this by the number of trials. Trials without a response were omitted from the latency analysis. These are therefore latencies given that a response occurred; latencies long enough to exceed trial duration were not recognized. Latencies were converted into a complementary measure, the proportion of the trial duration spent responding, $1 - L/T$, where L is the latency and T the trial duration (usually 10 s). This transformation gave changes in all of the dependent variables the same directional change with changes in response strength. The running rate (b) was calculated by taking the reciprocal of the average of interresponse times (IRTs) after the first response. The overall rate (B) was calculated by summing the total number of responses made over a session and dividing by the total time available for responding.

Summary Results

Figure 1 shows the key variables in the different experiments, averaged over pigeons and replications. They vary systematically in the expected directions with changes in the rate of reinforcement, satiation, and extinction. The measures shown in Figure 1 are conditional: the probability of a response given a trial; proportion of a trial spent responding given that a response occurred on that trial; rate of responding given that a response had occurred; response rate given that the

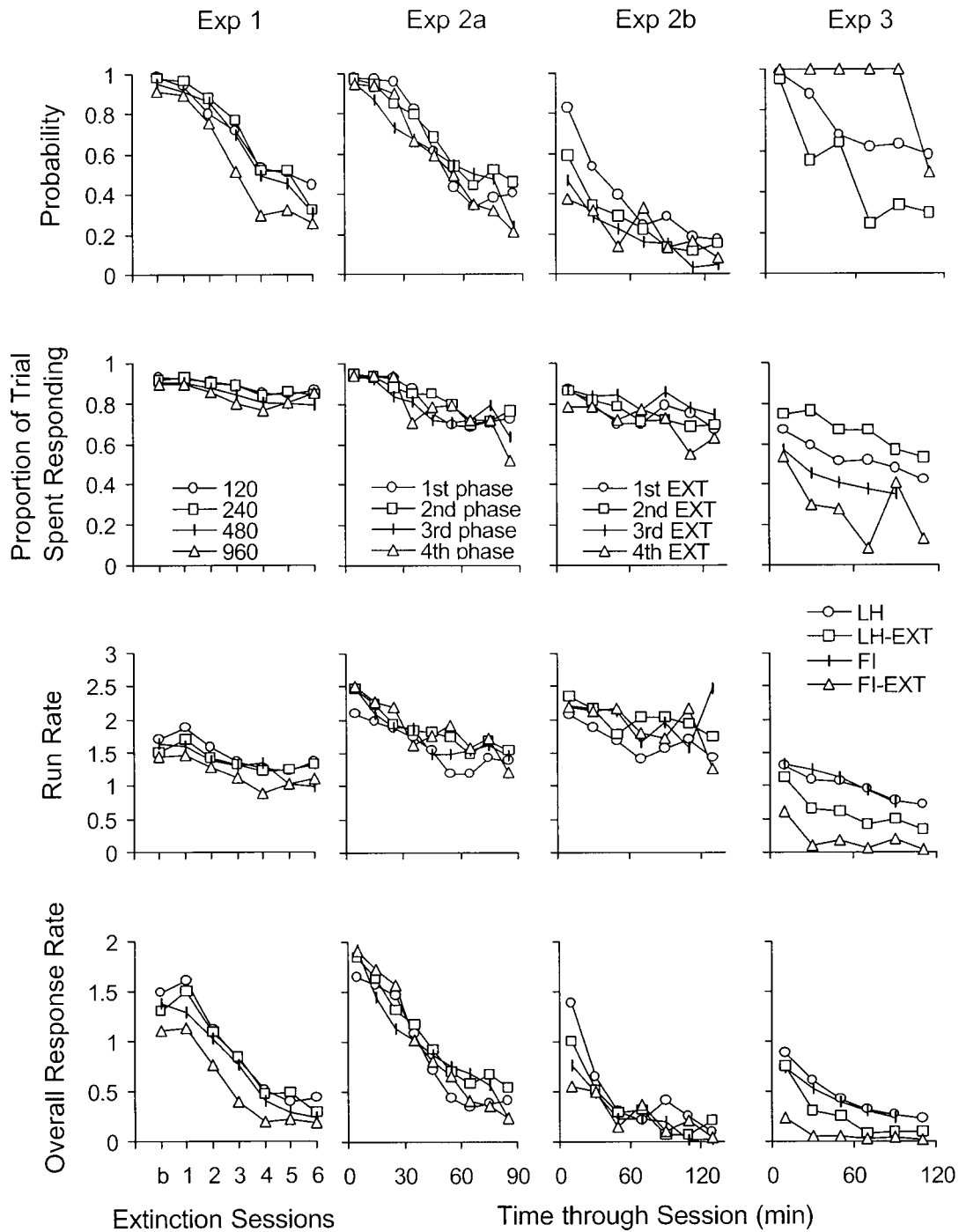


Fig. 1. Changes in the principal measures of the response averaged across 4 pigeons in Experiments 1 through 3 of Killeen and Hall (2001). The top row gives the probability of responding on any trial (p); the second gives the proportion of a trial in the response state (the complement of the latency, relative to the average trial duration: $1 - L/T$); the third gives the running rate (rate after the first response: b); the fourth gives overall response rate (total number of responses divided by the number of seconds available for responding: B). The first column shows baseline followed by successive sessions of extinction on VI schedules whose means are given in the legend. The second column shows satiation on VR 20; the third column shows multiple extinctions after that schedule. The fourth column shows performance on FI 20 s and FI 20 s with a 10-s limited hold over the course of satiation and extinction.

key was lit. Killeen and Hall (2001) showed that this ensemble of measures, when disturbed with varying rates of reinforcement, contingency of reinforcement, satiation, and extinction, were sufficiently well correlated (median $r = 0.92$) that it was economical to posit an underlying state variable—*strength*. This variable summarizes the effects of these operations on the organism. Being a composite measure, it is more robust and informative than any one of the measures by itself. Confirmatory factor analyses validated this claim and showed furthermore that overall response rate was an excellent predictor of that common factor of strength, having a factor loading of 1.0 on it for the group data. The probability of responding on a trial had a loading of 0.98, almost as good.

Whereas the principal components analyses underscored the functional relation among dependent variables as mediated by strength, it did not provide a molecular analysis of the workings of the dependent variables. That is provided here. In particular, it will be shown that even though overall response rate is often the best measure of strength, the best way to think about strength is as a probability—the probability of making a response conditional on the onset of a stimulus, or on the passage of time from that stimulus, or on the passage of time since the last response. Those probabilities are realized in a series of machines. This treatment permits a definitive statement of the relation between response rate and probability in the context of well-defined probability machines.

STOCHASTIC MODELS OF BEHAVIOR

Response Rate

Response rate is the number of responses in an epoch divided by the duration of that epoch. It is also a parameter of a distribution: It is the reciprocal of the average IRT. The first step of analysis is to inspect the distribution of the things averaged. Distributions require categorizations (bins), so that the frequency of different categories may be displayed. Figure 2 shows the distribution of IRTs at various “grains”—different sized categories of width Δ seconds, from the VI schedules reported as Phase 2 of Experiment

1 in Killeen and Hall (2001). The categories are placed above their midpoints, at abscissae of 0.5Δ , 1.5Δ , 2.5Δ , and so on. Multiplying these times by the frequencies displayed above them and summing gives the average IRT, the reciprocal of the average rate. All IRTs within a range of Δ seconds are considered elements of the same category. To further describe the distributions requires models, the simplest of which is the geometric progression.

Geometric distributions. A very coarse display of IRTs is shown in the top row of Figure 2, where $\Delta = 1,000$ ms. The data are from the last five baseline sessions of the VI 960-s condition. The observed decreases in the relative frequencies of emitting IRTs as a function of their length may be succinctly described as *geometric*. Geometric distributions occur when there is some probability, p , of a response at each unit of time (each bin on the x axis). A response will fall into the first bin with probability p , and miss it with probability $1 - p$. If it misses the first, it will fall into the second bin with probability p . The chance of both events happening is $p(1 - p)$. The chance of missing the first two bins and falling into the third is $p(1 - p)(1 - p)$. The chance of missing the first $n - 1$ bins and falling into the n th is $p(1 - p)^{n-1}$. This is a geometric distribution.

The mean of a geometric distribution is $\Delta/p - \Delta/2$. The subtrahend $\Delta/2$ situates the bins over their midpoints. Geometric densities account for more than 99.9% of the variance in the frequencies of the four coarse-grained IRT distributions shown in the top row of Figure 2. Geometric distributions are simple because the probability of an event in any epoch, given that the organism has gotten there, is the same for all epochs: It is p . The conditional probability of a response occurring in bin n , given that it had not yet occurred by bin $n - 1$, is p , which is true for any value of n . This “memoryless” character is unique to the geometric distribution and its continuous analogue, the exponential density. It indicates that knowledge of the time since the last response gives us no advantage in predicting the occurrence of the next response: To the extent geometric or exponential functions describe the distributions, time since the last response is not a causal variable. Probabilities conditional on

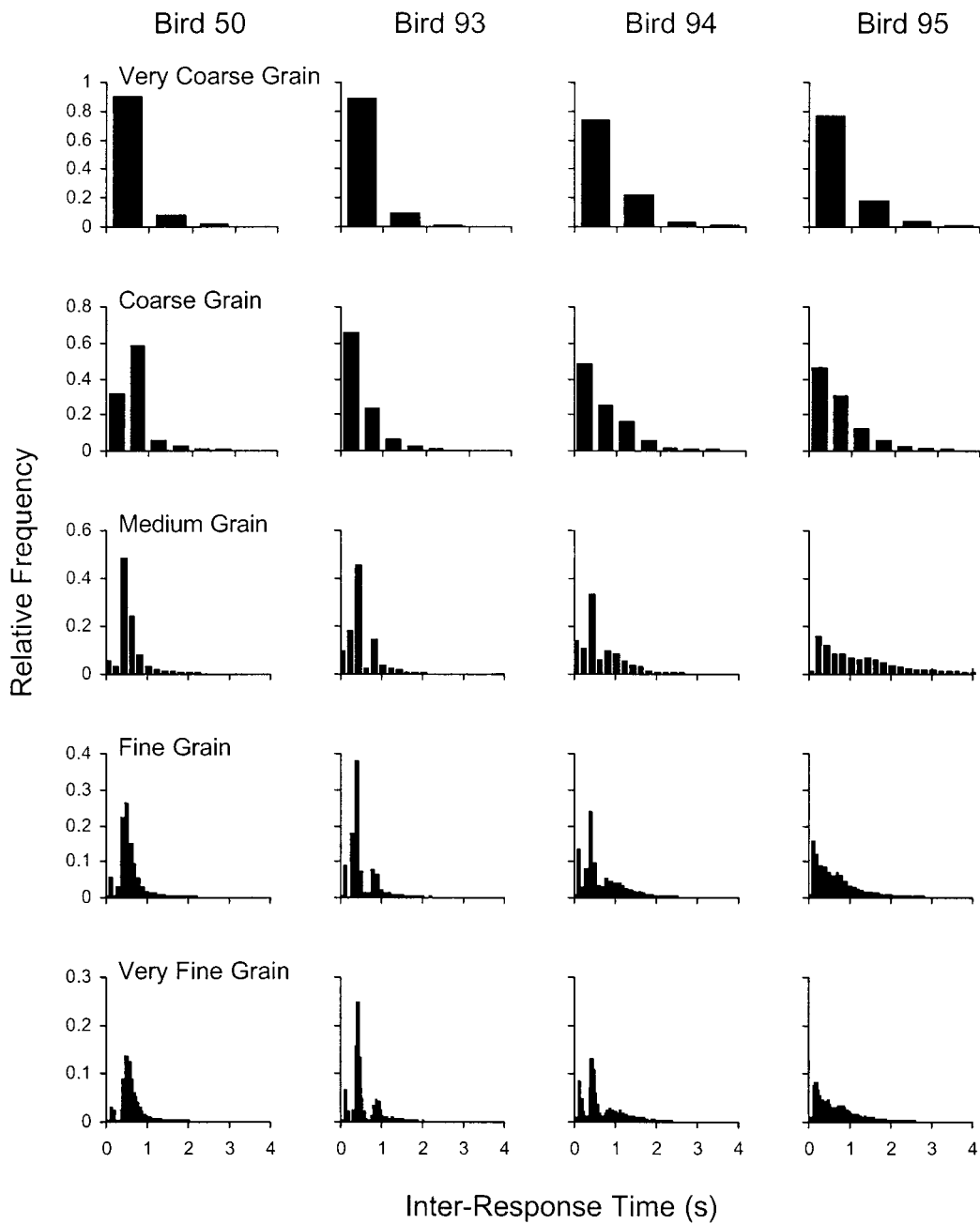


Fig. 2. Relative frequency distributions of interresponse times (IRTs) during the VI 960-s baseline shown in the first column of Figure 1. In the top row data are aggregated in 1-s bins (very coarse grain); in the second row, in 500-ms bins (coarse grain); in the third, 200-ms bins (medium grain); in the fourth, 100-ms bins (fine grain); and in the bottom row, 50-ms bins (very fine grain).

which temporal bin is under consideration are equal to base probabilities and may be ignored in accounts of behavior. Response rate (or its reciprocal, the average IRT, given

by the mean of the distribution) constitutes a complete description of those distributions, in that it completely specifies the single parameter in the geometric distribution.

Exponential distributions. As the bin size approaches 0, the geometric density approaches the exponential, $\lambda e^{-\lambda t}$. The parameter λ is the instantaneous rate of emission of responses. It has the dimension of the inverse of the unit in which t is measured, typically s^{-1} . An exponential distribution may be superimposed on the geometric by evaluating the exponential at the midpoint of each bin and multiplying that density times the width of the bin (Δ). The visual approximation of the geometric to the exponential improves as Δ decreases, down to the point at which they are equivalent. To get the geometric probability p from the exponential, integrate the latter over the bin duration: $p = 1 - e^{-\lambda\Delta}$. The inverse relation between the rate of the exponential and the probability of the geometric is $\lambda = -\ln(1 - p)/\Delta$. An animal averaging one response per second has a probability of .63 (i.e., $1 - e^{-1}$) of making a response during the first second after a previous response. The exponential provides a more general description than the geometric because it is not limited to predictions about a fixed category width, but, once λ is determined, can predict the probability of a response in a bin of any width.

Researchers content with the description at a level consistent with the geometric distribution already have in hand a map between rate and probability. The instantaneous rate of responses, λ , is the key parameter of the exponential; the average IRT is $1/\lambda$. The map between probability and that rate is $p = 1 - e^{-\lambda\Delta}$; its inverse is $\lambda = -\ln(1 - p)/\Delta$. Estes (1950) proposed this relation between rate and probability (cf. Bower, 1994), but it was not further developed.

The best (i.e., unbiased, maximum likelihood) estimator of the instantaneous rate, λ , is the reciprocal of the average IRT (Evans, Hastings, & Peacock, 1993); that is, the average response rate. Therefore, the best way to calculate response rates is in the traditional manner: Sum IRTs and divide by n . To the extent that the shape of the IRT distributions shown in the top row of Figure 2 is maintained at finer levels of analysis, the exponential distribution gives the map between probability and rate, and the average IRT provides the best measure of it.

As the categories become more refined and more details become visible, however, the

simple geometric description no longer suffices. At the next level in Figure 2—still a relatively coarse grain—it can be seen that short IRTs are not the most common for Pigeon 50, as required for a geometric process.

General Models for the Distribution of IRTs

Mathematical models are most satisfying when they act as the cords of an analogy, tying an abstruse process such as the strengthening or patterning of an operant to visualizable models in an explicit and testable manner (Miller, 1984). Machines that generate data consistent with the mathematical models of behavior provide a more accessible, intuitive grasp of possible mechanisms that underlie the response patterns. As nonbiological devices that demonstrate behavior similar to the biological systems we seek to understand, the machines constitute sufficient descriptions. They are not necessary descriptions because other machines can be fabricated that generate comparable output. The same is true for any scientific theory. A model that provides a sufficient description is a step forward. That step may encourage other steps. Once competing descriptions are available, considerations such as parsimony and elegance help to select among them.

The cyclic machine. Consider first a reflex circuit that generates perfectly periodic responses, ones that might issue from pacemakers such as a pendulum or a rotating wheel with a pawl that activates a switch with each revolution. Such machines generate the simplest IRT distributions, a single category at the mean IRT having a probability of 1 and a width approaching 0. The model has one parameter, δ , the period of cycle. Response rate is $1/\delta$ with zero variance.

This machine may be simulated on a computer using the flowchart shown at the top of Figure 3. The machine starts by setting an elapsed time variable t to 0. It then loops until the elapsed time equals δ , whereupon it emits a response, resets the timer to 0, and then waits to the end of the next epoch to make another response.

The stochastic cyclic machine: Very coarse grain. The next level of detail, exemplified by the top row of Figure 2, shows a skewed distribution of IRTs. A pacemaker that misses its strike with probability $1 - p$ generates such

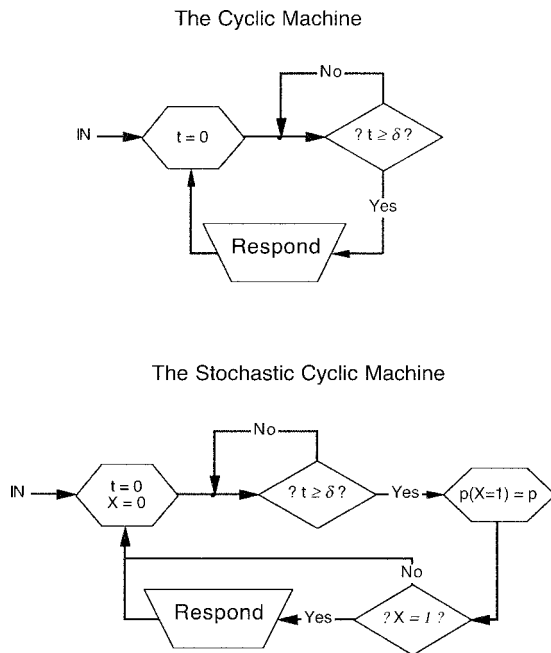


Fig. 3. The flowcharts for two machines that generate exemplary IRT distributions. Top: a machine that emits a response every δ seconds. Bottom: a machine that emits a response every δ seconds with probability p . It generates geometric distributions indistinguishable from those in the top row of Figure 2.

geometric distributions. If the pacemaker's period is fast relative to the bin size, the continuous analogue of the geometric distribution—the exponential with rate λ —provides a general description. If nothing other than the mean response rate (λ) is known, the most random distribution, and therefore the one most parsimonious of prior assumptions, is such an exponential (Kapur, 1989).

This machine may be simulated on a computer using the flowchart at the bottom of Figure 3. The machine starts by setting an indicator variable X and elapsed time t to 0. It then loops until the elapsed time equals δ , whereupon it sets the indicator to 1 with probability p . On those epochs when X is set to 1, the machine emits a response, reinitializes the variables, and then waits for the end of the next epoch. If δ is relatively large, the result is a geometric distribution, such as those shown in the top row of Figure 2. Letting the period δ become very small while the ratio p/δ remains constant at λ morphs the geometric distribution into an exponential with rate parameter λ .

The stochastic cyclic machine: Coarse grain. A finer grained analysis of IRT frequencies (the second row in Figure 2) shows distributions that may increase before decreasing. This is consistent with real cyclic machines having periods $\delta > 0$, as no pecks are possible before the pacemaker has recycled, causing dead, or *refractory*, times of δ seconds after each response, and yielding a maximum rate of $1/\delta$. All organisms have such ceilings on their response rates. As the bin size Δ decreases, the probability of observing responses in the first bin will decrease with it, until $\Delta < \delta$, where it must fall to 0. As bin size goes to 0, the geometric morphs into an exponential distribution with an origin at δ rather than at 0. This *refractory Poisson* model describes many response distributions. Its implications are consistent with other models of response rates under ceiling constraints (Killeen, 1982; Killeen & Bizo, 1998). It predicts a hyperbolic relation between responses that are evoked by reinforcement at a rate of λ and the rate, b , of those that are able to be emitted and measured: $b = \lambda/(1 + \lambda\delta)$ (Bharucha-Reid, 1960). This may be understood by writing the average IRT, $1/b$, as the time between responses plus the duration of a response: $1/b = 1/\lambda + \delta$. Taking reciprocals yields the above equation.

A special case holds when λ is proportional to the rate of reinforcement, R , relative to the period of the response generator, δ . Then $\lambda = aR/\delta$, and it follows that $b = kR/(R + 1/a)$, with $k = 1/\delta$. This is Herrnstein's hyperbola (de Villiers & Herrnstein, 1976; Herrnstein, 1974). The parameter a measures the motivation of the organism (Killeen, 1994). Thus molar measures and theories are closely related to these descriptive models of IRT distributions.

The refractory Poisson process provides an almost perfect fit to these and other data (e.g., Reynolds & Catania, 1961) when the bin size is no finer than $\Delta \approx 1/2$ s. Depending on how long the refractory period is relative to the bin size, one or more of the leading bins may be empty. The first bin to contain a response will have its leading edge clipped by the end of the refractory period, so it will generally show lower relative frequencies of responding than the second occupied bin. This effect is demonstrated below (in Figures 7 and 10).

As the bin size decreases relative to the period of the cycle, the stochastic cyclic machine may generate histograms with periodicities. The distributions will have higher-than-exponential probabilities at multiples of the period and lower-than-exponential probabilities between periods. These are visible in the medium- and fine-grained IRT distributions (middle and lower rows of Figure 2), and require the next class of models.

The stochastic cyclic machine with noise in output. Although the shifted origin of the above machines allows refractory periods between responses, it cannot account for relative frequencies of responding that increase for more than one bin before decreasing, as is the case for all but Pigeon 95 in the middle row of Figure 2. Nor can those machines account for more than two bins that fall between periodicities and have frequencies greater than zero. The stochastic cyclic machine is too precise for these things to happen; timing must be imperfect, that is, there must be noise in the model. Noise will permit varying degrees of smoothing between the discrete bins of the geometric process, causing it to approximate, to varying degrees, the data. There are two obvious ways to introduce noise.

The first way to introduce noise, developed by McGill (1962, 1963) and Wing and Kristofferson (1973), maintains a perfect pacemaker of period δ , but adds random delays between the start of action and the registration of a response. It is as though the pecking reflex is perfectly periodic, but a random variable intrudes between the initiation of a peck and its execution. This could be as simple as variance in the distance of the head from the key from one peck to the next. The predicted distributions of IRTs have a quick exponential ramp up to a maximum at the period of the pacemaker and a slower exponential decrease thereafter. This model provides an excellent fit at the first two levels of analysis. It makes two predictions that are not supported, however: With a constant-speed pacemaker, when one response takes longer than average, the next response should follow closer on its heels than average. Conversely, a short IRT should be followed by a longer IRT. This may be easily tested by correlating all IRTs with those that immediately follow (an autocorrelation Lag 1). The McGill mod-

el predicts those autocorrelations to be negative. For the 4 pigeons in the VI 960-s condition of the VI experiment, the autocorrelations were $r = -.01, .20, .15,$ and $.42$; all but the first are significantly greater than 0. The McGill model is therefore not isomorphic with these data. Neither is the simple exponential, for which none of the autocorrelations should be significantly different from zero. The McGill model generalizes the simple exponential in the wrong direction. A model developed by Shull, Gaynor, and Grimes (2001) generalizes it in the correct direction, as we shall see next.

Response bouts. Another way of testing the McGill (1962, 1963) and exponential models is in terms of conditional probabilities. If every IRT less than the median is called *short* and the others are *long*, what is the probability that a short response will occur given that one had just occurred $p(\text{short}|\text{short})$? For a random process this should be $.50$. For the 4 pigeons it was $.58, .56, .61,$ and $.73$. These are significantly greater than $.50$. There is thus a tendency for the pigeons to continue in a high response-rate mode given that they are in that mode. The pigeons engage in bouts of responding, and being in a bout (i.e., emitting a short IRT) is a better-than-chance predictor that it will stay in that mode and emit another short IRT. This conceptualization of response rates as a mixture of distributions was developed and validated by Shull et al. (2001), and will be represented as a probability machine after a crucial subcomponent is developed.

The clocked Bernoulli module (CBM). In Figure 4 the stochastic cyclic machine is generalized so that it may be used as a module in constructing other machines. It is a slightly revised version of the stochastic cyclic machine shown in Figure 3: It utilizes generic parameters, τ in place of δ and π in place of p , that are not associated with any particular behavioral process. Rather than respond and recycle, it sets a flag (the state indicator goes from 0 to 1) and exits. Because of the general utility of such a module, it is given a name: *clocked Bernoulli module* (CBM). A Bernoulli process is one that metaphorically tosses a coin with a probability π of heads, and exits when it obtains a head. It is the basis of many distributions, such as the binomial and its offspring, the normal, Poisson, and exponential.

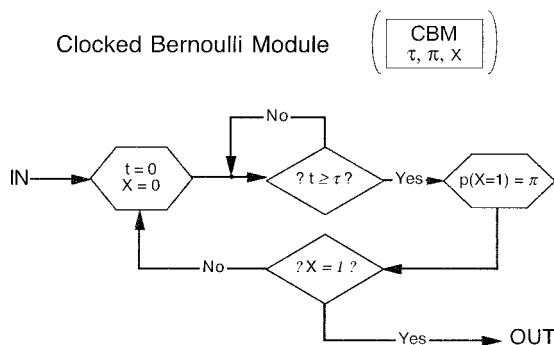
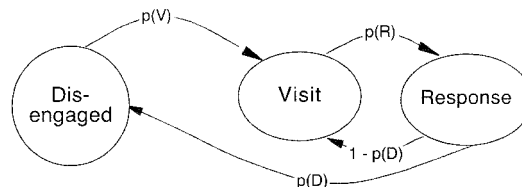


Fig. 4. A mechanism that is used in many of the probability machines, the clocked Bernoulli module (CBM) is a generalization of the machine shown in the bottom of Figure 3. The CBM essentially flips a coin every τ seconds. The coin has a probability π of landing heads; if it does so, it sets the state variable X to 1 and exits. The CBM is represented by the icon at the top of the figure.

The CBM is clocked because it adds a temporal element τ that must elapse between tosses. Such a temporal element is necessary to use Markov processes as models of temporal phenomena. The CBM is essentially a probabilistic pacemaker of period τ . τ defines the interval over which the probability is measured. If τ is small it need not be assigned a particular value, because it is the ratio of π/τ that is then measured as a mean rate of output (λ), and that mean is invariant over proportional changes in π and τ .

Figure 5 shows the Markov model used by Shull et al. (2001) to describe the responding of rats under a variety of schedules of reinforcement and motivational conditions. It is similar to one used by Heyman (1988) for responding maintained by VI schedules. Shull et al. found that motivational operations primarily affected the probability of initiating a response bout, $p(V)$, whereas manipulating the contingencies of reinforcement affected the probability of ending a bout, $p(D)$. These two operations correspond to arousal and coupling manipulations in the mathematical principles of reinforcement (Killeen & Bizo, 1996, 1998). The probability machine associated with the Shull et al. model is shown in the bottom of Figure 5, with reinforcement added to their basic model. The IRT machine is a simple CBM with probability parameter $\pi = p(R) = b$. The temporal parameter is undefined and therefore takes the unit in which time is measured (for Shull

Shull Gaynor & Grimes (2001)



Probability Machine

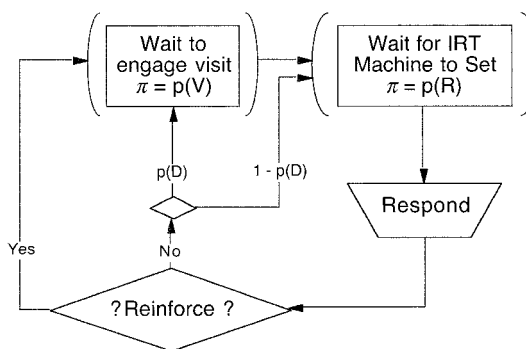


Fig. 5. Top: a Markov model for responding consisting of bouts of target responses interspersed with non-responses. Bottom: the flowchart for the corresponding probability machine. The component modules are described in Figure 4.

et al., this was 1 s). The hiatus from bouts (the “disengaged” state) is also a CBM, with parameters $\pi = p(V)$ and $\tau = 1$ s.

Finer Analyses

For both the McGill model and the geometric-exponential model, frequencies should be monotonically decreasing for all IRTs longer than the period of the pacemaker. The deeper levels of analysis in Figure 2 show that this is not the case: There is residual periodicity in the IRTs, as shown by Blough (1963) and others (e.g., Gentry, Weiss, & Laties, 1983; Palya, 1991, 1992; Ray & McGill, 1964). This periodicity is more clearly visible in Figure 6.

The stochastic cyclic machine with noise in the period. The periodicity in the finer grained distributions is captured by another model that adds noise, not to the motor latency but

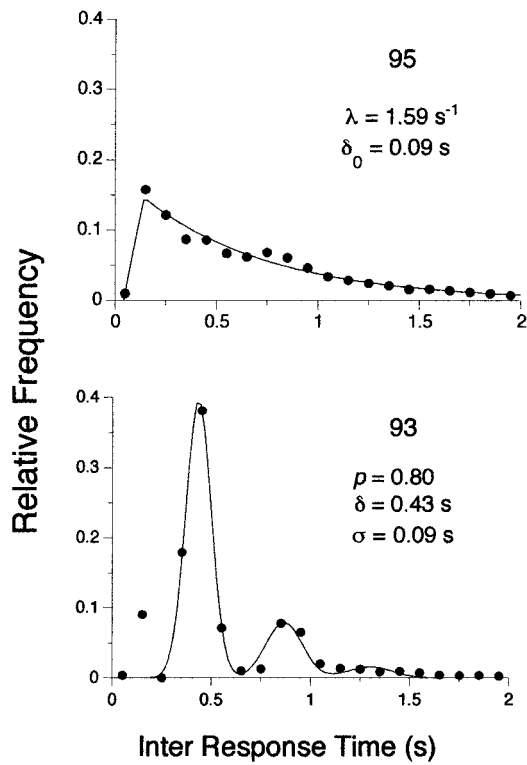
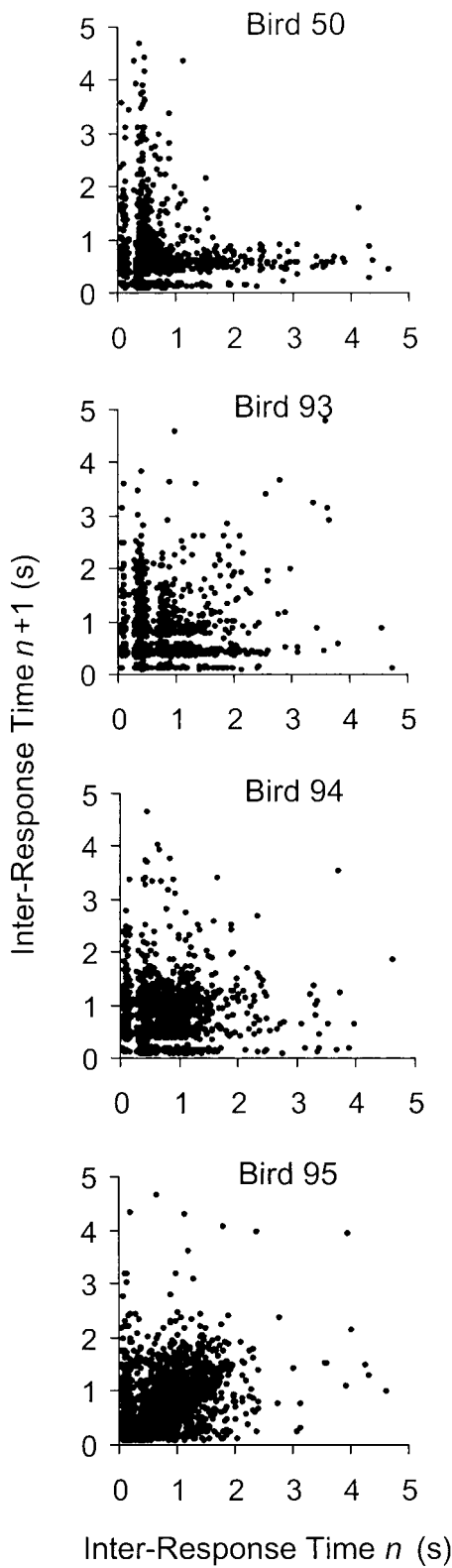


Fig. 7. Top: the shifted exponential fit to the fine-grained distribution of Pigeon 95. Some residual periodicity is visible. Bottom: the Palya distribution fit to the fine-grained distribution of Pigeon 93, with parameters as shown and $\delta_0 = 0$ s. The function is a mixture of Gaussian distributions with first mode at $\delta + \delta_0$, second at $2\delta + \delta_0$, and third at $3\delta + \delta_0$; with variances of σ^2 , $2\sigma^2$, and $3\sigma^2$; and with areas of p , $p(1 - p)$, and $p(1 - p)^2$. The return plots and IRT machines corresponding to these graphs are shown in Figures 6 and 8.

to the period of the pacemaker. It is as if the wheel of the pacemaker is not perfect, but wobbles on its axis, generating a “smear” around the periodic bins. This is most obvious for the data from Pigeon 93 from Killeen and Hall (2001) redrawn in the bottom of Figure 7. Call the standard deviation of the fundamental period σ . When a peck is missed, it is as though it occurred “off key” (Bachrach, 1966). If a peck occurs at the next

←

Fig. 6. The number of IRTs of a duration given by the y axis, following an IRT of a duration given by the x axis. Data for these return plots are from the last five sessions of the VI 960-s baseline condition of Killeen and Hall (2001).

opportunity, its average IRT will be 2τ , and its standard deviation will be $\sqrt{2}\sigma$, and those around subsequent increments, located at 3τ , 4τ , $n\tau$, will increase as the square root of n . This means that the banding will be less distinct around longer IRTs, as seen in Figure 6. This machine is identical to the model of Palya (1992). Because it is an important and accurate model, we call such a probabilistic cyclic machine with variability in its period a *Palya machine*, and describe it more fully below.

The simplicity of the two-parameter refractory Poisson process makes it the preferred model whenever the periodic ripples in the distributions are minimal (e.g., in the top panels of Figures 7 and 10). Just as a black-and-white sketch is sometimes preferred over a full-color portrait, a scientist may choose a more economical description, such as two-parameter distributions, knowing that all descriptions are, in any case, approximations.

Two-state Palya machines. At the fine and very fine levels of analysis (Figure 2), another feature is manifest in the data that is absent from the above models. The first mode in the data for all pigeons except 95 is smaller than the second mode, a feature observed by other investigators (e.g., Mazur & Hyslop, 1982; Rau, 1997). Various accounts may be given of this feature: Perhaps it is caused by a peculiar pecking topography, such as an opening of the beak on the completion of one response to cause a quick second response. Ploog and Zeigler (1996) found that pigeons made two to four times as many beak-opening (gaping) movements as were recorded as key pecks (their Figure 1). If some pigeons make these closer to the key than others, this would account for the presence of early minor modes in their distributions. We have not seen this feature in the IRT distributions of rats (see below).

Multistate Palya machines such as that invoked for the data of Shull et al. (2001) could accommodate minor early modes and provide a more accurate description of the character of response rates, including the autocorrelational data. Complete descriptions are, of course, more prolix than cursory ones.

The Structural Similarity of the Models

The parameters of these increasingly fine descriptions are closely related, and are uni-

fied by the generic IRT machine diagrammed in Figure 8.

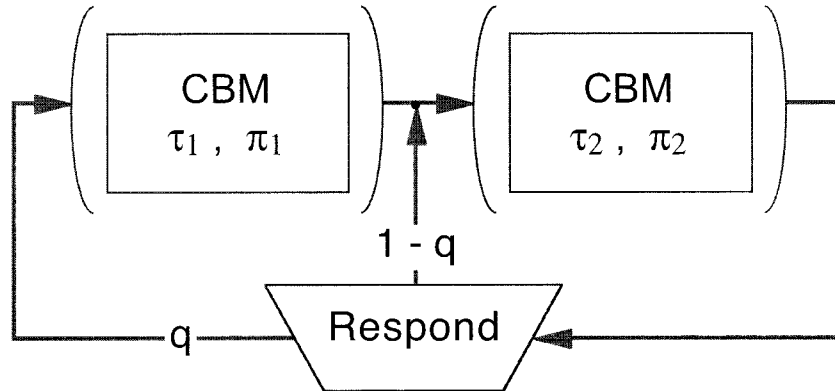
The IRT for the cyclic machine is δ , and thus the average running response rate is $1/\delta$ with variance 0. It may be characterized in Figure 6 by a point at $x = \delta$ s and $y = \delta$ s. In Figure 8 the first CBM is taken out of play by setting the probability of quitting, q , to 0. The second CBM generates a deterministic period of δ seconds. This machine has one free parameter, δ .

The cyclic machine becomes a stochastic (probabilistic) cyclic machine if $p < 1$. The first CBM in Figure 8 is deleted by setting q to 0, and the second generates a geometric distribution of IRTs. The mean number of cycles required for a response is $1/p$, each of duration δ . The mean IRT is δ/p . The average response rate is the reciprocal of the average IRT: p/δ . As p approaches 1.0, these statistics approach those of the simple cyclic machine. This is a strongly periodic machine. It may be characterized in Figure 6 by a lattice, most dense at $x = \delta$ s and $y = \delta$ s and thinning at multiples of these values according to $p(1-p)^{n-1}$. This machine has two free parameters, δ and p .

As their period approaches 0, stochastic cyclic machines function as a Poisson machine, which generates a simpler exponential distribution of IRTs. The Poisson machine has a very short period τ and a probability of $\pi = \lambda\tau$, where λ is the rate of initiation of responses during the epoch τ . The ratio of π/τ determines the single free parameter of this system, its rate λ .

The refractory Poisson machine generates an exponential distribution shifted by δ from the origin. The shift corresponds to the time required to make a response, a time during which other responses are impossible—a refractory period. The average IRT is the mean of the exponential distribution from the second CBM ($1/\lambda$) plus the refractory period: $(1/\lambda + \delta)$. The average response rate is the reciprocal of this average IRT, $\lambda/(1 + \lambda\delta)$, which approaches the rate of initiation of responses λ as the dead time δ goes to 0. This model suffices to describe the IRTs of rats' lever presses; it suffices for pigeons except at fine-grained levels of analyses, for which a Palya machine is more accurate. It may be characterized in Figure 6 by a gap to $x = \delta$ s and $y = \delta$ s, and exponentially thinning

IRT Machines



Machine	Parameter		Constraints		
	τ_1	π_1	τ_2	π_2	q
Cyclic	-	-	δ	1	0
Geometric	-	-	δ	p	0
Exponential	-	-	≈ 0	$\lambda\tau$	0
Shifted Exponential	δ	1	≈ 0	$\lambda\tau$	1
Palya	δ_0	1	$\delta \pm \sigma$	p	1
Shull	≈ 0	$\lambda_1\tau$	≈ 0	$\lambda_2\tau$	q

Fig. 8. IRT machines based on two sequenced CBMs. With appropriate parameters, these clocked Bernoulli processes emulate all of the models of IRTs developed in this paper. The rate of leaving each CBM is $\lambda = \pi/\tau$. All are exponential in the limit as $\tau \rightarrow 0$, in which case the single parameter λ characterizes the module. For the first three machines, only the second CBM is operative. Cyclic machines dwell in the second state for δ seconds and then respond with a probability of 1. Geometric machines dwell in the second state for δ seconds before leaving with probability p . Exponential distributions are generated when $\tau \rightarrow 0$ with the rate of exit constant at $\lambda = \pi/\tau$. The shifted exponential machine adds a fixed dead time δ , corresponding to the minimal response time. The Palya machine is like the shifted exponential, with an offset of δ_0 and a variable period for the second CBM. When animals engage in bouts of responding and hiatus, τ_1 gives the time away from the response state before there is an opportunity to revisit it. If $\pi_1 = 1$, the hiatus is of fixed duration. If $\pi_1 < 1$, it is a geometric-exponential hiatus. After each response the animal begins a hiatus with probability q , or stays in the bout with a probability of $1 - q$. While in the bout state, responses are exponentially distributed in the case $\tau_2 \rightarrow 0$.

marks above these values, as is approximately the case for Pigeon 95. It permits a direct bridge to contemporary theories of molar response rate. It is implemented in Figure 8 by bringing the second CBM into play after a delay of δ s in the first CBM, and is otherwise the same as the exponential arrangement. It has two free parameters, δ and λ .

The Palya machine adds variance to the period of the stochastic cyclic machine. It was defined and shown to provide a good description of pigeon IRTs by Palya (1992). Because the variance is symmetric around multiples of the period, the mean IRT is simply that of the geometric distribution: δ/p . The difference between the shifted exponential and the Palya machines is that the former predicts a constant marginal probability of a response any time after the refractory period δ , whereas the latter predicts periodicity in responding. As the variance of the pacemaker increases, so does the resemblance of the Palya model to the exponential: Variance smoothes the natural periodicity of the cyclic machine. This is manifest in Figure 7 for Pigeon 95, whose data are more parsimoniously represented by the shifted exponential model than by a Palya machine. The Palya machine may be characterized in Figure 6 by a blurred lattice, dense at $x = \delta$ s and $y = \delta$ s and thinning and blending at multiples of these values. It is charted in Figure 8. The mean period of the second CBM, δ , probably corresponds to cyclic pecks and off-key pecks. This period has a standard deviation of σ . In the first CBM, δ_0 corresponds to the extra time necessary to make a measured response, and appears as a small adjustment of the location of the origin of IRT distributions. For simple Palya machines effective and ineffective responses have the same period, so that $\delta_0 = 0$. If δ is much smaller than δ_0 , then the Palya machine acts like a refractory Poisson and generates a shifted exponential. There are four free parameters in the Palya machine and three in the simple Palya machine.

The two-state Markov model proposed by Shull et al. (2001) is charted in Figure 8 with parameters that hold the animal in the first state for a hiatus of mean duration $1/\lambda_1 = \tau_1/\pi_1$, and then move it to the response state at which it emits target responses at a rate of λ_2 . After a response it quits the response state with probability q . This entails that response

bouts have mean durations of $1/(q\lambda_2)$. In the case $q = 1$, the Shull machine reduces to a Palya or Poisson machine with an exponential refractory period, and generates IRT distributions that are a mixture of two exponential distributions. There are three free parameters in the Shull machine.

Other models are readily constructed and identified by their parameters: $\text{IRT}(\tau_1, \pi_1, \tau_2, \pi_2, q)$. $\text{IRT}(\tau_1, 1, \tau_2, 1, 1)$ strictly alternates two IRTs. $\text{IRT}(\tau_1, 1, \tau_2, 1, q)$ generates a mixture of two IRTs. For exponential CBMs, the description reduces to $\text{IRT}_e(\lambda_1, \lambda_2, q)$. $\text{IRT}_e(\lambda_1, \lambda_2, 1)$ gives a generalized gamma distribution (McGill & Gibbon, 1965) used to describe the time course of adjunctive responding (Killeen, 1975).

Some models of IRTs, such as the variable-rate model of McGill (1963) and Wing and Kristofferson (1973) mentioned above, require that Figure 8 be rewired. In these models the emission of a response does not affect the first CBM: The first CBM is free running, so its output returns to restart it, as well as start the second CBM. The first CBM is deterministic, so $\pi_1 = 1$. The second CBM is assumed to represent motor variance. The model generates a sharply peaked LaPlace distribution of IRTs. The tau-gamma model of Church, Broadbent, and Gibbon (1992) adds a third CBM to the McGill model; this rounds off the peak of the LaPlace and fits a variety of data.

A major characteristic of these machines is the approximately geometric or exponential decrease in the probability of emitting IRTs as a function of their length. The marginal probabilities (the probability of making a response during the epoch t to $t + \Delta$, given that the animal has gotten to t) are constant over time (a) for the exponential; (b) for the shifted exponential after the refractory period; (c) for the geometric model at multiples of the period; (d) for the Shull model within hiatus and within bouts; and (e) for the Palya machine at multiples of the period but with smear into the intervals between multiples. The information embodied in these tails is efficiently represented by the average response rate. When the probability of responding within an observational epoch is small (say, $p < .25$), all models make the same prediction: Response rate is approximately proportional to response probability.

Hazard functions (Luce, 1986) are graphs of these marginal probabilities as a function of t , and thus characterize the distribution. Anger's (1956) IRTs per opportunity measure is a hazard function. The hazard function of the exponential is a horizontal line at a height of λ , which provides an approximation of many data at the very coarse and coarse levels of analysis. The hazard function of the shifted exponential is a step up at δ s from a line at 0 to one at λ . That of a Shull machine is a step down from a line at λ_1 to one at λ_2 . An alternative representation, the log survival function (Shull, 1991), is used to advantage by Shull *et al.* (2001) and Conover, Fulton, and Shizgal (2001).

Application of Models to Other Data

Pigeons on VR schedules. The data from Killeen and Hall (2001, Experiment 2, Phase 1, first 200 trials) may be analyzed in a manner similar to that shown in Figure 7. Figure 9 shows the fine-grain (100-ms bins) IRT distributions for 4 pigeons from Killeen and Hall. The distributions are similar to those obtained for VI schedules (cf. Figure 2). The hint of periodic ripples in the 4th, 8th, and possibly 12th bins would befit a mechanism such as the Palya machine.

NEW EXPERIMENTS

Additional experiments were conducted to establish the generality of the results just summarized. Rats were used to assess species generality, and both lever pressing and instrumental licking were used to assess the impact of response topography.

Experiment 1: Fixed-Ratio (FR) Schedules

The procedures were modeled after those of Bizo and Killeen (1997).

Subjects. Six male Sprague-Dawley rats (*Rattus norvegicus*) with previous experimental histories were maintained at $85\% \pm 10$ g of their ad libitum weights and were fed supplemental rat chow, if needed, at 6:00 p.m. Rats were housed individually with free access to water in a colony illuminated on a 12:12 hr light/dark cycle with dusk at 6:30 p.m.

Apparatus. The experiment was conducted in a Lehigh Valley Electronics chamber, measuring 18 cm by 21 cm by 31 cm, enclosed in

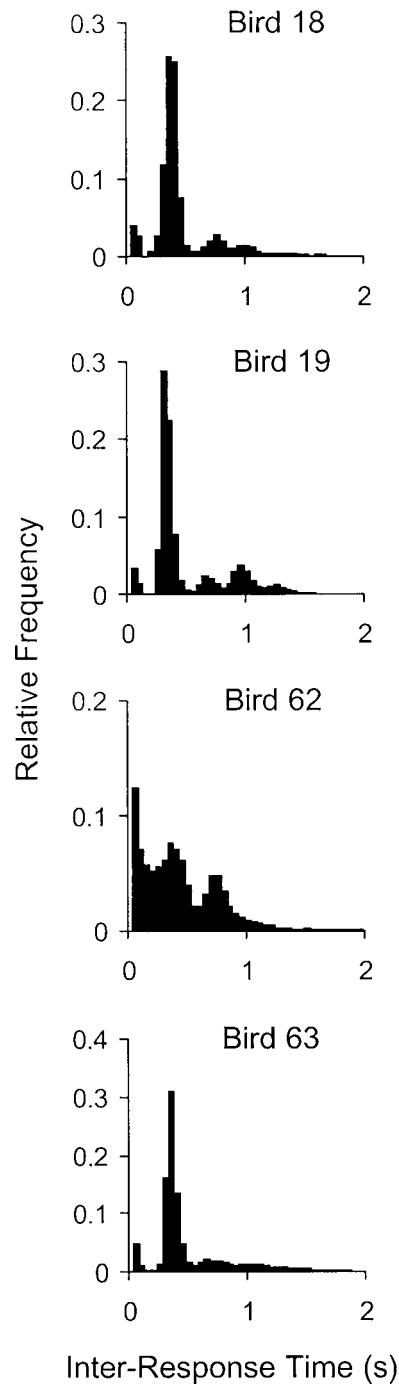


Fig. 9. The IRT distributions of 4 pigeons reinforced on random-ratio schedules. Echoes of the fundamental peck rate identify the generator as a Palya machine, with fundamental period δ between 0.3 and 0.4 s. Pigeon 19 showed minimal variability around this cycle, whereas Pigeon 62 showed substantial variability.

a sound-attenuating box. Access to 45-mg food pellets was available in a magazine measuring 2 cm by 4 cm by 5 cm, located 3 cm from the floor. A Coulbourn Instruments retractable lever was located 3 cm to the left of the magazine and 3 cm from the floor. A force of 0.22 N was required to activate a microswitch, registering an effective lever press. A houselight, 12 cm from the floor and 9 cm from the left wall, illuminated the chamber. An IBM-compatible 286 computer controlled the experiment. A ventilation fan yielded an ambient noise level of approximately 71 dB.

Procedure. Sessions began with a 15-min habituation period in which the houselight was illuminated and the lever was retracted. After this period, the lever was extended into the box and lever pressing was reinforced according to an FR schedule with one 45-mg Noyes pellet. Each reinforcer was followed by a 20-s intertrial interval in which the lever was retracted. Sessions were terminated following 100 pellet deliveries or 2 hr. Rats experienced each of the following sequence of ratio values for two sessions: 1, 2, 3, 4, 6, 8, 10, 12, 15, 20, 25, 30, 35, 40, 50, 60, 70, 80, 90, and 100.

Results. Response distributions were collected from the FR 80 condition. On this schedule 79 sequential IRTs were available uninterrupted by reinforcement, and response rates were high enough to provide many such sequences over the course of a single session. The relative frequency distributions were based on a fine-grain (100-ms) bin size. Figure 10 shows that distributions from 5 of the rats were similar. The curve through the data is a shifted exponential with median values of $\delta = 0.05$ s and $\lambda = 5.6$ s⁻¹. Although exposure to the various schedules was intentionally brief, with only 36 sessions on various FR schedules prior to data collection, this did not compromise the ability to describe the resulting IRT distributions with a simple model.

Notice that data are plotted at values less than the refractory period δ . Consider R1 in the top left panel, with $\delta = 0.059$ and a relative frequency of 0.16 at the abscissa 0.05. This abscissa is the midpoint of a bin ranging from 0 to 0.10; it therefore captures the first 41% of the IRT distribution. The rising limb of the theoretical distribution is the relative frequency of responses expected in a bin with a lower limit of 0 and a midpoint given by its

abscissa, as the upper limit is continually increased.

The shifted exponential provides a good model at this and coarser levels of analyses for all rats except R6, shown in the bottom right panel. For R6 a more accurate and aesthetic model is provided by a simple Palya machine with parameters $p = .92$, $\tau = 0.14$ s, and $\sigma = 0.06$ s, and no shift in the origin ($q = 0$).

Experiment 2: Instrumental Licking

Subjects. Two naive male Sprague-Dawley rats (*Rattus norvegicus*) were given varying amounts of access to water at the end of each day (ranging from 6 to 30 min). They were housed individually with free access to food in their home cages, and the colony was illuminated on a reversed 12:12 hr light/dark cycle with dusk at 6:30 a.m.

Apparatus. The experiment was conducted in a Coulbourn Instruments chamber measuring 28 cm by 30 cm by 22 cm. A spout connected to a lickometer was located 11 cm from the left wall, 5 cm off the floor, and extended 1 cm from the wall. Two lights, each measuring approximately 1 cm in diameter, were centered 1 cm above the spout. A houselight was centered in the box 1 cm from the ceiling. An IBM-compatible 386 computer controlled the experiment. The chamber was enclosed in a sound-attenuating box suffused with white noise and fan noise yielding an ambient noise level of approximately 76 dB.

Procedure. Rats were trained to lick the spout under an FR 1 schedule of water reinforcement. After two training sessions, the experimental sessions began with a 100-s habituation period, after which a green light above the spout signaled availability of water contingent on 50 licks of the dry spout. The rats had 20 min of access to water the evening prior to experimentation. Reinforcement consisted of the delivery of 0.04 ml of water through the spout, signaled by changing the light above the spout to red. Sessions were terminated after the delivery of 100 reinforcers. Twenty-six sessions were conducted, with data taken from the last session. This selection was arbitrary, in that data sampled from the prior 15 sessions were not noticeably different in character.

Results. Frequency distributions were constructed for rats' licking in 20-ms bins. These

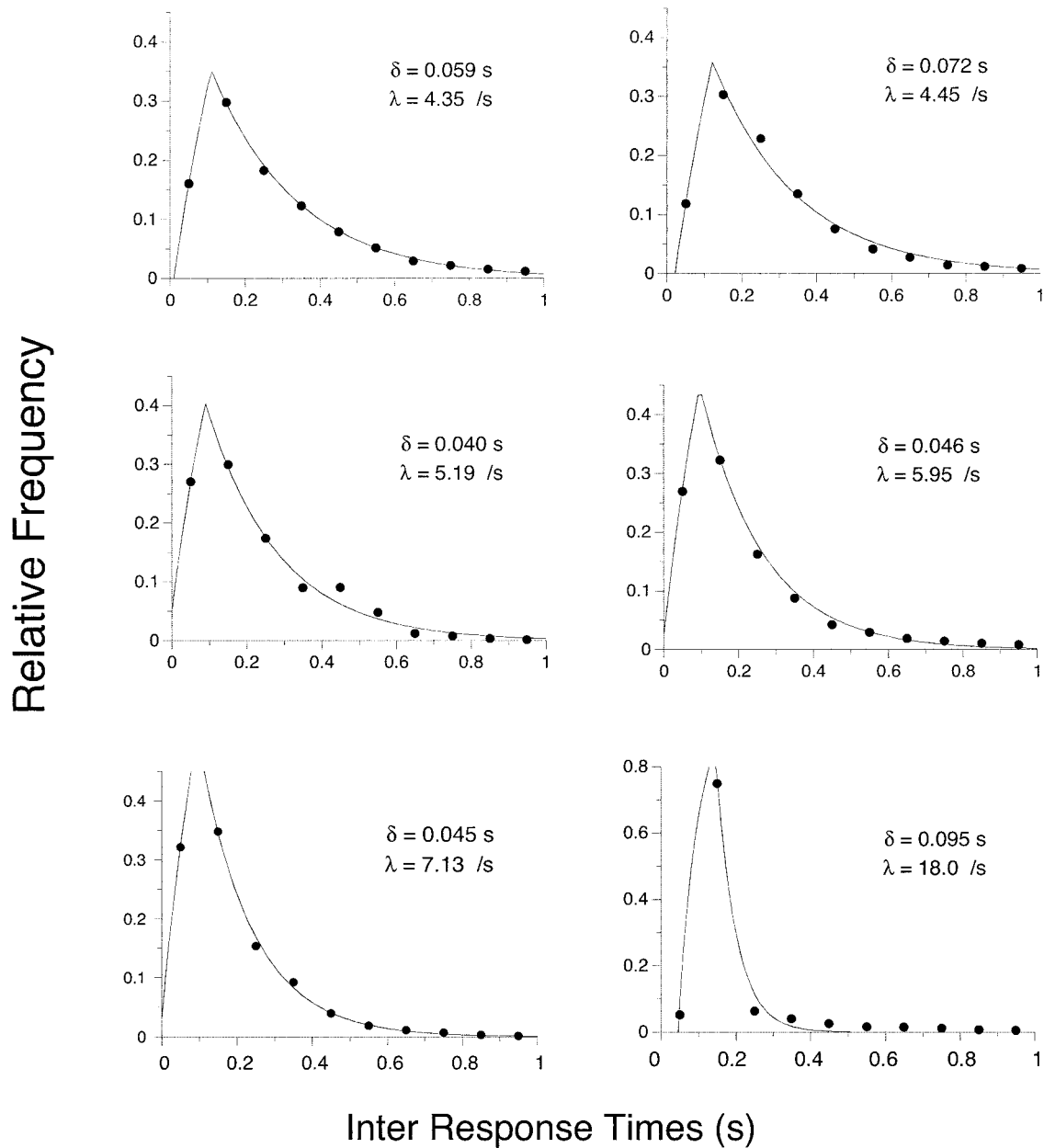


Fig. 10. The IRT distributions of 6 rats reinforced with 45-mg food pellets for lever pressing on FR 80 schedules in Experiment 1. The curves are shifted exponential distributions (the refractory Poisson process) with parameters δ (refractory period) and λ (intrinsic rate of initiating responses).

are displayed in Figure 11. The distributions were similar for the 2 rats and were combined in a relative frequency display. The curve through the data is the signature of a simple Palya machine ($q = 0$) with parameters $p = .90$, $\delta = 154$ ms, and $\sigma = 34$ ms.

This concludes the analysis of response rate

in terms of IRT machines. Although different operant responses may interact with reinforcement contingencies in different ways (e.g., Hemmes, 1975), the basic patterning of pecks, presses, and licks all exemplify the same machinery (Figure 8), with different parameter settings characterizing animal or op-

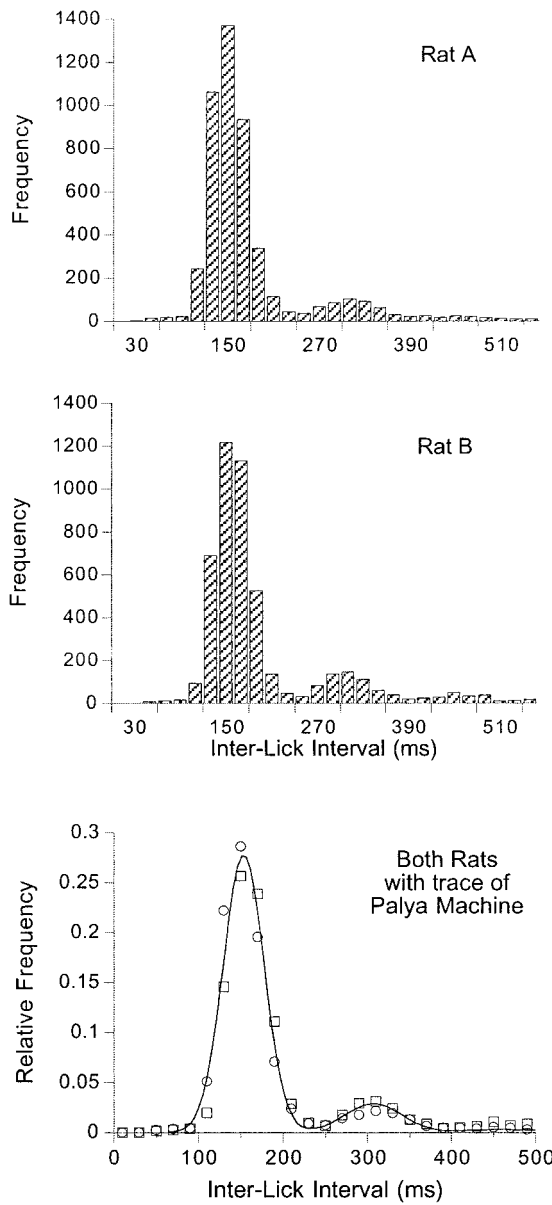


Fig. 11. The IRT distributions of 2 rats from Experiment 2 licking a spout on FR 50 schedules for water. The top panels show the frequency distributions of interlick intervals from one session comprising 5,000 responses. The bottom panel represents those data as relative frequency distributions. The curve is the output of a simple Palya machine.

erandum. In these machines, probability is defined over the epoch τ , and iterative sampling of this epoch generates a response rate. In a later section of this paper, the probability of observing a response in an arbitrary ob-

servations epoch Δ will be predicted from measured response rate.

RESPONSE LATENCY

Figure 12 shows the distributions of latencies (the time to the first response on those trials when a response occurred) calculated from the pigeons' data in Killeen and Hall (2001, Experiment 1). These are plotted as a function of time into the interval. There was little variation in the means (in seconds) of the distributions as a function of the baseline rate of reinforcement (left panel); they were (with standard deviations in parentheses) 0.62 (0.24), 0.64 (0.09), 0.74 (0.16), and 0.71 (0.06), respectively, for Intervals 120, 240, 480, and 960 s. The curves through the points are generated by a Poisson process and are called *gamma* or *Erlang* distributions.

The Poisson process and gamma or Erlang distributions. The best way to describe a Poisson process is by reference to the bottom of Figure 13. It shows a CBM operating at a period ϵ sufficiently short that only the rate, λ , is a relevant parameter. When it fires it increments a counter. When the counter equals a criterion n , the latent period terminates with a response and reset of the counter. This is a Poisson process. If the criterion is an integer, the resulting distribution is called the Erlang. The Erlang is a special case of the gamma distribution, for which n can be a positive real number. The distribution of the number of counts registered over some period of time, for many replications, is called a Poisson distribution. The distribution of times at which a criterion count n is attained is called an Erlang (when n is an integer) or gamma (when n is real) distribution. The Poisson process may also be thought of as a progression through n CBMs, each with the same rate parameter. If the distribution results from a progression through n CBMs, each of which has a different rate parameter, the generalized Erlang (or generalized gamma) distribution results. McGill developed its use as a stochastic latency mechanism (McGill, 1963; McGill & Gibbon, 1965), and Killeen (1975) used it to describe the distribution of adjunctive responding on periodic schedules. Although the generalized gamma could fit these data, it would require more parameters than the simple Erlang. All of these distributions ap-

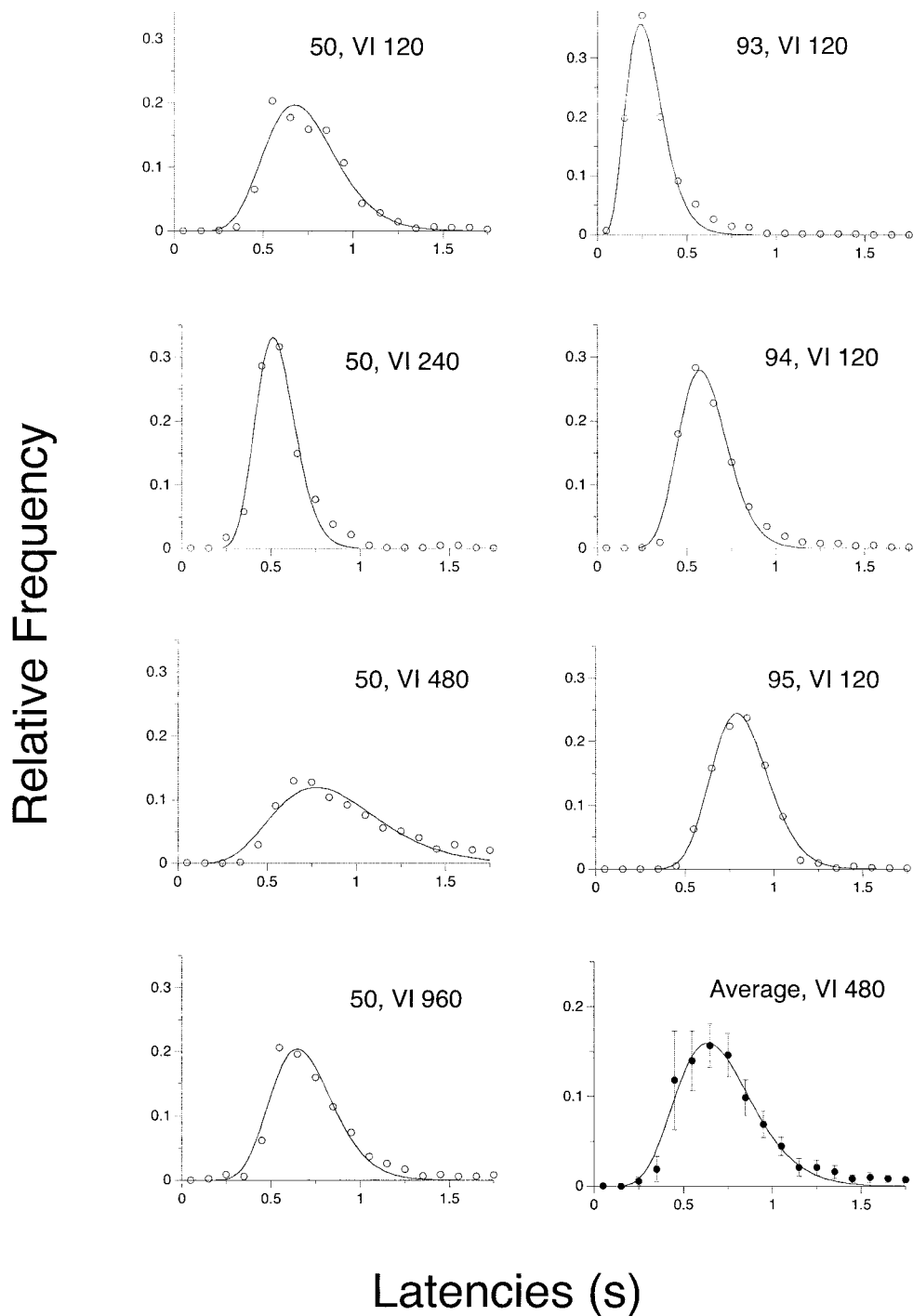
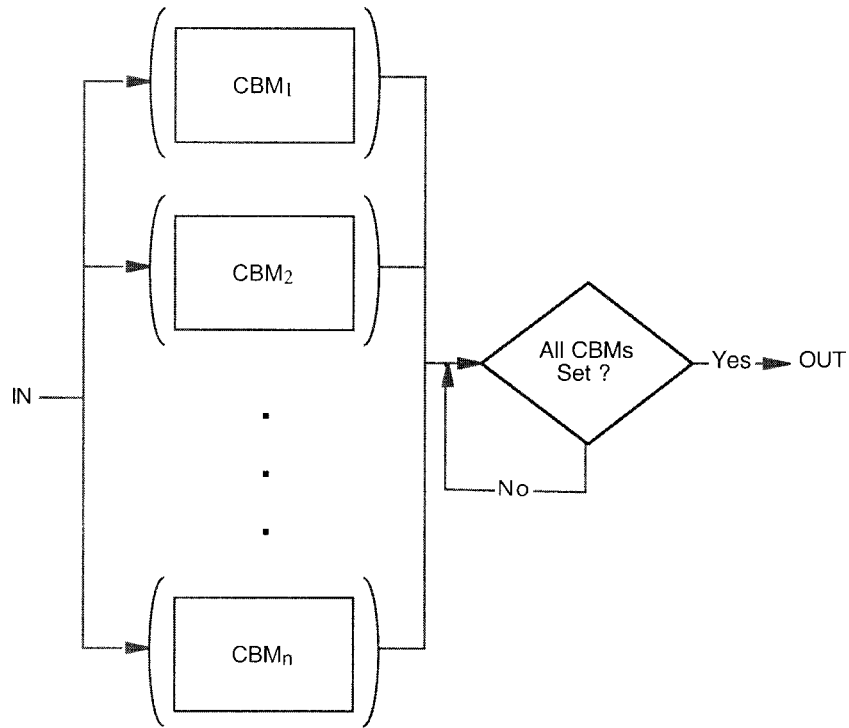


Fig. 12. Latency distributions for Pigeon 50 in each baseline condition in Killeen and Hall's (2001) Experiment 1; for each of the other pigeons in the VI 120-s baseline; and for the average over pigeons from the VI 480-s condition, with standard deviation bars. The curves are gamma densities (Equation 1).

Latency Machines

Gumbel Machine



Gamma Machine

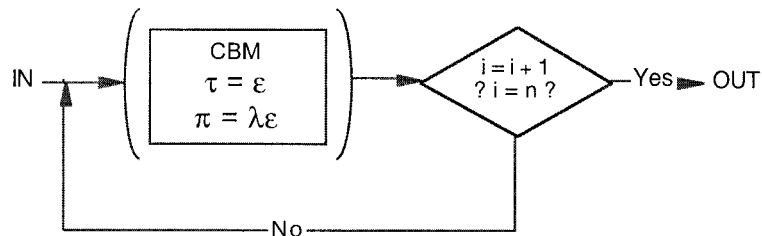


Fig. 13. Two latency machines. The Gumbel or EV machine at the top involves a number of CBMs started in unison. As soon as all have set, the machine triggers a response and exits. The gamma machine at the bottom, described by Equation 1, was used to model the data shown in Figure 12. A CBM cycles until it has fired n times, whereupon it emits a response and exits. If $n = 1$, the machine generates geometric distributions; as the period ϵ decreases, these approach the exponential distribution. For the data analyzed here, approximately a dozen states were necessary for both types of model.

proximate the normal as n increases. The Erlang density is

$$f(t) = \frac{(t/\tau)^{n-1} e^{-(t/\tau)}}{\tau(n-1)}. \quad (1)$$

This model drives the curves through the data in Figure 12. The mean (across subjects) periods of the process for increasing VIs were 42, 34, 54, and 39 ms, with mean values for n of 16, 18, 16, and 18, respectively.

If an additional, exponential, CBM is wired in series with the Erlang machine, and the number of cycles on the Erlang is large ($n > 10$), the resulting distribution is *exponential Gaussian* (exgauss), which is commonly used to model human reaction-time distributions. The exgauss distribution provided a comparable account of these data, but required a third parameter (the rate of the last CBM) to do so.

Extreme values. Another distribution that provides an equivalent fit to these data is the *extreme value* (EV) or Gumbel distribution (Evans *et al.*, 1993; Gumbel, 1958). This function was developed to describe the densities of rare events, such as floods. It is most easily explained in terms of CBMs (see top of Figure 13). Whereas the Erlang may arise from a series of sequential transitions through states (CBMs) with each transition having the same rate constant, in the EV process a number of simultaneous operations of CBMs occur, each having the same rate constant. For the largest EV process, the last CBM to fire determines when the response occurs. It is as though all of the relevant causal factors must be satisfied before the response is emitted, and each factor is represented as a Poisson process. The largest EV is a model of the interaction of necessary causal factors, the slowest of which on any trial sets the delay. The smallest EV is a model for the interaction of sufficient causal factors, any one of which may occasion a response. (The smallest EV process exits when the first CBM fires; it is not used here.) Both have been used as models of behavioral processes (Killeen, 2001).

The Gumbel distribution is the limiting EV distribution for a large number of component CBMs. The EV density for n factors is

$$f(t) = e^{-(t/\tau)} (1 - e^{-(t/\tau)})^{n-1}. \quad (2)$$

The distribution of variances accounted for by this model is not significantly different

from that of the gamma density. In both cases, if there is only one CBM ($n = 1$), the latency machine generates geometric-exponential distributions. In all cases, concatenated CBMs (Figure 13) provide relevant machines.

Long-latency responses. We thought that poorly motivated animals would miss responding on a trial because a portion of the latency distribution would be clipped by the end of the trial. That is not what happened. Decreasing probabilities of responding were caused by a combination of minor shifts in the gamma and larger increases in the probability that the first response occurred later in the trial, increases not predicted by the gamma. For the data shown in Figure 12, the proportion of latencies greater than 2 s increased as the VI mean increased, with the average proportions being 0.02, 0.03, 0.06, and 0.08 for the increasing VIs. This is predicted by neither the EV nor the gamma distributions, for which the probability of a response after 2 s is very close to 0. For representative parameters, 99.9% of the responses should have been emitted by $t = 2$ s into the trial. This calculation suggests that there are actually two processes governing the initial response on a trial: an almost immediate, relatively invariant, reflex-like response, and a desultory delayed response that is more closely related to the probability of reinforcement.

Mixtures. It is thus plausible that a mixture of latency distributions underlie these data: a gamma-like process that generates the majority of first responses and whose parameters are relatively invariant, plus another distribution of responses that are evoked with a constant probability during the course of the trial. It is this latter distribution that becomes dominant at lower reinforcement probabilities and gets truncated by the end of the trial.

This distinction between two classes of responses has precedent (Hearst, 1975; Kimble, 1967). Baron and Herpolsheimer (1999) found that increases in FR requirements increased the skew of postreinforcement-pause distributions rather than shifting the entire distribution. The presence of long-latency responses in human eyelid conditioning has actually been taken as evidence of “voluntary” instrumental responding, and was used to eliminate subjects from the experiments (Coleman, 1985; Coleman & Webster, 1988). The distinction between these two types of re-

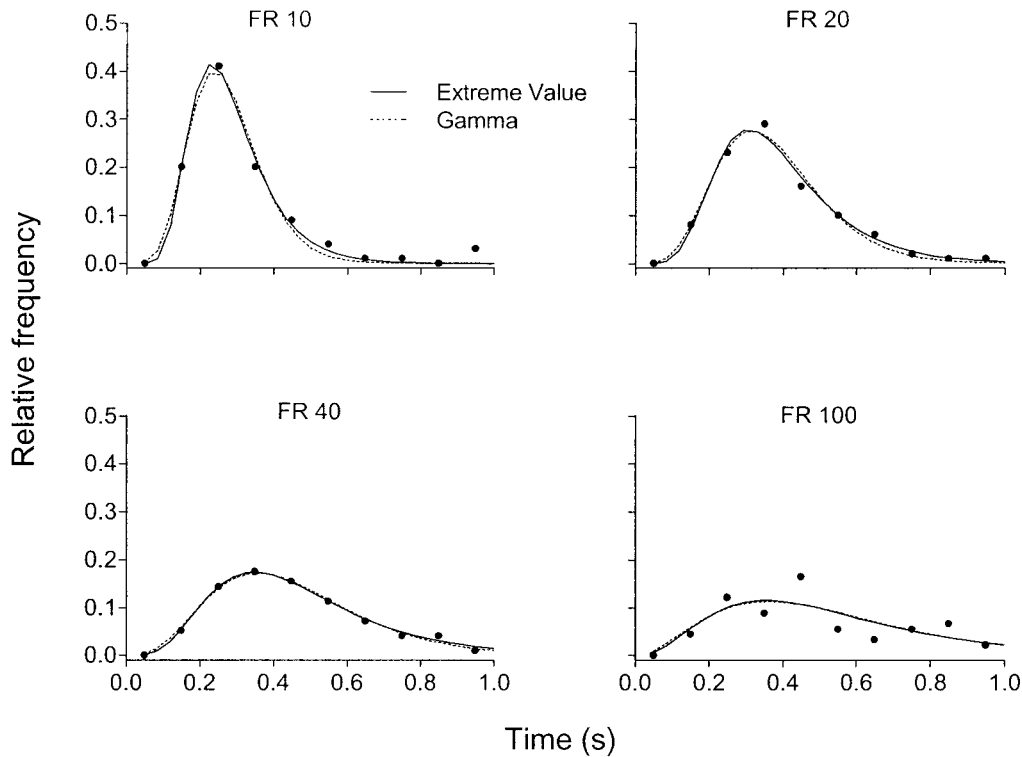


Fig. 14. Latency distributions averaged across rats at FR 10, FR 20, FR 40, and FR 100 from Experiment 1. The dashed curves are gamma densities (Equation 1); the solid curves are EV densities (Equation 2).

response resonates with Skinner's (1938) assertion that, in trials experiments, responses under control of the discriminative stimulus are "pseudo-reflexive."

In the previous section, autocorrelational analyses revealed that response run rates consist of bouts of higher likelihood of responding (of period δ) mixed with bouts of pausing. Shull et al. (2001) also identified a second, slower process in IRT distributions collected in free-operant experiments. The present analysis of latencies suggests that trial performances as a whole consist of mixtures of trials of high-probability quick responses to the onset of the conditional stimulus plus trials of late or nonresponding.

Rats on FR schedules (from Experiment 1). Figure 14 shows latency distributions for rats lever pressing on FR 10, FR 20, FR 40, and FR 100 schedules, averaged across rats. The gamma and EV densities are fit to the data, and show an equivalent ability to imitate them. With increasing size of the FR, the periods ($1/\lambda$) of the gamma were 0.03, 0.06, 0.10,

and 0.17 s and periods of the EV were 0.09, 0.13, 0.19, and 0.27 s. The number of processes (n) were 8, 7, 5, and 3 for the gamma and 15, 10, 6, and 4 for the EV. The mean and variance of the distributions (n/λ and n/λ^2 for the gamma) increase with increases in schedule value. Figure 14 shows that latency data from rats on ratio schedules may be characterized with the same arrangement of CBMs as latency data from pigeons on interval schedules.

Summary. The latency of the first response of the rats and pigeons in these experiments is well described as sequenced Poisson processes yielding gamma (Erlang) distributions. An equivalent description is provided by the EV (Gumbel) distribution. In Experiment 1 the mean and variance of the distributions generated by pigeons increased only slightly as probability of reinforcement decreased from 8% (VI 120) to 1% (VI 960). The decreasing probability of responding on a trial was greater than could be accounted for by those shifts in the distribution; trials without

a response are not simply trials with a latency drawn from the right tail of the latency distribution. This leads to further inquiry concerning the variables of which responding on a trial is a function.

THE PROBABILITY OF RESPONDING DURING A TRIAL

Figure 1 (left column) showed that the probability of responding decreased slightly in the VI 960-s condition in Experiment 1, and decreased more significantly with trials of extinction. Such was the case in the other experiments. Is the probability of responding on a trial constant from one trial to the next, or does it change as a function of the number of trials since the last response? If the probability that a trial contains at least one response (p) is constant, then the probability of two trials in a row with a response is $p \times p$, and the probability of three in a row is p^3 . The probability of one trial without a response and two with a response is $(1 - p)p^2$, and so on. This is the simplest Markov model. How well does the assumption of independence characterize these data? The last five sessions of the VI 960-s condition were analyzed to determine the probability of occurrence of each of the possible patterns of responding on all triplets of trials, and on sequences of up to seven trials in a row containing responses. There were 998 triplets for each of the 4 pigeons. The assumption of constant probability accounted for an average of only 83% of the variance in the data, ranging from 70% to 96% for the different pigeons.

There were systematic deviations from the pattern predicted by independence, as these middling coefficients of determination suggest: The probability of three trials in a row with responses was higher than predicted, as was the probability of three trials in a row without responses. A better model is invited.

The next-simplest Markov model that might account for the data posits that the pattern of probabilities resulted from a mixture of two states: When animals were in a response mode, they responded with high probability; but sometimes they went into episodes of not responding, perhaps simply turning away from the key. A two-state Mar-

kov process, in which the probability of a response given a response on the previous trial is .96 and the probability of a response given no previous response is .37, accounted for 99% of the variance in the average probabilities of these sequences (see the Appendix for the calculations, Figure 15 for the state diagram, and Figure 16 for the predictions). These two states may correspond to ones in Timberlake's behavioral system (Shettleworth, 1994; Silva, Timberlake, & Koehler, 1996; Timberlake, 1994), to paying attention or not, or simply to facing the key or not. They may correspond to B_0 , the behavior reinforced by R_0 in Herrnstein's (1970) hyperbolic law of strength. They correspond to Shull *et al.*'s (2001) two-state Markov model, with responses occurring on a trial if the animal happened to be in a visit state when the trial started and not otherwise. They may correspond to the short- and long-latency distributions described above.

Summary. The probability of responding on a trial depends on whether the subjects had responded on the previous trial, as though they were in modes, or states, of responding or not responding. The average sequences of responding were well predicted by the assumption that if they had not responded on a trial, there was only a 37% probability that they would respond on the next trial; if they had responded on a trial, there was a 96% probability that they would continue to do so. As was the case with running rates in which there were bouts of responding and hiatus from it, and as was the case with latencies in which there was a mixture of two distributions, it requires a two-state Markov model to adequately describe the probability of responding on a trial.

Each of the CBMs is characterized by two parameters, probability (π) and period (τ ; or $1/\lambda$ for the exponential models). In the case of the probability of responding from one trial to the next, it is trial onset, not time, that queries the probability gate. The machines are differentially called into play by the contingencies of reinforcement and contingencies of measurement. In the next section the general relations among these machines, their parameters, and the principal dependent variable in our field, response rate, are determined.

Response State Selection Machine

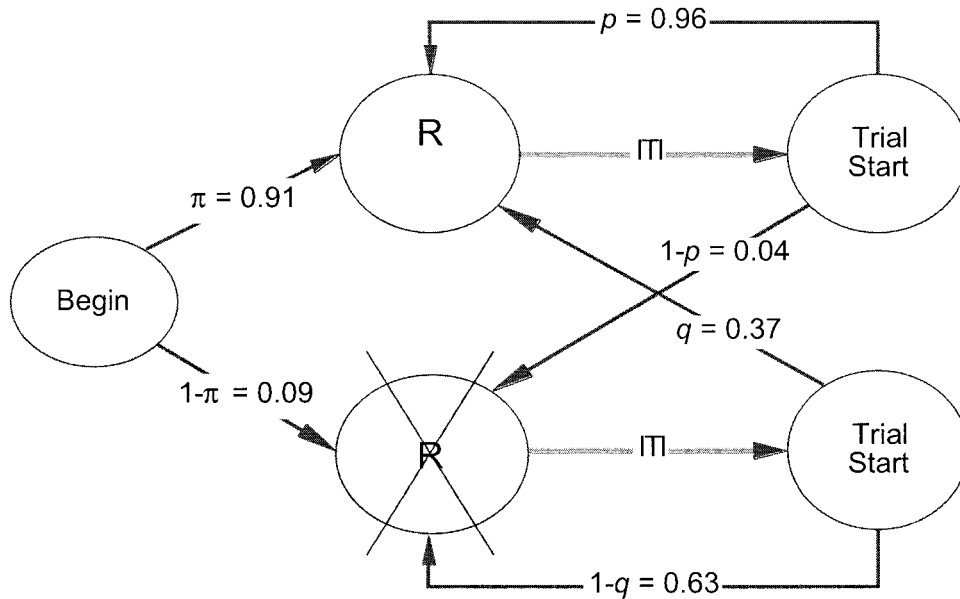


Fig. 15. The probability of responding on any trial depends on whether the organism is in a response state. This Markov model describes the data from Experiment 1 of Killeen and Hall (2001). The base probability of responding on a trial is .91; if the animal responds, then on the next trial the probability of a response increases to .96. If the animal does not respond, the probability of a response on the next trial is .37.

MAPPING RATE TO PROBABILITY

The refractory Poisson machine falls between the simple exponential model and the Palya machine in complexity, accuracy, and parsimony. It, and the shifted exponential IRT distribution that characterizes it, provide the basis for a functional relation between rate and probability. Recall that for the refractory Poisson machine each response requires δ seconds for its emission; after this refractory period there is a constant probability of emitting a response in subsequent epochs. The hazard function of this machine is 0 for δ seconds, and then abruptly rises to a horizontal line at a level of $\lambda = \pi/\delta$, indicating a rate of response initiation of λ . Analyses will start with the simplest limiting conditions of this model and build to the complete model as necessary for high probabilities of responding.

Consider first the case in which measured

response rate (b) is far below its ceiling; then the probability of observing a response in that epoch is approximately proportional to response rate:

$$p \approx \Delta b \quad 0 \leq \Delta b < 0.25. \quad (3a)$$

Conversely, within the same range $b \approx p/\Delta$.

At higher response rates the probability falls below Δb , requiring the integral of the exponential IRT density:

$$p \approx 1 - e^{-b\Delta}. \quad (3b)$$

Equation 3b is the area under an exponential decay function from 0 to Δ . Conversely, $b = -\ln(1 - p)/\Delta$.

Equation 3b applies to the complete range of probabilities only if the refractory period after a response (δ) is negligible. It reduces to Equation 3a for small Δb . When δ is non-negligible, then as response rates approach their maximum ($b_{\max} = 1/\delta$), the time available for emitting a response is increasingly

Probabilities of Responding on Sequences of Trials

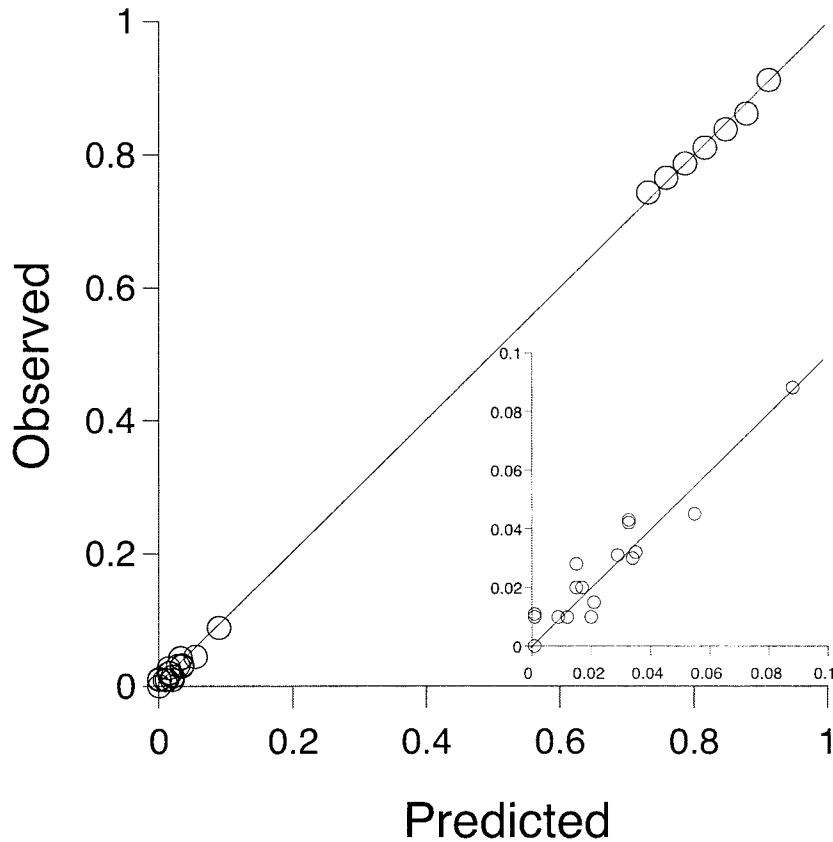


Fig. 16. The probabilities of responding on various sequences of trials are plotted against those given by the Markov model shown in Figure 15. The data are the averages over 4 pigeons from the last 1,000 trials in the VI 960-s condition of Experiment 1 of Killeen and Hall (2001). The highest probabilities correspond to 1, 2, 3, . . . , 7 trials in a row containing at least one response. The inset panel magnifies the lower decant and contains all the singlets, doublets, triplets, and quadruplets of sequences that involve at least one trial without a response.

constricted by the time required for the completion of each response. If each response requires $\delta = 0.25$ s, an animal responding at a rate of 2 per second has only $1 - 2 \times 0.25 = 0.5$ s available out of every second for initiating a response. In this case Equation 3b becomes

$$p = 1 - e^{-\lambda\Delta}, \quad (3c)$$

with the rate of initiating responses, λ , estimated as (Cox & Miller, 1965)

$$\lambda = b/(1 - \delta b), \quad b < b_{\max}. \quad (3d)$$

As the refractory period (δ) or response rate (b) decreases, this converges on the simpler

exponential map (Equation 3b), which also suffices for brief observation intervals Δ . For sufficiently low response rates and short epochs, Equation 3c reduces to Equation 3a. These maps are shown in Figure 17 for $\Delta = 1$ s, for various values of δ . The converse of Equation 3d is

$$b = \lambda/(1 + \lambda\delta), \quad (3e)$$

which predicts observed rates from instantaneous rate and dead time. This is equivalent to Shull *et al.*'s (2001) Equation 1, with $b = V'$, $\lambda = V$, and $\delta = NW$, the time occupied by a bout.

Equation 3e is sometimes called a hyper-

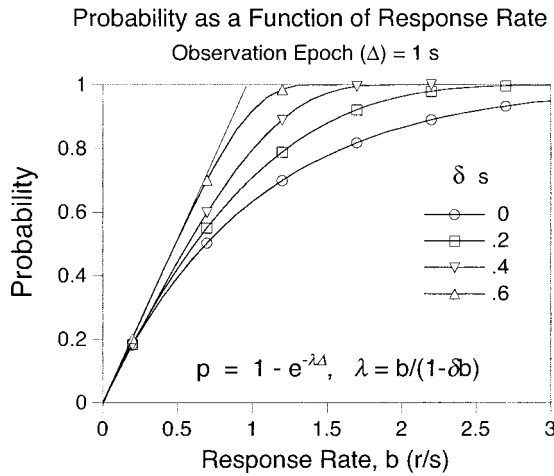


Fig. 17. The maps between response rate and response probability generated by Equation 3a (the straight line from 0,0 to 1,1); Equation 3b (circles); and Equation 3c with different values of refractory period δ , measured in seconds. Measured response rate b is given in responses per second.

bolic function. If λ is proportional to rate of reinforcement ($\lambda = aR$), Equation 3e becomes a fundamental equation in the mathematical principles of reinforcement (Killeen, 1994). The converse of the complete probability model for rate is $b = 1/[\delta - \Delta/\ln(1 - p)]$.

Figure 18 gives the map between the probability of a response within a Δ -s epoch for each pigeon in Experiment 1 of Killeen and Hall (2001) plotted against response rate. The data are from the first three sessions of extinction from the VI 120-s schedule, using only trials with at least one response. From these approximately 1,000 trials, a random process in the computer selected a trial, and the running rate on that trial was calculated. Another random process selected an observation epoch of Δ s randomly from the portion of that trial after the latency. The presence or absence of a response in that epoch was registered. This process continued for 1,000 samples with replacement. The probability of a response was calculated as the proportion of observations from trials with a running rate of $b \pm 0.1 \text{ s}^{-1}$ that contained at least one response. This was plotted against the running rate b . The curves are from Equation 3c, with δ fixed at 0.29, 0.20, 0.27, and 0.10 s for Pigeons 50, 93, 94, and 95, respectively.

The same values were used for all observation epochs Δ .

Notice in Figure 18 that a straight line from the origin to (1, 1) will account for a sizable amount of the variance in the data at low response rates. That is the approximate prediction of all of these models: At low response rates, $p \approx b\Delta$. It validates Skinner's sense that response rate is important because it permits an estimate of response probability: "Our basic datum is the rate at which such a response is emitted. . . . Such a datum is closely associated with the notion of probability of action" (Ferster & Skinner, 1957/1997, p. 7), and "Perhaps most important of all, frequency of response is a valuable datum just because it provides a substantial basis for the concept of probability of response—a concept toward which a science of behavior seems to have been groping for many decades" (Skinner, 1961, p. 74).

Summary. The shifted exponential distribution corresponding to the refractory Poisson process provides a good description of the data and approximates the more precise but less wieldy Palya model. The key parameter of the Poisson is its intensity, or instantaneous rate, λ . When the refractory period after a response (δ) is small, response rate measured in the conventional manner (b) is a good estimate of λ . When response rates approach their ceiling ($1/\delta$), however, the hyperbolic relation between measured rate (b) and instantaneous rate (λ) becomes salient, as shown in Equation 3e.

There is a close relation between probability and rate, but it is not one of equivalence. The probability of emission of a response in any epoch is a concave function of the instantaneous response rate, approaching 1.0 as rates approach their ceiling (Figures 17 through 19). At low response rates ($b\Delta \ll 1$), the probability of a response in any epoch of length Δ is $p \approx b\Delta$ (Equation 3a). At higher rates, $p = 1 - e^{-\lambda\Delta}$ (Equation 3b), with $\lambda \approx b$. At still higher rates, account must be taken of the responses that would have been emitted but fell into the refractory period (Cox & Miller, 1965). This is accomplished by Equation 3c. The refractory Poisson process embodied by Equation 3e is consistent with a hyperbolic relation between response strength and response rate posited by various theories of motivation (e.g., Killeen, 1994).

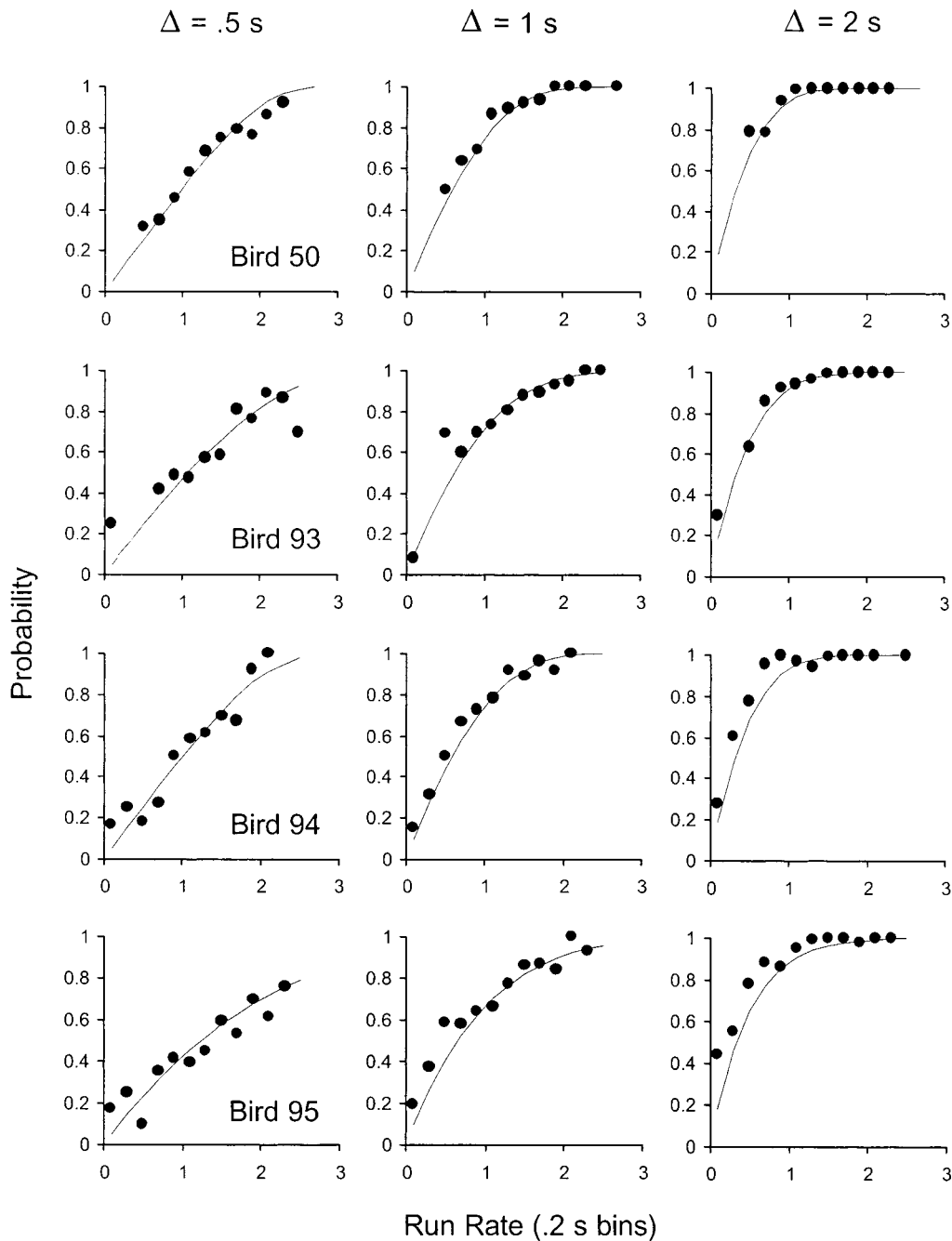


Fig. 18. The probability of making a response during a Δ -second epoch as a function of running rate on each trial for the first three extinction sessions after the VI 120-s condition of Experiment 1 of Killeen and Hall (2001) for each pigeon. The curves are from Equation 3c.

Global Response Probability

The probability of seeing a response during a brief epoch Δ randomly sampled from a session may be found by combining the

probability of choosing a trial with a response (p), the probability of selecting the run state of that trial ($1 - L/T$; where L is latency and T trial duration), and the probability of not

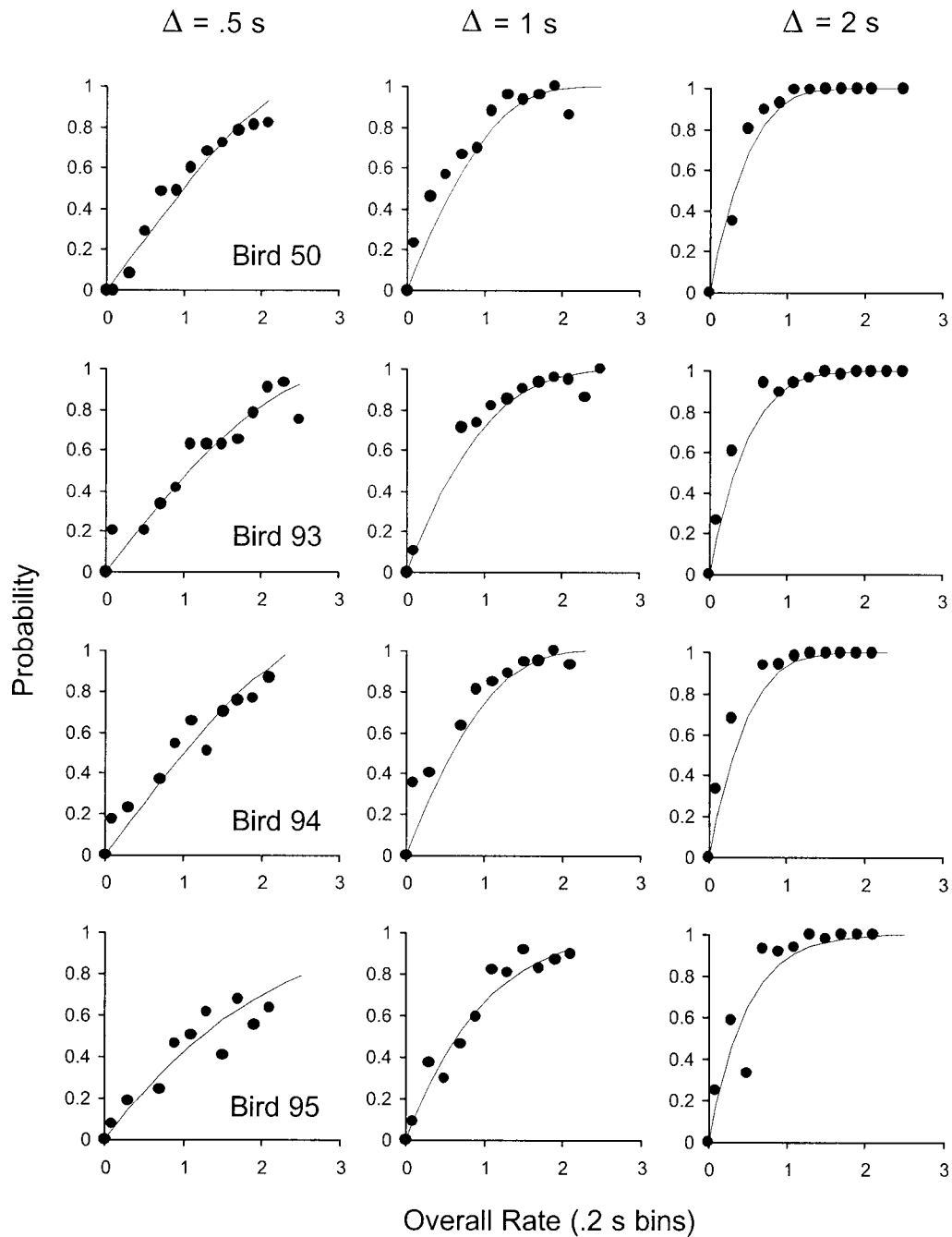


Fig. 19. The probability of making a response during a Δ -second epoch as a function of overall response rates on each trial. The data are from the same sessions as those shown in Figure 18. The curves are from Equation 3c, with b now signifying overall response rate, and with δ fixed at 0.29, 0.20, 0.27, and 0.10 for the top, second, third, and fourth rows, respectively (the same values used in Figure 18).

falling into an interresponse interval (Equation 3c):

$$P = p(1 - L/T)(1 - e^{-\lambda\Delta}). \quad (4)$$

Equation 4 combines the three key factors: trial probability, latency, and the rate-to-probability map. It is definitive, but puts perhaps too fine a point on the analyses, requiring more information than necessary for a good description of the present data. This is because the three factors in Equation 4 are usually positively correlated (Killeen & Hall, 2001). We may shift the burden of the first two terms in Equation 4 onto the third, Equation 3c and 3d. With b now representing overall response rate in those equations (i.e., the number of responses divided by time available for responding: B), it provides a map between overall response rate and global probability of responding. This is shown in Figure 19, in which the probability of encountering a response in 1,000 random samples of Δ -s epochs over all parts of all the trials of the first three sessions of extinction from Experiment 1 of Killeen and Hall (2001), VI 120 s, are predicted from the overall response rate on those trials. The same analytic procedure and the same values of refractory period δ were used as in Figure 18, with all trials in a session sampled. It is manifest that global response rates both capture the contributions of the static temporal properties (Equation 4) and permit prediction of global response probabilities (Figure 19). Because rates often vary within the course of a session due to extinction, spontaneous recovery, satiation, and so forth, the success of such a rate-to-probability map over these heterogeneous conditions is remarkable.

GENERAL DISCUSSION

Levels of Analysis

Many different levels of descriptions of behavior—from qualitative through molar to molecular—are possible. The current article provides several levels of analysis, offering characterizations at one level that are then improved at the next. The cost of refinement is paid in parameters and complexity, which eventually frustrate intuition.

Overall response rates comprise three components, based on conditional probabilities:

1. The probability of responding on any sequence of trials is well predicted by the probability of responding conditional on the presence of a response on the previous trial (Figure 15).

2. The probability of the first response given trial onset is well described by gamma and EV distributions (Figures 12 through 14). In addition to the responses falling within these distributions, there was a collection of late responses that accounted for about 5% of the trial initiations under baseline conditions and an increasing percentage through extinction.

3. The probability of a response given a response in the previous Δ seconds is given by the IRT distribution. These distributions (e.g., Figure 11) were most precisely characterized as the output from a Palya machine in which pulses from a pacemaker with a variable period are registered probabilistically. The refractory Poisson emitter provides a simpler model that gives rise to a shifted exponential distribution of IRTs (e.g., Figure 10 and the top panel of Figure 7), and is consistent with the hyperbolic law of response strength (de Villiers & Herrnstein, 1976).

These variables may be combined into overall response rate or probability (Equation 4). Depending on experimental context, each will weigh more or less heavily in their contribution to response rate and in the information they provide concerning response strength. As is the case for any summary statistic, details of the particulars are sacrificed to the generality and simplicity of an omnibus measure. Overall rate both correlates highly with the latent dimension strength (Killeen & Hall, 2001) and predicts probability of responding within any epoch of time (Figure 19). It does so even though it is a gloss over nonhomogenous epochs—trials without a response, latent periods, refractory periods, and bouts of responding at high rates.

Probability is not rate, but it arises as the result of sampling the output of an emitter that operates some proportion of the time (during bouts); when operating, it does so at some rate (λ), and its output is realized as a measurable response some proportion of the time (p). The summary measure of response rate provides an excellent description of the state of many of the static properties of the response, and possesses both internal and external validity. Response rate is not an inter-

val scale of strength; due to physical limitations on rate, it is concave, with increases at the high end requiring stronger motivational operations than equal increases at the low end. Instantaneous response rate λ , which corrects response rate for this constraint, may be such an interval scale, but establishing that requires a network of relations among operations, measurements, and outcomes that has yet to be accomplished.

Probability Machines

The probability of emitting a response within Δ seconds after the previous response is p . This is the variable that controls the shape of the exponential tail of the IRT distribution, which has a rate parameter $\lambda = -\ln(1 - p)/\Delta$, which is approximately equal to p/Δ . Response rates give us information about p , the probability of emitting a response. At low rates p is proportional to response rate, but at higher rates it bends beneath its unit ceiling. The exponential integral (Figure 17) describes that curvature. When an observational epoch is an arbitrary temporal interval (say, $\Delta = 1$ s) drawn randomly from the trials and contingent only on the presence of a lit center key, the global probability of a response is well predicted from overall response rates, as given by Equation 3c and shown in Figure 18. Further analysis based on categorizing the data according to certain conditions reveals finer structures. The conditionals—the givens—that were effective in the present analysis were the presence of a response on the previous trial, the trial onset, the occurrence of a prior response on that trial, and the lapse of a refractory period. For Palya machines, the time since the last response was also informative. The stochastic processes associated with these levels are given by the associated probability machines and their flowcharts.

The final proof of successful analysis is a synthesis of the components into the whole (Teitelbaum & Pellis, 1992). Figure 20 shows how the various stochastic machines may be concatenated to emulate the behavior of organisms.

Behavior as Mixtures of States

Timberlake (e.g., 1999, 2000) has regularly reminded us that the free operant is at less liberty than its name claims. Animals natural-

ly engage in a portfolio of behavior, describable with an ethogram, that often overwhelms analyses. The majority of articles in this journal describe aspects of a single operant under the control of a single kind of reinforcer. A second response key is sometimes employed to give the animal a second route to the same goal. Sometimes alternate kinds of behavior are recorded by transducers that may still leave too much information (e.g., Pear, 1985) or too little (e.g., Killeen & Bizo, 1998) for the analyst to formulate. Sometimes the alternate behavior is merely inferred by gaps in the target responding (Herrnstein, 1970; Rachlin, Battalio, Kagel, & Green, 1981; Shull et al., 2001). When explicit alternative responses are identified, the analysis often becomes rich and surprising: The statistics of alternation between bouts of responses may indicate stress due to parasitism or pregnancy, as revealed by their fractal dimension (Alados, Escos, & Emlen, 1996); the motivational control of elements of a sequence may vary as a function of their location in the sequence (Balleine, Garner, Gonzalez, & Dickinson, 1995); patterns of responding may be embedded in larger patterns in a self-similar structure (Cole, 1995); and the relative frequency of independent responses may be distributed as an exponential function of their activation energy (Hanson, 1991).

Multiplication of essentials. Understanding the state of an organism is key to understanding how and why it behaves (Timberlake & Silva, 1994). The *state of the organism* has too often been a generic reference to organismic variables that are not understood—a euphemism for error variance. But it can have a more precise meaning, as the principal component of a constellation of variables whose specification increases our ability to predict and control behavior. It is what is to be *given* in a conditional statement. If an animal responded on the previous trial it is in a different state than if it had not (Figure 15)—it behaves in a predictably different manner. Immediately after one response another response is impossible; *refractory* is the name of a state whose specification increases our ability to predict the absence of responses. *Strength* is the measure of a state that binds together our key dependent variables. Whereas such essential constructs should not be

Synthesizer

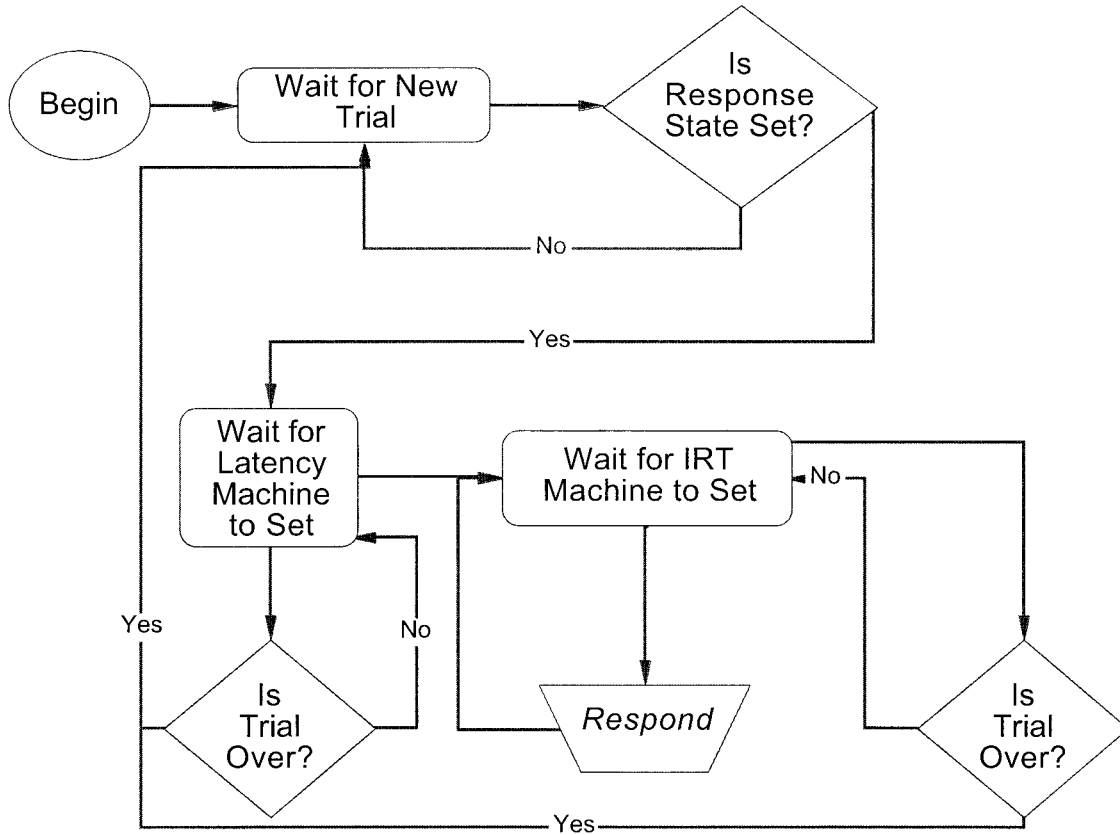


Fig. 20. The probability, latency, and IRT machines are assembled by a synthesizer, which generates a stream of responses emulating those of the subjects in these experiments.

multiplied beyond their necessity, when necessary they should not be omitted.

State transition diagrams such as the behavior systems theory diagrams of Timberlake; the Markov models of Myerson and Hale (1988), Gibbon (1995), and Shull et al. (2001), and probability machines such as those displayed here, are steps toward non-gratuitous characterization of states. The biometrics literature provides more technical treatments of sophisticated state-space models (e.g., Mangel & Clark, 1988); Staddon (1993) makes a case for their application in behavior analysis. The transition between such states occurs stochastically when certain conditions are met. Those conditions may be

easily identified, or they may be obscure. If the transition *always* occurs given x , we speak of *cause*; if it *never* occurs, we speak of *inhibition*; if it *sometimes* occurs, we speak of *probability*. As a measure of strength, probability is a fundamental variable in the analysis of behavior.

Probability Versus Rate

Probability machines provide a characterization that is the dual of response rate, much as nuclear particles are duals of waves. Which is a better measure, probability or rate? This depends on the operations available for their measurement.

Probabilities are measured by repeatedly

sampling the behavior of a system over experimentally defined epochs. The epoch should be brief enough so that the response will not occur many times within it, or the measure saturates; uniform probabilities of 0 or 1 are uninformative. The information transmitted by an observation is maximum when epochs are chosen so that all outcomes of interest are equally likely to occur within it. If the outcomes are binary, such as presence or absence of a response, this is 50%. At that probability, solution of Equation 3b predicts a response rate $b \approx 0.7/\Delta$, assuming no refractory period ($\delta = 0$). For $\delta = 0.25$ s and an observation epoch of 1 s, rate decreases to $b \approx 0.6/s$. The predictions in Figures 18 and 19 provide a map between rate and probability that can be traveled in either direction.

Probability assessments are often metaphorical extensions of the sampling operation that defines objective probabilities. Subjective probabilities are based on aggregation of real or imagined events into a population. A number between 0 and 1 is then assigned as a measure of confidence in a particular outcome, based on the number of relevant instances in the imagined population, the latency or difficulty in imagining relevant instances, and so on. Subjective probabilities are useful in everyday life; but because the construction of the relevant population and the sampling from it are not verifiable or uniform across individuals (Keren & Teigen, 2001), the resulting estimates reflect the history of the individual as much as objective probabilities.

Response rates are measured by summing responses and dividing by elapsed time. Rates provide an exhaustive sampling without replacement of the interval of interest, and are thus the most efficient possible estimators of the corresponding probability. Measured response rate (b) is not the same as the rate of initiating responses (λ), just as IRTs are not the same as the time between responses. The difference is δ . The average time between the start of one response and the start of the next, $1/b$, equals the average time between responses, $1/\lambda$, plus the response duration δ : $1/b = 1/\lambda + \delta$. Take the reciprocal to derive the relation between the two rates b and λ given by Equations 3d and 3e: $b = \lambda/(1 + \lambda\delta)$. This is sometimes called a hyperbolic relation. If the rate of initiating responses is

proportional to the rate of reinforcement ($\lambda = aR$), then $b = (R/\delta)/(R + R_0)$, where $R_0 = 1/a\delta$. The fundamental hyperbolic relation between reinforcement rate and response rate (Herrnstein, 1974, 1979) may thus derive its shape from this simple fact of measurement.

In sum, probabilities tell us the relative frequency with which responses of a similar nature will fall within an observation epoch. The reciprocal of a probability tells us the number of epochs that we must expect to sample before seeing the event. Probabilities are useful for observations that form a natural epoch—trials or other occasions initiated by a discriminative stimulus—and for brief epochs from a stream of recurrent responses. When available, rates provide an efficient measure of response strength, which predicts other variables such as response latency and probability (Equations 3 and 4). Observational conveniences and efficiencies determine the best way to measure strength.

Summary and Conclusions

Behavior at large is a constellation of response states, within which different target responses occur with some probability. The onset of a salient stimulus is often a good predictor of a new state, and time since the onset of the state is often a good predictor of the probability of a response. If the response does not cause the animal to exit the state, responding may continue at a characteristic rate. States and transitions between them are fundamental units for the analysis of behavior, because they define homogenous epochs for observation. The three-term contingency of discriminative stimulus, response, and reinforcer is a formula for describing behavioral states, and thus is a verbal allusion to Figure 20. Contingencies of reinforcement specify the rules by which the experimenter or nature shapes organisms to transit between states. Parsing behavior into states increases the predictability of behavior. Accurate state diagrams epitomize our understanding of behavior. It is within this context that rate and probability provide dual descriptions of response strength.

Both the latency of the first response in a state and the ensuing interresponse times may be analyzed with probability machines, a common element of which is the clocked Ber-

noulli module (CBM). This device adds time to Markov models (Figure 4). The instantaneous rate parameter of CBMs, λ , specifies the rate of state transition. In IRT machines this is transition from a nonresponse to a response. Measured response rates may be less than λ due to refractory periods (δ), or to probabilistic disengagement and reengagement in the response bouts—that is, transition among states with different parameters. A general model of IRT machines is shown in Figure 8, and examples of latency machines appear in Figure 13.

Response rate is an efficient but biased estimator of the instantaneous rate, λ : It must be corrected for the physical limits on rate due to refractory periods. Response rate is also a good estimator of probability, but must be corrected for the nonlinear relation between the probability of responding during an epoch and the rate during the state from which that epoch is drawn. These corrections are provided in Equation 3 and are graphed in Figures 17 through 19. Killeen and Hall (2001) showed that response probability, latency, and rate were highly correlated with one another and with a factor called strength; and that overall response rate is an excellent predictor of strength. This paper provides a common modeling language based on units of probability (p and π) and time ($1/\lambda$ and τ) for those dependent variables. Future research may show that operations that affect strength modify those units in simple and predictable ways. It is our hypothesis that the rate of initiating responses, λ , will prove to be the best single index of strength.

REFERENCES

- Alados, C. L., Escos, J. M., & Emlen, J. M. (1996). Fractal structure of sequential behavior patterns: An indicator of stress. *Animal Behaviour*, *51*, 437–443.
- Anger, D. (1956). The dependence of interresponse times upon the relative reinforcement of different interresponse times. *Journal of Experimental Psychology*, *52*, 145–161.
- Bachrach, A. J. (1966). A simple method of obtaining a scatter distribution of off-key pigeon pecking. *Journal of the Experimental Analysis of Behavior*, *9*, 152.
- Balleine, B. W., Garner, C., Gonzalez, F., & Dickinson, A. (1995). Motivational control of heterogeneous instrumental chains. *Journal of Experimental Psychology: Animal Behavior Processes*, *21*, 203–217.
- Baron, A., & Herpolsheimer, L. R. (1999). Averaging effects in the study of fixed-ratio response patterns. *Journal of the Experimental Analysis of Behavior*, *71*, 145–153.
- Bharucha-Reid, A. T. (1960). *Elements of the theory of Markov processes and their applications*. New York: McGraw-Hill.
- Bizo, L. A., & Killeen, P. R. (1997). Models of ratio schedule performance. *Journal of Experimental Psychology: Animal Behavior Processes*, *23*, 351–367.
- Blough, D. S. (1963). Interresponse time as a function of continuous variables: A new method and some data. *Journal of the Experimental Analysis of Behavior*, *6*, 237–246.
- Bower, G. H. (1994). A turning point in mathematical learning theory. *Psychological Review*, *101*, 290–300.
- Church, R. M., Broadbent, H. A., & Gibbon, J. (1992). Biological and psychological description of an internal clock. In I. Gormezano & E. A. Wasserman (Eds.), *Learning and memory: The behavioral and biological substrates* (pp. 105–128). Hillsdale, NJ: Erlbaum.
- Cole, B. J. (1995). Fractal time in animal behavior: The movement activity of *Drosophila*. *Animal Behaviour*, *50*, 1317–1324.
- Coleman, S. R. (1985). The problem of volition and the conditioned reflex. Part I: Conceptual background, 1900–1940. *Behaviorism*, *13*, 99–124.
- Coleman, S. R., & Webster, S. (1988). The problem of volition and the conditioned reflex. Part II: Voluntary-responding subjects. *Behaviorism*, *16*, 17–49.
- Conover, K. L., Fulton, S., & Shizgal, P. (2001). Operant tempo varies with reinforcement rate: Implications for measurement of reward efficacy. *Behavioural Processes*, *56*, 85–101.
- Cox, D. R., & Miller, H. D. (1965). *The theory of stochastic processes*. New York: Wiley.
- de Villiers, P. A., & Herrnstein, R. J. (1976). Toward a law of response strength. *Psychological Bulletin*, *83*, 1131–1153.
- Dews, P. B. (1981). History and present status of rate-dependency investigations. In T. Thompson, P. Dews, & W. McKim (Eds.), *Advances in behavioral pharmacology* (Vol. 3, pp. 111–118). New York: Academic Press.
- Estes, W. K. (1950). Toward a statistical theory of learning. *Psychological Review*, *57*, 94–107.
- Evans, M., Hastings, N., & Peacock, B. (1993). *Statistical distributions* (2nd ed.). New York: Wiley.
- Ferster, C. B., & Skinner, B. F. (1997). *Schedules of reinforcement*. Acton, MA: Copley Publishing Group. (Original work published 1957)
- Gentry, G. D., Weiss, B., & Laties, V. G. (1983). The microanalysis of fixed-interval responding. *Journal of the Experimental Analysis of Behavior*, *39*, 327–343.
- Gibbon, J. (1995). Dynamics of time matching: Arousal makes better seem worse. *Psychonomic Bulletin & Review*, *2*, 208–215.
- Gumbel, E. J. (1958). *Statistics of extremes*. New York: Columbia University Press.
- Hanson, S. J. (1991). Behavioral diversity, search and stochastic connectionist systems. In M. L. Commons, S. Grossberg, & J. E. R. Staddon (Eds.), *Quantitative analysis of behavior: Neural network models of conditioning and action* (pp. 295–344). Hillsdale, NJ: Erlbaum.
- Hearst, E. (1975). The classical-instrumental distinction: Reflexes, voluntary behavior, and categories of associative learning. In W. K. Estes (Ed.), *Handbook of learning and cognitive processes* (Vol. 2, pp. 181–223). Mahwah, NJ: Erlbaum.
- Hemmes, N. S. (1975). Pigeons' performance under dif-

- ferential reinforcement of low rate schedules depends upon the operant. *Learning & Motivation*, 6, 344–357.
- Herrnstein, R. J. (1970). On the law of effect. *Journal of the Experimental Analysis of Behavior*, 13, 243–266.
- Herrnstein, R. J. (1974). Formal properties of the matching law. *Journal of the Experimental Analysis of Behavior*, 21, 159–164.
- Herrnstein, R. J. (1979). Derivatives of matching. *Psychological Review*, 86, 486–495.
- Heyman, G. (1988). How drugs affect cells and reinforcement affects behavior: Formal analogies. In M. L. Commons, R. M. Church, J. R. Stellar, & A. R. Wagner (Eds.), *Quantitative analyses of behavior: Biological determinants of reinforcement* (Vol. 7, pp. 157–182). Mahwah, NJ: Erlbaum.
- Johnson, L. M., & Morris, E. K. (1987). When speaking of probability in behavior analysis. *Behaviorism*, 15, 107–129.
- Kapur, J. N. (1989). *Maximum entropy models in science and engineering*. New York: Wiley.
- Keren, G., & Teigen, K. H. (2001). The probability-outcome correspondence principle: A dispositional view of the interpretation of probability statements. *Memory & Cognition*, 29, 1010–1021.
- Killeen, P. R. (1975). On the temporal control of behavior. *Psychological Review*, 82, 89–115.
- Killeen, P. R. (1982). Incentive theory. In D. J. Bernstein (Ed.), *Nebraska symposium on motivation, 1981: Response structure and organization* (pp. 169–216). Lincoln: University of Nebraska Press.
- Killeen, P. R. (1994). Mathematical principles of reinforcement. *Behavioral and Brain Sciences*, 17, 105–172.
- Killeen, P. R. (2001). Writing and overwriting short-term memory. *Psychonomic Bulletin & Review*, 8, 18–43.
- Killeen, P. R., & Bizo, L. A. (1996). The response dimension. In K. H. Pribram & J. King (Eds.), *Learning as self-organization* (pp. 141–154). Mahwah, NJ: Erlbaum.
- Killeen, P. R., & Bizo, L. A. (1998). The mechanics of reinforcement. *Psychonomic Bulletin & Review*, 221–238.
- Killeen, P. R., & Hall, S. S. (2001). The principal components of response strength. *Journal of the Experimental Analysis of Behavior*, 75, 111–134.
- Kimble, G. A. (1967). The concept of reflex and the problem of volition. In G. A. Kimble (Ed.), *Foundations of conditioning and learning* (pp. 144–154). New York: Appleton-Century-Crofts.
- Luce, R. D. (1986). *Response times: Their role in inferring elementary mental organization*. New York: Oxford University Press.
- Mangel, M., & Clark, C. (1988). *Dynamic modeling in behavioral ecology*. Princeton, NJ: Princeton University Press.
- Mazur, J. E., & Hyslop, M. E. (1982). Fixed-ratio performance with and without a postreinforcement timeout. *Journal of the Experimental Analysis of Behavior*, 38, 143–155.
- McGill, W. J. (1962). Random fluctuations of response rate. *Psychometrika*, 27, 3–17.
- McGill, W. J. (1963). Stochastic latency mechanisms. In R. R. Bush & E. Galanter (Eds.), *Handbook of mathematical psychology* (Vol. 1, pp. 309–360). New York: Wiley.
- McGill, W. J., & Gibbon, J. (1965). The general-gamma distribution and reaction times. *Journal of Mathematical Psychology*, 2, 1–18.
- Miller, A. I. (1984). *Imagery in scientific thought*. Boston: Birkhäuser.
- Myerson, J., & Hale, S. (1988). Choice in transition: A comparison of melioration and the kinetic model. *Journal of the Experimental Analysis of Behavior*, 49, 291–302.
- Palya, W. L. (1991). Laser printers as powerful tools for the scientific visualization of behavior. *Behavior Research Methods, Instruments, & Computers*, 23, 277–282.
- Palya, W. L. (1992). Dynamics in the fine structure of schedule-controlled behavior. *Journal of the Experimental Analysis of Behavior*, 57, 267–287.
- Pear, J. J. (1985). Spatiotemporal patterns of behavior produced by variable-interval schedules of reinforcement. *Journal of the Experimental Analysis of Behavior*, 44, 217–231.
- Ploog, B. O., & Zeigler, H. P. (1996). Effects of food-pellet size on rate, latency, and topography of auto-shaped key pecks and gapes in pigeons. *Journal of the Experimental Analysis of Behavior*, 65, 21–35.
- Rachlin, H., Battalio, R., Kagel, J., & Green, L. (1981). Maximization theory in behavioral psychology. *Behavioral and Brain Sciences*, 4, 371–417.
- Rau, J. C. (1997). *Molecular structure in multiple schedules*. Unpublished masters thesis, University of Canterbury, Christchurch, New Zealand.
- Ray, R. C., & McGill, W. (1964). Effects of class-interval size upon certain frequency distributions of interresponse times. *Journal of the Experimental Analysis of Behavior*, 7, 125–127.
- Reynolds, G. S., & Catania, A. C. (1961). Behavioral contrast with fixed-interval and low-rate reinforcement. *Journal of the Experimental Analysis of Behavior*, 4, 387–391.
- Ross, S. R. (1997). *Introduction to probability models* (6th ed.). San Diego, CA: Academic Press.
- Shettleworth, S. J. (1994). Commentary: What are behavior systems and what are they for? *Psychonomic Bulletin & Review*, 1, 451–456.
- Shull, R. L. (1991). Mathematical description of operant behavior: An introduction. In I. H. Iversen & K. A. Lattal (Eds.), *Experimental analysis of behavior* (Vol. 2, pp. 243–282). New York: Elsevier.
- Shull, R. L., Gaynor, S. T., & Grimes, J. A. (2001). Response rate viewed as engagement bouts: Effects of relative reinforcement and schedule type. *Journal of the Experimental Analysis of Behavior*, 75, 247–274.
- Silva, F. J., Timberlake, W., & Koehler, T. L. (1996). A behavior systems approach to bidirectional excitatory serial conditioning. *Learning and Motivation*, 27, 130–150.
- Skinner, B. F. (1938). *The behavior of organisms*. New York: Appleton-Century-Crofts.
- Skinner, B. F. (1961). *Cumulative record* (Enlarged ed.). New York: Appleton-Century-Crofts.
- Skinner, B. F. (1988). What is the experimental analysis of behavior? *Journal of the Experimental Analysis of Behavior*, 9, 213–218.
- Staddon, J. E. R. (1993). The conventional wisdom of behavior analysis. *Journal of the Experimental Analysis of Behavior*, 60, 439–447.
- Teitelbaum, P., & Pellis, S. (1992). Toward a synthetic physiological psychology. *Psychological Science*, 3, 4–20.
- Timberlake, W. (1994). Behavior systems, association-

- ism, and Pavlovian conditioning. *Psychonomic Bulletin & Review*, 1, 405–420.
- Timberlake, W. (1999). Biological behaviorism. In W. O'Donahue & R. Kitchener (Eds.), *Handbook of behaviorism* (pp. 243–284). New York: Academic Press.
- Timberlake, W. (2000). Motivational modes in behavior systems. In R. R. Mowrer & S. B. Klein (Eds.), *Handbook of contemporary learning theories* (pp. 155–209). Mahwah, NJ: Erlbaum.
- Timberlake, W., & Silva, F. J. (1994). Observation of behavior, inference of function, and the study of learning. *Psychonomic Bulletin & Review*, 1, 73–88.
- Wing, A. M., & Kristofferson, A. B. (1973). Response delays and the timing of discrete motor responses. *Perception and Psychophysics*, 14, 5–12.

Received September 24, 2001
Final acceptance May 5, 2002

APPENDIX

Extreme Value (Gumbel) Distributions

The EV distribution is the limiting distribution of the largest observation of a set of independent identical exponential variates. Its distribution function is

$$F(t) = \exp[-\exp[-(t - m)/s]]. \quad (\text{A1})$$

Changing the sign of t gives the distribution of the smallest observation of the set. The density is the derivative of the distribution function, which in this case is simply $f(t) = \exp[-(t - m)/s]/s * F(t)$. This is the asymptotic distribution for a large n . The probabil-

ities may also be calculated directly, as given by Equation 2 in the text.

Two-State Markov Model for Probability of Responding on Any Trial

Call the base probability of making a response on any trial π . In the two-state Markov model the probability of making a response on a trial given that the animal had just responded on the previous trial, $p(1|1)$, is p ; the probability of a response given that the animal had not responded on the previous trial, $p(1|0)$, is q . These are displayed in Figure 15. The complementary probabilities are $1 - p$ for a nonresponse given a response [$p(0|1)$] and $1 - q$ for a nonresponse given a nonresponse [$p(0|0)$]. The long-run probability of observing a response on a trial when it is not known whether a response occurred on the previous trial is π , and this is related to p and q as $\pi = q/(1 + q - p)$ (see, e.g., Ross, 1997, p. 172 ff.). The probability of a string such as 111, $p(111)$, is then $\pi p p$; $p(101) = \pi(1 - p)q$; $p(110) = \pi p(1 - p)$; $p(011) = (1 - \pi)q p$; and so on. Two degrees of freedom are utilized to set the parameters: π was set equal to the obtained base rate (0.912), and p was inferred from the series 11111 [viz., as $(0.787/\pi)^{1/4}$]. Figure 16 plots the predictions of this Markov model against the data.