The Predictably Unpredictable Operant

Allen Neuringer and Greg Jensen
Reed College and Columbia University

Animals can learn to repeat a response when reinforcement is contingent upon accurate repetitions or to vary when reinforcement is contingent upon variability. In the first case, individual responses can readily be predicted; in the latter, prediction may be difficult or impossible. Particular levels of variability or (un)predictability can be reinforced, including responses that approximate a random model. Variability is an operant dimension of behavior, controlled by reinforcers, much like response force, frequency, location, and topography. As with these others, contingencies of reinforcement and discriminative stimuli exert precise control. Reinforced variability imparts functionality in many situations, such as when individuals learn new responses, attempt to solve problems, or engage in creative work. Perhaps most importantly, reinforced variability helps to explain the voluntary nature of all operant behaviors.

Keywords: Operant Variability; Voluntary; Determinism; Random; Choice

B. F. Skinner (1938) identified orderly relationships between environmental events and operant responses by defining the responses in terms of their outcomes rather than their individual characteristics. For example, when a rat presses a lever, contacts are closed in a microswitch, resulting in a food pellet, and a pigeon's key-pecks produce grain, and so on, and operant analyses ignore whether the rat pressed with the left paw, right paw, or snout, or whether the pigeon pecked the key from the left side or the right. Skinner and most other operant psychologists analyze behaviors at the level of outcome-defined generic classes that consist of families of responses. The individual responses in a class may or may not resemble one another, but all produce the same reinforcing outcome. Research over the past 80 years ably documents many orderly relationships between such classes and environmental variables.

But behavioral analyses can proceed at different levels, including the individual responses that produce common proximate effects and therefore comprise a class. In some cases, each individual meets criteria that can be described in terms of physical dimensions of the response. Thus, for example, the movement of the rat's paw must occur at a particular location, on top of the lever, and be in a particular direction, generally downward, and with a force greater than some minimal value, all in order to activate a microswitch. In many operant-conditioning experiments, microswitch closures (the proximal outcome) are studied as a function of environmental events, such as reinforcers and discriminative stimuli. In other cases, response dimensions may differ widely but all have a common proximal effect, such as in expressing our understanding of another person (as indicated by a glance or verbal reply) when you point to the wine bottle or ask to pass the wine or look sadly at your empty wine glass.

An expanded view of operant behavior is obtained when we study the individual instances that comprise operant classes: how they become members of a class, what causes emission of one or another instance, how instances are organized, and how they are generated (e.g., Pear, 1985). One way to do this is to focus on observed variability (or predictability) of within-class instances. We will show that variability is a reinforceable dimension of behavior – variability is itself an operant – and that environmental conditions re-

sult in behaviors that are predictably unpredictable. ("Predictably unpredictable" may seem to be an oxymoron but please read on.) We begin with some definitions and then turn to examples of variability that have been selected across evolutionary time (i.e., variability that is elicited or induced by environmental stimuli and contexts).

Definitions. Variability is an attribute of a set of instances: In this paper, the instances include responses, response sequences, and other dimensions of a response (such as response rates). Variability often implies noise, high dispersion, or unpredictability but the term is also used to refer to a continuum, from repetitive or predictable to stochastic or random. Context will indicate the intended meaning.

The terms *stochastic* and *random* are often used interchangeably, but we will use random to refer to cases of maximum unpredictability, where alternatives are uniformly distributed or equiprobable and predictions of individual responses cannot be better than chance. An intuitive sense can be gained if you imagine an urn filled with 1000 colored balls, 500 red and 500 green. The urn is well shaken and one ball is blindly selected. After selection, the ball's color is noted, returned to the urn, and the selection process repeated. Prediction of each ball's color will be no better than chance (in this case .5). This process of selecting balls provides a model of a random process and the outcome represents a random sequence.

Stochastic will be used as the more general term to apply as well to unequal or biased sets, for example if the urn were filled with 800 red and 200 green balls (see Nickerson, 2002, for discussion). Prediction accuracy could then rise to .8 (if one always predicted red). However, the process and output are described as stochastic because conditional probabilities (e.g., red given red, green given green, etc.) provide no more information than the first-order probabilities of .8 and .2. Stated differently, the selection of a green ball imparts no information as to when the next green might be obtained. Another example is seen when a responder prefers one of two responses, such as the left versus right alternative, but where conditional probabilities impart no more information than the baseline distributions. A third case occurs when a responder alternates from red to green or green to red more frequently than if responses were randomly generated, such as when switching is likely whenever a run of one of the colors is greater than three (i.e., three reds in a row or three greens), but responses are otherwise no more predictable than if based on first-order probabilities. Statistical analyses (e.g., the U-value statistic to be described below and other measures of entropy) provide indices of the level of unpredictability.

Induced Variability: Three Examples

Variability is sometimes induced or elicited by environmental events. The inducing stimuli and forms of variation often differ across species but are typical within the species. Induced variability is not learned under the selective influences of reinforcement contingencies, but, as will be seen, often interacts with reinforced variability.

Kineses. A simple example is seen in *E. coli* bacteria. They exhibit two types of movement: straight-line swimming and random tumbling (Macnab & Koshland, 1972). When the food gradient improves across time, a bacterium swims straight ahead. When the food concentration decreases, tumbling becomes more probable. Tumbling results in movements in random directions with the combination of straight-line and random movements resulting in a kind of hill climbing in the direction of nutritive substances. This simple example shows stimulus-controlled induction of two levels of variability: unpredictable responding (tumbling) and completely predictable straight-ahead movement.

Protean behaviors. A strange observation befuddled researchers for many years. If keys were jangled in a laboratory room that contained caged rats, some of them would run around and jump frenetically. Why might noise produce what came to be referred to as "audiogenic seizures?" Chance (1957) noticed that if the chamber contained a small box, the rats would hide there whenever the keys jangled and would not run wildly. So began research on protean behavior, named after Proteus the Greek god, who could change his shape unpredictably so as to elude pursuers. Driver and Humphries (1970) document protean behaviors engaged in by many different species. Among the functions served by protean behaviors, survival is primary: a prey animal is more likely to survive if it responds unpredictably when confronted by a potential predator. For example, insects, fish, birds, and small mammals will move in highly erratic fashion in the presence of a predator. The hare will zigzag left or right, or move straight ahead, mixing the three unpredictably when chased by a fox. Across species, variable behaviors include unpredictable changes in direction, speed, form and type or topography of movement (Driver & Humphries, 1988).

There are other reasons to befuddle other animals. One example is the so-called 'crazy dance' seen when a weasel attempts to capture a vole. Having spied a vole, the weasel may jump about this potential meal, roll on the ground, twirl in a circle, do somersaults – all while moving around the vole – until finally it pounces on the motionless (some might say, astounded) vole. Australian aborigines do similarly crazy-seeming displays when hunting kangaroos.

The evolutionary bases of protean responses are indicated by the commonality of responses within a species as well as

the more general tendency to respond unpredictably across many different species. One researcher writes, "Along with directional fleeing, protean escape behaviors are probably the most widespread and successful of all behavioral antipredator tactics, being used by virtually all mobile animals on land, under water, and in the air" (G. F. Miller, 1997, p. 319). The controlled or selected nature of protean behavior is indicated by the fact that when a predator is distant, the potential prey may simply run away – strategies differ when the predator is far vs. near – demonstrating stimulus control. Driver and Humphries (1988, p. 157) write that protean unpredictability is "not so random as to be formless; it is a structured system within which predictability is reduced to a minimum." This point parallels one that will be emphasized throughout the present paper: Phylogenetic selection pressures and ontogenetic reinforcers establish sets of *functional responses* from which instances emerge *stochastically*. Orderliness and predictability are provided by the functionality of the responses and unpredictability by their stochastic emission.

Bird song. "Variations attract" characterizes mating preferences in some songbird species (Catchpole & Slater, 1995). Female mockingbirds prefer males who sing complex songs; female sparrows display sexually more in the presence of variable songs than repetitive ones; and female great tits demonstrate sexual interest in males with the largest song repertoires. This implies that birds can discriminate among different levels of variability and experimental analyses with pigeons confirm this conjecture (Young & Wasserman, 2001). Also implied is that male song variability is influenced by environmental contexts. In support, Searcy and Yasukawa (1990) observed that when male red-winged blackbirds were presented with a female dummy, song variability increased. In some species, such as zebra finches, variable songs are generated by males in the absence of females but once a female is attracted, the male's songs become female-directed and more stereotyped (Sakata, Hampton, and Brainard, 2008). Whether males increase or decrease song variability when females are present or expected, birdsong variability is an evolved characteristic, with levels controlled by environmental contexts.

Genetic Variability

The above section showed that evolved phenotypic variability is related in orderly ways to environmental events; variability is selected and constrained. Similar effects are seen at the level of genes. Changes in DNA molecules have many causes including errors during replication, mutations caused by chemicals or radiation, jumps or transpositions of genetic materials early in the developing fetus and in adult brain (transposons), and other spontaneous changes. Lewis Thomas highlighted the importance of genetic variability:

"The capacity to blunder slightly is the real marvel of DNA. Without this special attribute, we would still be anaerobic bacteria and there would be no music" (quoted in Pennisi, 1998, p.1131). There are additional important contributors to individual variations in all sexually reproducing organisms: variability during gamete formation. High levels of constrained variations are produced when genetic material in sperm and egg cells divide: there is random and independent assortment within individual chromosomes and random crossings between portions of maternal and paternal chromosomes. Mutations, jumps, assortments, and crossings occur stochastically, without regard to the current needs of an organism. However, the processes that permit and maintain genetic variability have themselves evolved under selection pressures. "(T)he genome…(has an) ability to create, focus, tune and regulate genetic variation and thus to play a role in its own evolution" (Caporale, 1999, pp. 15). A combination of variation and selection at work within the genome may best be described as *selected (or bounded) stochasticity*, with mutations, mixings and variations occurring stochastically and unpredictably, but within a confined milieu that has been selected and conserved over evolutionary time. As will be seen, operant response variability is similarly selected, but this process is instead driven ontogenetically by experiences with reinforcing feedback.

Operant Variability: Overview

Behavioral variability is often assumed to have one of three causes: unrelated events within the environment or organism, induction from such things as aversive events or extinction, or unexplained variance. Not only in behavioral psychology but also in most sub-fields of psychology, variability is treated as a nuisance, something to be minimized because it obscures relationships. There is a fourth contributor, at least as important as any of the other three, and one that leads to a revision of our views of operant behavior and its voluntary nature. To state the case simply: Variable responding is produced and maintained by reinforcers contingent upon it. Variability does not always decrease with learning, this being counter to initial theories of reinforcement. Of most importance, *particular levels of variability* are engendered by reinforcers contingent upon those levels. Variability is a dimension of behavior analogous to other *operant dimensions*, such as response rate, force, and topography. Support for these claims will be outlined below, but we begin with an overview of the methods and analyses used to document reinforcement of variability.

Operant Variability: Basic Procedures

Methods

In most of the experiments to be described, two alterna-

tive responses are possible and a sequence consists of a fixed number of responses per trial, with the possible patterns constituting the response class. For example, if a trial consists of 4 responses on Left (L) and Right (R) operanda, then the operant class would comprise 16 instances (2^4), or LLLL, LLLR, LLRL, LLRR, and so on. If trial length were instead 8 Ls and Rs in length, then the operant class would contain 256 (2^8) possible instances. The main question asked in many of the studies is whether high levels of sequence variations can be generated and maintained by *reinforcers contingent upon the variability*. A number of different procedures have been employed and the most common will be described.

(i) Under *recency methods*, reinforcement is contingent upon a sequence that had not occurred across a given number of previous trials (Page & Neuringer, 1985). The lag procedure is a common example. Under lag 5, the current sequence will be reinforced only if it had not been emitted during any of the previous 5 trials. In a variation of the lag procedure, Machado (1989) kept track of the number of intervening trials before a given sequence was repeated, defining this as the "recurrence time," and combined it with a percentile reinforcement contingency to generate high levels of variability (see also Machado, 1992). Percentile reinforcement contingencies base the criterion for reinforcement on the subject's own performance over a previous set of responses (see Galbicka, 1994). Another variant is the novel response procedure in which a response is reinforced upon its first observed occurrence (Pryor, Haag, & O'Reilly, 1969) or first occurrence within a given session (Goetz & Baer, 1973). Similarly, radial-arm maze procedures reinforce only initial (within a given session) entries into arms of a maze (Olton & Samuelson, 1976).

(ii) *Frequency* or *threshold procedures* reinforce responses that have occurred with relatively low frequencies. Denney and Neuringer (1998) provide an example in which trials consisted of four responses by rats on L and R levers. A running tally was kept of the frequencies of each of the 16 possible sequences. If the relative frequency of the current sequence – the number of its occurrences divided by the total occurrences of all 16 sequences – was less than a specified threshold value, in this case, .05, the rat was rewarded. Recently emitted sequences contributed more to the maintained tally than non-recent because after each reinforcement, all 16 counters were multiplied by a weighting coefficient equal to 0.95. Therefore the contributions of particular trials to the running tally counters decreased exponentially with successive reinforcements. One variant is the *least-frequent* response procedure (Blough, 1966; Schoenfeld, Harris, & Farmer, 1966; Shimp, 1967) that reinforces only the single response or sequence that is currently least frequent. Another variant is the *frequency dependence* procedure (Machado,

1992; 1993) in which the probability of reinforcement is a continuous function of relative response frequency, the more frequent a response, the less likely it is to be reinforced.

(iii) *Statistical evaluation* procedures compare a subject's performance to that of a stochastic model. Neuringer (1986) provided human participants with feedback from 10 statistical tests of randomness. In variations of this method, Platt and Glimcher (1999) and Lee, Conroy, McGreevy, and Barraclough (2004) performed on-line statistical analyses of monkey choices and reinforced only those choices that were not predicted by the computer's statistical analyses.

Evidence from each of the methods just discussed supports the hypothesis that variability can be reinforced (Neuringer, 2002).

Measures

Among the many measures of variability and randomness (Knuth, 1969), *U-value* is commonly employed (Machado, 1989; Page & Neuringer, 1985; Stokes, 1995). U-value is based on the distribution of relative frequencies, or probabilities of a set of responses. For a set of 16 possible responses, U-value is given by the following equation:

$$U = \frac{-\sum_{n=1}^{i=16} P_i \cdot \log_2(P_i)}{\log_2(16)}$$

[Eq. 1]

Here p_i represents the probability (or relative frequency) of a response sequence i. U-values approach 1.0 when relative frequencies approach equality, as would be expected over the long run from a random process, and 0.0 when a single instance is repeated.

U-value is closely related to Shannon's measure of information entropy, typically denoted by the symbol H (Shannon, 1948), which in the above example takes the following form:

$$H = -\sum P_i \cdot \log_2(P_i)$$

[Eq. 2]

Unlike U-value, Shannon information has no upper limit. Thus, for example, in the case of L and R alternatives and 4-response trials, if each of the 16 possible sequences is emitted equally often, H would equal 4; but if there were three possible responses, L, R, and C (for center), again with 4-response trials, 81 different sequences would be possible, and equality of emission would yield an H value of 6.34. The advantage of using U-value instead of H is that U provides a common scale where equality of responding will

yield a U-value of 1.0, independent of number of possible responses or sequences.

Other measures include percentages or frequencies of trials in which variability contingencies are met (Page & Neuringer, 1985); percentages of alternations vs. stays (Machado, 1992); conditional probabilities of responses (Machado, 1992); frequency of novel responses (Goetz & Baer, 1973; Schwartz, 1982); frequency of different responses in a session (Machado, 1997; Schwartz, 1982); Markov analyses (Machado, 1994); and a variety of statistical tests used to assess the randomness of a finite sequence of outputs (Neuringer, 1986). Research employing all of these measures converges on the conclusion that variability can be reinforced.

### Operant Variability: Experimental Evidence

Blough (1966) performed one of the first studies to show that highly variable responses can be reinforced, in this case interresponse times (IRTs). Blough was attempting to design an alternative to variable-interval and variable-ratio reinforcement schedules as a baseline to measure effects of other variables. He proposed that reinforcement of randomly occurring responses might provide a statistically stable and reproducible baseline. Pigeons were rewarded if the time since their previous peck, that is, if the current IRT, had occurred least frequently over the recent past. To see what this implies, imagine that responses were occurring randomly with a .5 probability during each second since the previous response, resulting in an exponential distribution of IRTs[1], much like the random emissions of an atomic emitter. Each response resets the IRT timer, and therefore response probabilities would be .5 during the first second, .5 during the next second (assuming that a responses had not occurred during the first second), and so on. In a set of 1000 responses, approximately 500 would occur in the 0 to 1-sec IRT bin, 250 in the 1-2 s IRT bin, 125 in the 2-3 s bin, and so on. Blough created 16 IRT bins, adjusting their sizes so that a random generator would produce equal numbers in the counters associated with each bin. Because, as just described, a random responder generates many more short IRTs than long ones, thereby resulting in an exponential distribution, bin size was small at the short end and large at the long, increasing systematically across the IRT range. To be reinforced, a pigeon's IRT had to fall in a bin that contained the lowest number of prior entries, compared to all of the other bins, across a moving window of 150 responses. Other aspects of the procedure increased the likelihood of exponentially distributed IRT frequencies and also controlled rates of reinforcement. The result was that the birds learned to approximate the exponential distribution, but with some biases, i.e., they responded stochastically. Very short IRTs (<0.5 s) were more frequent than if responses were randomly

generated and there was a tendency for long IRTs to follow long, and short to follow short, more than expected from a random emitter, but these might have been due to aspects of the particular reinforcement contingencies. Despite these problems, the pigeons' distributions of intervals approximated the exponential distribution expected from an atomic emitter. This was the first clear experimental demonstration that highly unpredictable responding could be reinforced. Blough's study also showed how biases (in this case for short IRTs) could affect stochastic emission, another finding that was supported by research in the years to follow.

Page and Neuringer (1985) provided additional evidence and important control conditions. Variability of response *sequences* was the measure of interest. Pigeons pecked L and R keys, 8 responses per trial. In one sub-experiment, a trial ended with food if the sequence in that trial differed from the sequences in each of the previous five trials, a Lag 5 schedule. If the current sequence was the same as any one (or more) of the previous 5, then a brief timeout (chamber dark and keys inactive) followed. Although this criterion could be met in many different ways, the birds generated stochastic sequences, e.g., as measured by the U-value statistic. Furthermore, as is the case for other operant dimensions, variability was shown to be sensitive to schedule parameters. As lag values increased from 1 to 25, variability increased. Other studies have confirmed the control by reinforcement contingencies over levels of variability in rats (Grunow & Neuringer, 2002), pigeons (Blough, 1966; Neuringer, 1992) and people (Jensen, Miller, & Neuringer, 2006; Maes, 2003), and with responses as diverse as lever presses, eye-movement saccades (Madelain, Chaprenaut, & Chauvin, 2007), vocalizations by birds (Manabe, Staddon, & Cleaveland, 1997) and walruses (Schusterman & Reichmuth, 2008), and instances of categories generated by human participants (Neuringer & Jensen, 2010).

But a key question remained: Was the *contingency* between reinforcers and variability responsible for the variable responding? Alternative hypotheses had to be considered. For example, when variability is reinforced, absence of variability results in the withholding of reinforcement. But we know that low reinforcement frequencies *induce* variability (see below). Thus, the variability observed may have been caused by decreased reinforcement or brief periods of extinction. To test this possibility, Page and Neuringer compared two conditions. The first was a lag 50 where a sequence was required to differ from each of the previous 50 sequences (this condition referred to as Var). As in most other experiments from our laboratory, variability was assessed continuously across sessions such that a sequence at the beginning of one session had to differ from each of the terminal 50 sequences in the previous session, and so on. High levels of variability were generated. In a second condition (referred

to as Yoke), each pigeon experienced exactly the same intermittency of reinforcers as in Var but under a self-yoked procedure. If, for a particular bird, the 1st, 5th, and 8th trials in a Var session produced reinforcers, then the 1st, 5th, and 8th trials would be reinforced under the yoke condition but *independently of whether sequences met the lag contingency or not*. Thus, under yoke, the pigeons had to continue to emit 8 responses to complete each trial, and reinforcement was identical to Var, but variability was not required. The result was a marked decrease in sequence variability in the Yoke condition and an increase in sequence repetitions. Return to lag 50 resulted in variability again increasing and a return to Yoke again caused a decrease (Figure 1). This result, which has been replicated in experiments in many different laboratories (see Neuringer, 2002, for review), demonstrates that the reinforcement *contingency* is responsible for the high levels of variability, or, in other words, that variability is an operant dimension of behavior.
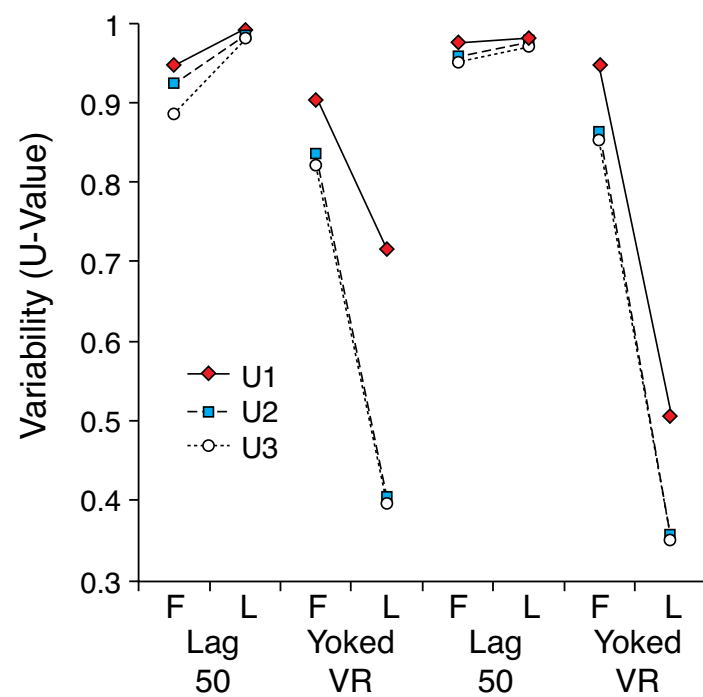

Figure 1. Variability of pigeon responses under Lag 50 conditions (where reinforcers depend upon the current sequence differing from each of the last 50 sequences) and yoked variable ratio (Yoked-VR, where reinforcers are provided independently of sequence variability). Three measures of response uncertainty (or entropy) are shown: U1 = Responses evaluated one at a time; U2 = Responses evaluated in pairs; U3 = Responses evaluated in triplets. F = Averages over the first 5 sessions and L = averages over the final 5 sessions of each condition. (Adapted with permission from Page, S. & Neuringer, A. (1985). Variability is an operant. Journal of Experimental Psychology: Animal Behavior Processes, 11, 429-452.)

Reinforcers control more than whether responses vary; they also define the set or class from which variations emerge. The result is controlled, selected, or bounded variability. This controlled nature of operant variability is shown in a number of ways. For example, when many operanda are present in an experimental chamber, if reinforcement is based on variations among a subset, then responses are generally confined to the reinforced subset (Neuringer, Kornell, & Olufs, 2001). Thus, rats learn *what* to vary as well as *how much* to vary. Similarly, when 4-response sequences (across L and R levers) constituted a trial under lag contingencies, but only trials that began with RR were reinforced, the emitted sequences, while varying, were generally limited to the reinforced set (Mook, Jeffrey, & Neuringer, 1993).

Extraordinary evidence for the controlled nature of operant variability was seen when the reinforcement contingencies required variability along two dimensions of a response while, simultaneously, repetitions along a third dimension (Ross & Neuringer, 2002). Human participants were instructed to draw rectangles on the screen of a computer so as to earn points. Those were the only instructions. Participants in the first of three groups were rewarded for rectangles whose screen locations (indicated by the centroids) varied as did forms (square, or rectangles that were long in the horizontal or vertical direction, etc.) while sizes were approximately the same, trial after trial. Participants in a second group were reinforced for repeating location while varying size and form. A third group was reinforced for repeating form while varying size and location. Each group learned to respond appropriately – to vary *and* repeat, as required by the contingencies (Figure 2). Many of the participants, while realizing that points depended on their drawing of rectangles, could not identify the underlying criteria. This provides a striking example of the power of binary feedback (reinforce or not) to control variations and repetitions along multiple dimensions of a response, and to do so concurrently and independently.

Discriminative stimuli provide additional evidence for the controlled nature of operant variability. For example, Page and Neuringer reinforced pigeons for varying 5-response sequences under a Lag 10 contingency in the presence of blue keylights and for repeating a single 5-response sequence, LRRLL, when the key color was red. The birds learned this discrimination and when the contingencies were reversed, so that now they had to vary in red and repeat in blue, their performances changed appropriately (see, also, Cohen, Neuringer, and Rhodes, 1990, for similar results with rats).

Additionally, Denney and Neuringer (1998) showed that comparison with a fixed, repeated sequence was not necessary to demonstrate stimulus control. Rats were required to vary under one stimulus, Var, but variability was not re-
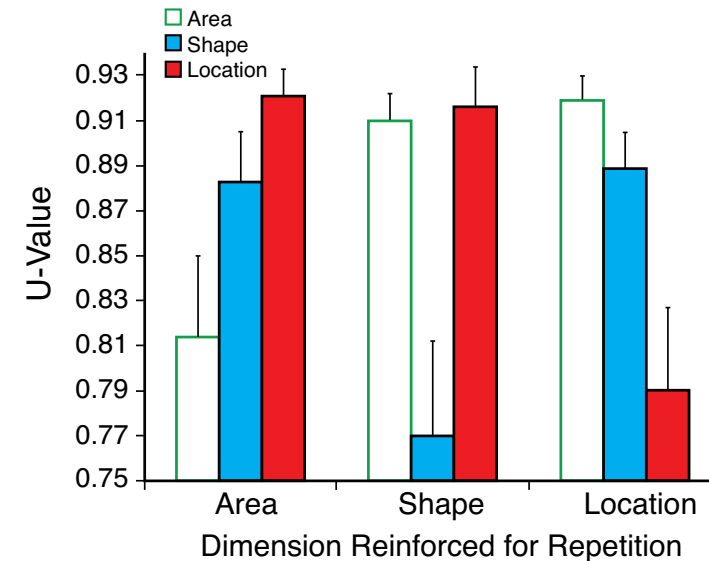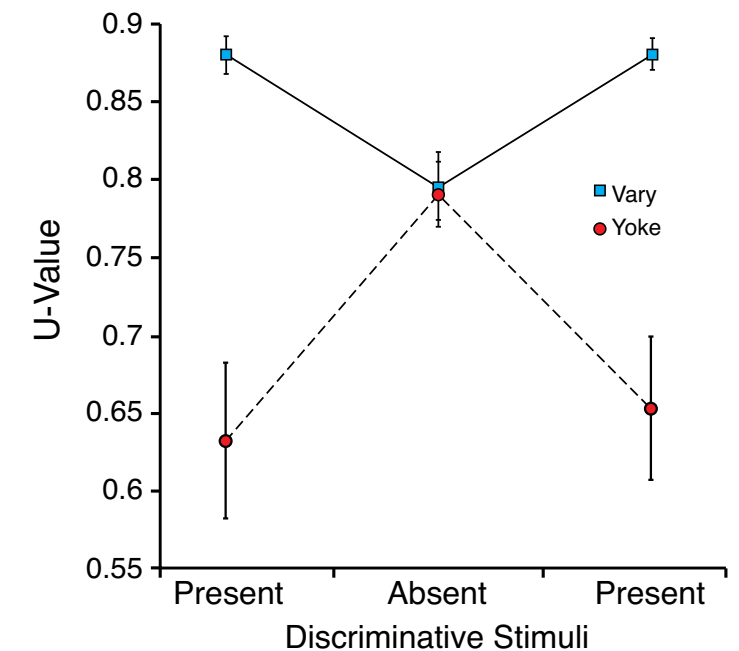
Figure 2. Variability, given by U-value, for each of three dimensions of rectangles – area, shape, and location – drawn by human participants on the screen of a computer. The x-axis indicates three separate groups: Reinforced for repeating rectangle areas (left set of bars), or repeating shapes (middle set of bars), or locations on the screen (right set of bars). For each group, variations were required for the other two dimensions. Error bars indicate standard errors. (Adapted with permission from Ross, C. & Neuringer, A. (2002). Reinforcement of variations and repetitions along three independent response dimensions. Behavioural Processes, 57, 199-209.)

quired under a second stimulus, a Yoke condition such that reinforcement was approximately equivalent in the two stimuli. Figure 3 shows the results. The outer points, left and right, show the large differences in variability when discriminative stimuli were present, with the squares representing responding during the Var stimulus and the circles showing variability during Yoke. U-values were high when variability was required and much lower when variability was simply permitted (but not required). The center points show levels of variability when the discriminative stimuli were removed. Now, absent any cues as to when to vary or not, the rats intermixed high levels of variability with low, and did so throughout the session. Ward, Kynaston, Bailey, and Odum (2008) showed similar control by vary and yoke stimuli in pigeons. These studies demonstrate conclusively that levels of operant variability are controlled by discriminative cues.

Induced Variability of Operant Behaviors

Earlier in the paper, we described species-typical variability that is induced by particular stimuli, such as a predator. Variability is also induced by presenting and withholding reinforcers. Terminology can be confusing and we use "induced operant variability" to refer to effects on operant (and

Figure 3. Response variability, given by U-value, for rats in a discrimination-learning task. The data are averages across 20 rats during three single sessions, two in which discriminative stimuli were present, and an intervening session in which discriminative cues were absent. When the Vary stimulus was present (squares and solid lines), only infrequently emitted response sequences were reinforced. The same frequency of reinforcement was provided during the Yoke stimulus (circles and dashed lines) but independent of whether sequences varied. The Vary and Yoke contingencies were continued during the "Absent" phase, but the discriminative cues were no longer present. (Adapted with permission from Denney, J. & Neuringer, A. (1998). Behavioral variability is controlled by discriminative stimuli. Animal Learning & Behavior, 26, 154-162.)

reinforced) responses that are *independent of the contingency between responses and reinforcers*. As will be seen, control conditions are necessary to separate induced effects from those due to the contingency. Insofar as variability emerges even when reinforcement is not directly contingent upon it, a purely contingency-based account is insufficient to explain variability's role in operant behavior. No less importantly, induction procedures are often used to generate the variability from which new responses can be reinforced, in both therapeutic and learning contexts to be described below.

Extinction-induced variability. After a period of reinforced responding, withholding of the reinforcers (extinction) leads to increased variability. This variability is not learned since it occurs upon the first extinction experience and it is not reinforced. Examples of extinction-induced variability include variability of response location (Anto-

nitis, 1951; Eckerman & Lanson, 1969), response force (Notterman & Mintz, 1965), topography (Stokes, 1995), and number (Mechner, 1958). As with other examples of induced variability (such as protean behaviors), extinction-induced variations are bounded and orderly: they are primarily selected from the same, or similar response class as was established during original learning. For example, if lever presses produced food pellets, a rat may vary the ways in which it presses when food is withheld with much of the behavior directed towards the lever (e.g., Stokes, 1995).

This constrained nature of extinction-induced variability was shown by Neuringer, Kornell, and Olufs (2001). Rats were reinforced for a particular 3-response sequence across 3 operanda, namely (L)eft lever ➔ response (K)ey ➔ (R)ight lever (LKR), in that order. After the rats had learned the sequence, reinforcement was withheld. The top panel of Figure 4 shows the relative frequencies of each possible 3-response sequence (proportions of occurrences) during the reinforcement phase (filled circles) and during extinction (open circles). LKR was most frequent when it was reinforced, as expected. During extinction, LKR continued to be the most frequently emitted sequence. (Note that these graphs show *relative* frequencies. Absolute rates of response were much lower during extinction than during the reinforcement phase.) Shown at the bottom of the figure are the ratios of response proportions during the reinforcement and extinction phases (that is, the ratio of the two curves in the upper graph). Thus, while the same ordering of sequence proportions was maintained during extinction as during conditioning, variability increased in extinction due to the unusual or highly unlikely sequences (for related findings, see Bouton, 1994; Pear, 1985). Neuringer, Kornell, and Olufs obtained similar results from a second group of rats that had been reinforced for sequence variations: The ordering of sequence probabilities was maintained while low probability sequences became (slightly) more frequent. Souza, Abreu-Rodrigues, and Baumann (2010) obtained the same results with human participants. We conclude that extinction results in a "… combination of generally doing what worked before but occasionally doing something very different… (This) may maximize the possibility of reinforcement from a previously bountiful source while providing necessary variations for new learning" (Neuringer et al., 2001, p. 79).

Variability induced by distance from reinforcement. Responding becomes increasingly repetitive and predictable as a reinforcer is approached in time, space or effort. This was shown for sequence variability (Cherot, Jones, & Neuringer, 1996), lever-press duration variability (Gharib, Gade, & Roberts, 2004), and movement variability (Akins, Domjan, & Gutierrez, 1994; see also Craig, 1918).
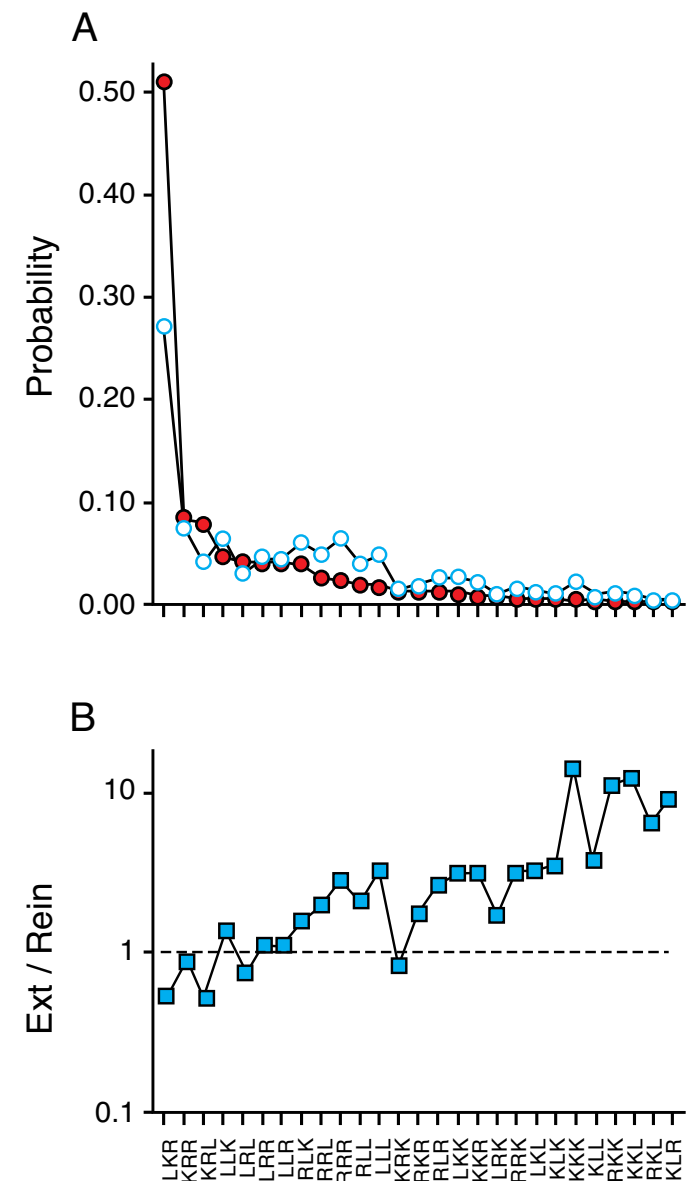
Figure 4. The top graph shows probability of each of 27 possible 3-response sequences when only the LKR sequence was reinforced (filled circles) and when extinction was imposed, i.e., no reinforcers were provided (open circles). The bottom graph shows the ratios of these same probability, in other words, extinction probabilities divided by reinforcement probabilities, on a logarithmic y-axis. (Adapted with permission from Neuringer, A., Kornell, N., & Olufs, M. (2001). Stability and variability in extinction. Journal of Experimental Psychology: Animal Behavior Processes, 27, 79-94.)

Variability induced by reinforcement frequencies. In general, response variability is high when reinforcers are infrequent, and lower when reinforcers are frequent (see Lee, Sturmey, & Fields, 2007, for a review). One interpretation is that low expectation (or anticipation) of reinforcers induces variability (Gharib et al., 2004), a description that can also be applied to the just-noted distance-inducing effects.

Functionality of induced operant variability. Early studies by Thorndike (1911) and Guthrie and Horton (1946) documented the emergence of new operants (referred to as instrumental responses) from a substrate of induced variable behaviors. For example, Thorndike observed that a cat scrambled about the cage, clawed at the wall and roof, but eventually a response succeeded in opening a door that provided access to food. Across trials, the time to escape from the cage shortened, and the form of the response became increasingly predictable. The conclusion was that learning of a response caused variability to decrease, a conclusion that has been extended to suggest that reinforcement necessarily narrows and constrains responses (Schwartz & Lacey, 1982). Among the many ways in which Skinner extended this early work was to describe the shaping of new responses. In shaping, the criterion for reinforcement changes across time or responses and, as in Thorndike's experiments, variability contributed importantly to the learning process. In each of these cases, response variability was induced – it occurred for reasons other than explicit reinforcement of the variability itself. An obvious question is whether direct reinforcement of variability would, in fact, facilitate acquisition of new operant responses, a question to which we return later.

Interactions

The same reinforcers that produce response variability (because of the contingency between variability and reinforcer) may interfere with it (because of induced effects of the reinforcer). As a related example, it is exceedingly difficult to reinforce 'standing still' in a hungry pigeon because the induced excitement and motivation of the food reinforcers work in opposition to standing still. There are many cases of induced or evolved influences interfering with operant behavior, and, of course, many cases where facilitation is observed. Interference and facilitation also occur with respect to the variability operant.

Cherot, Jones, and Neuringer (1996) showed that when one group of rats was reinforced for repetitions and another for variations, anticipation of reinforcement facilitated performance in the first case, and interfered with it in the second. The first group (Rep) was reinforced for repeating sequences of 4 responses across two levers and the second group (Var) for sequence variability. The novel aspect of the Cherot et al. experiment was that only every fourth correct sequence provided a food reinforcer, i.e., a Fixed Ratio 4 was superimposed on the Rep and Var contingencies, with the first three correct sequences producing only a conditioned stimulus. As expected, animals in the Var group responded much more variably overall than the Rep animals, indicating control by the variability/repetition contingencies (Figure 5, bottom). However, as each reinforcer delivery
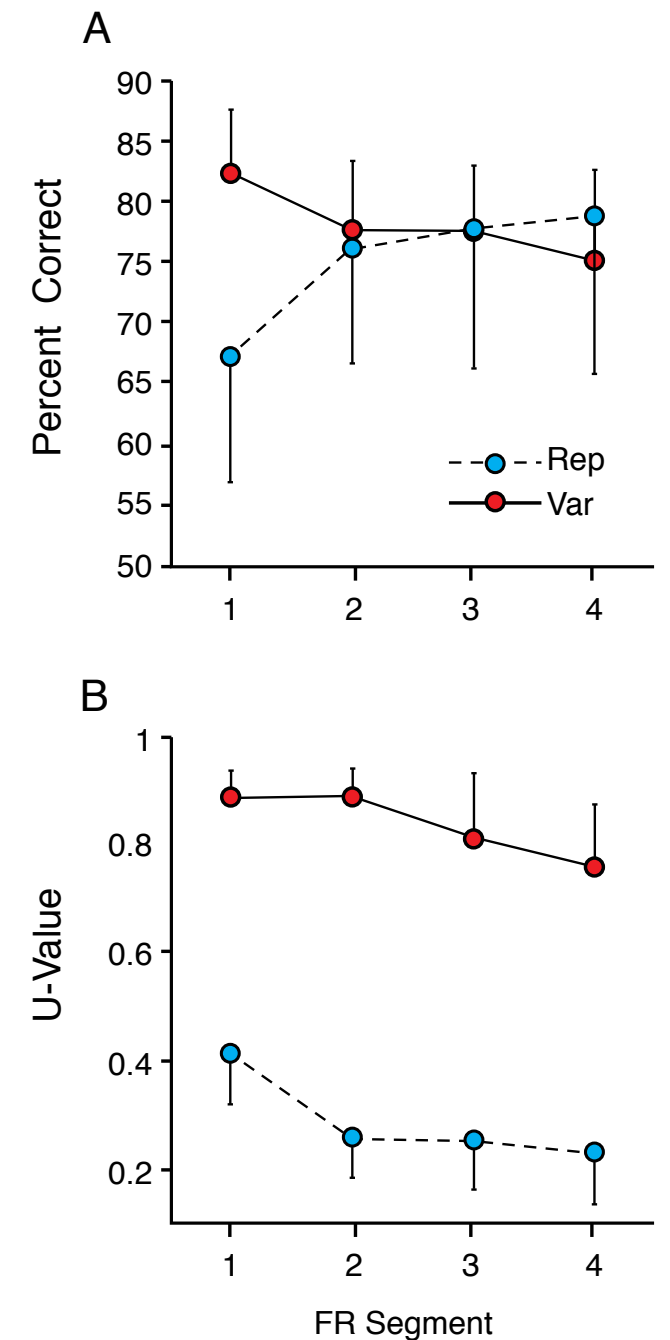

Figure 5. (Top)The percentage of sequences that met a variability-reinforcement contingency (red circles and solid lines represent 7 rats in the Var group) or repetition-reinforcement contingency (blue circles and dashed lines represent 7 rats in the Rep group) as a function of the location within a fixed-ratio (FR 4) schedule where four correct sequences were required for each food reinforcer. Lines connect group means, and error bars show group standard deviations. (Bottom) U values, a measure of sequence variability, as a function of location within the fixed ratio. (Adapted with permission from Cherot, C., Jones, A., & Neuringer, A. (1996) Reinforced variability decreases with approach to reinforcers. Journal of Experimental Psychology: Animal Behavior Processes, 22, 497-508.)

was approached (i.e., as the last of 4 successful sequences was neared) levels of variability *decreased* for both Var and Rep groups. Recall the expectancy-of-reinforcement effects described above. In this case as well, variability decreased as reinforcers were approached, thereby *facilitating* correct responding in the Rep group but *interfering* with it in the Var (Figure 5, top). Similar interactions between contingency and induction may help to explain other findings, including those related to creative behaviors. For example, the commonly reported detrimental effects of anticipated rewards on creative activities may in part be due to the proximity effects just described (Amabile, 1983; Neuringer, 2003). But overall levels of variability – and perhaps the creativity of the activity as well – are higher when reinforced than when not. Thus, rather than concluding that reinforcement is generally detrimental to creativity, it will be more helpful to identify reinforcing and inducing effects.

Grunow and Neuringer (2002) showed that the magnitude and direction of induced variability depend upon *levels of reinforced variability*. Four groups of rats were reinforced for different levels of variability, from high to relatively low, across three operanda, two levers and a key. The leftmost points in Figure 6 (CRF or continuous reinforcement) show that the variability contingencies exerted strong control: The group reinforced for high variability responded most variably, the group reinforced for low variability responded with low levels, and so on for the intermediate groups. Overall frequencies of reinforcement were then systematically lowered by superimposing a Variable Interval (VI) schedule-of-reinforcement requirement atop the variability contingency. Thus, in one phase, reinforcers (for varying) were available only on the average of once per minute (VI 1 min), and in another phase, once every 5 min (VI 5 min). Only after the VI interval elapsed would meeting the variability contingency be reinforced, with all other trials leading to brief timeouts. As outlined above, we know that lowering reinforcement rate often increases response variability. Grunow and Neuringer asked whether the same would be found across the different levels of operantly reinforced variability. The results showed that the size and direction of the induction effects depended on the variability contingencies. When low variability was reinforced, decreases in reinforcement resulted in increased variation, consistent with most previous reports. However, when high variability was reinforced, lowering of reinforcements had the opposite effect: response variability decreased as reinforcements became increasingly rare. The intermediate groups showed intermediate effects. This is another case in which response variability depends on a combination of variability-contingent (or reinforced) and induced (rate of reinforcement) influences. Such interactions may help to explain additional observations from outside of the lab. Many work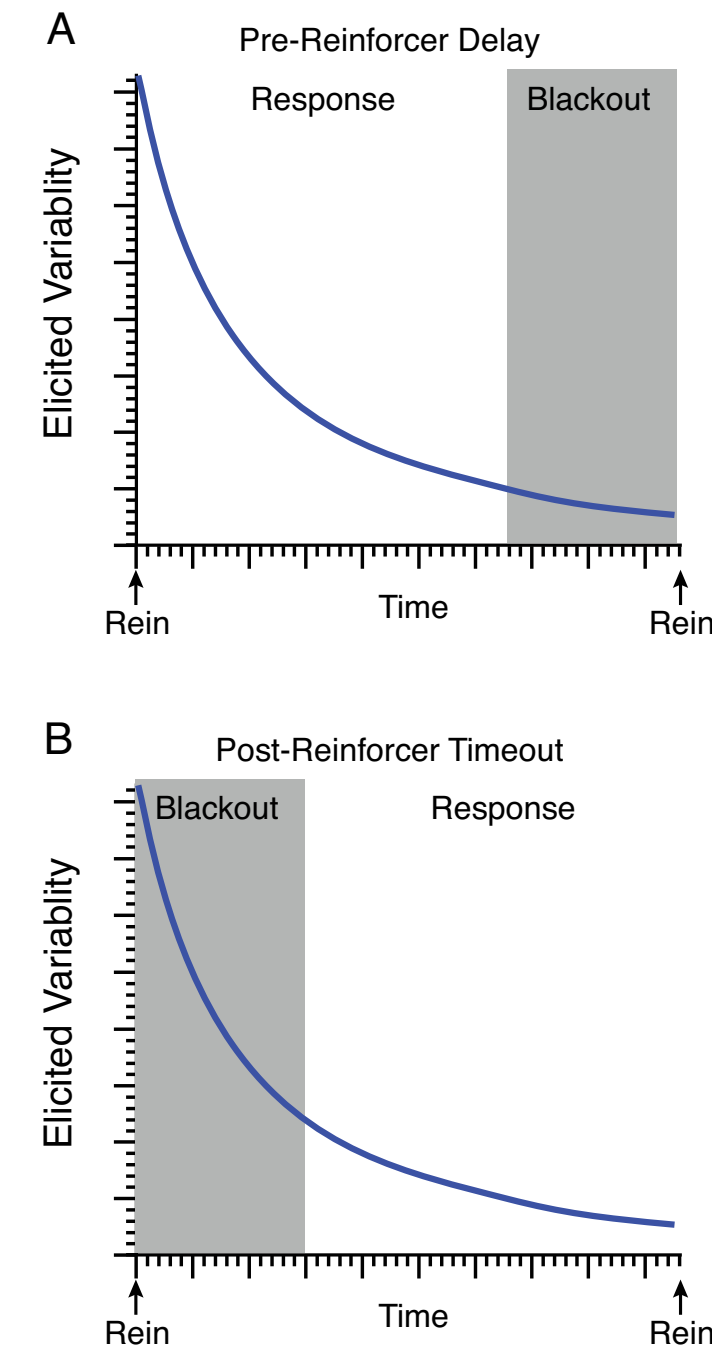ers engage in repetitive behaviors, e.g., factory workers, mail carriers, and fare collectors; but variable behaviors are the norm for others, e.g., inventors, fashion designers, and artists. Lowering pay or withholding positive feedback may affect behaviors differently in these two cases.
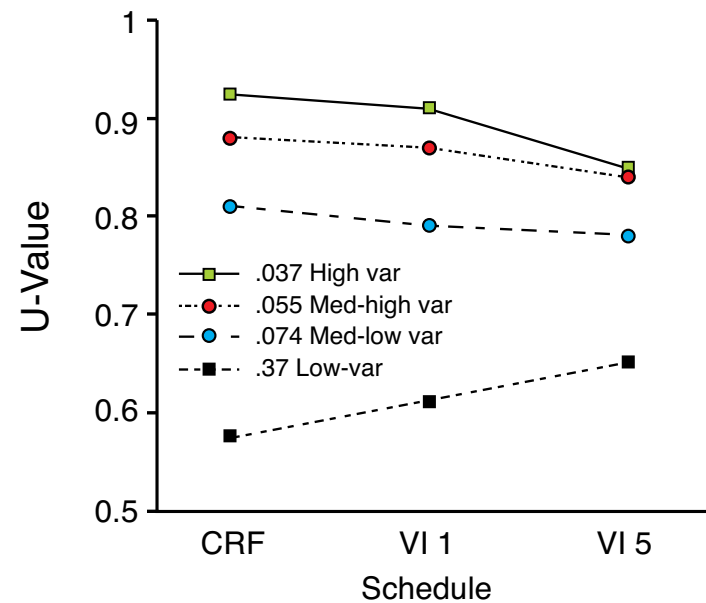
Figure 6. Average U values for four groups of rats (10 each) that differ in terms of the levels of variability required for reinforcement: 0.037 = very high variability required; 0.37 = very low variability required; and the other two groups, 0.055 and 0.074 = intermediate levels required. The x-axis shows 3 phases: CRF = reinforcement every time variability contingency met; VI 1 = reinforcement for meeting respective variability contingencies no more than once per minute, on average; VI 5 = reinforcement no more than once every 5 min. (Adapted with permission from Grunow, A. & Neuringer, A. (2002). Learning to vary and varying to learn. Psychonomic Bulletin & Review, 9, 250-258.)

Delays of reinforcement (periods imposed between responses and reinforcers) also induce changes in responding that depend upon reinforced levels of variation. Wagner and Neuringer (2006) reinforced different groups of rats for low, medium, and high response-sequence variability with a trial consisting of three responses across four active operanda – two levers and two keys. The authors asked two questions: do levels of reinforced variability influence the effects of reinforcement delays; and do delays *prior to reinforcers* have different effects on variability than the same periods *following reinforcers* (post-reinforcement timeouts). The main results were that both levels of reinforced variability and the location of the delays/timeouts influenced the induced effects. Delays increased variability when low variability was reinforced and decreased it when reinforcers depended upon high variability. At low variability levels, delays

and timeouts had opposite effects: pre-reinforcement delays *increased* variability and post-reinforcement timeouts *decreased* it (see, however, Odum, Ward, Barnes, & Burke, 2006). Figure 7 suggests a reason why. Delays remove the opportunity to respond in the interval leading up to the reinforcer, when repetitions are most likely to be induced (as shown by Cherot et al., 1996). Post-reinforcement timeouts do the opposite by removing the opportunity for responding early during the next period, when induced variability is high. Thus knowledge of the subtle interactions between the inducing and strengthening effects of reinforcement helps to explain response variability. There may be important lessons here for application: when variability is desirable, as when shaping a new response, or reinforcing creativity or problem solving, then imposition of pauses, rest periods or timeouts *following* reinforcers may increase helpful variations. On the other hand, pauses should not *precede* reinforcers that are contingent upon variability or upon successive approximations to the desired response. In brief, both contingent and induced effects of reinforcers must be considered when attempting to influence levels of response variability during the shaping process.

Cognitive Variability

Operant response classes have much in common with cognitive categories (Murphy, 2002). Categories contain multiple instances, and these instances often demonstrate a hierarchical ordering, from high to low probability (Rosch, 1978). Apples are a more likely response to "name a fruit" than kumquats. However, the ordering of within-category instances – or in operant terms, the probabilities of response – differ, both within individuals at different times or under different circumstances, and across similar individuals at the same time in identical circumstances (Barsalou, 1987). Furthermore, distributions of within-category probabilities change with environmental demands, much as is the case for within-class operant variability. To state this differently, under some circumstances, normally low probability responses may be emitted with high probability and vice-versa. This was shown when college students were asked to generate instances of a verbal category that were highly likely to be given by other individuals (e.g., when asked simply to name a fruit), unlikely to be given, and levels in between. Participants could readily generate low-to-high probability instances of both common categories (e.g., animals and fruits) and ad-hoc categories (e.g.,things to eat on a diet, something to do during a lecture, and things that might fall on one's head) (Neuringer & Jensen, 2010). More than any other operant domain, language demonstrates an ability of users to vary the predictability of instances. Linguistic variability, while extraordinary in its range, may be established and controlled in the same ways as other types of operant variability.

Graded structure is characteristic of operant response classes as well, even when variability is reinforced. Reinforcement may flatten the within-class probability distributions, but as seen in many studies (Hunziker, Saldana, &

Figure 7. Depiction of how elicited, or induced, variability, along the y-axis, changes with interreinforcement time, along the x-axis, when reinforcers are delayed (blackouts precede reinforcers) (top) and when blackouts follow reinforcers (bottom). The shaded portions indicate blackout periods when responses were not possible, and the open portions indicate the periods where responses could be emitted. (Adapted with permission from Wagner, K. & Neuringer, A. (2006) Operant variability when reinforcement is delayed. Learning & Behavior, 34, 111-123.)

Neuringer, 1996; Jensen & Neuringer, 2009; Neuringer et al., 2001; Pesek-Cotton, Johnson, & Newland, 2011), differences in probabilities often remain. Contributing to the within-class hierarchies are types of operanda, distances among operanda, distances to the reinforcer dispenser, etc. In short, both operant responses and category instances appear to be probabilistically generated, with the probabilities organized and influenced by environmental variables including, importantly, feedback from reinforcers.

Functionality of Operant Variability

Conditioning of difficult-to-learn responses. Skinner (1981) suggested that operant behaviors are selected by reinforcers from a substrate of varying behaviors in a way analogous to the evolutionary process of variation and selection. Others have supported the parallel (Baum, 1994; Catania, 1995; Hull, Langman, & Glenn, 2001; Staddon & Simmelhag, 1971). As discussed above, selective pressures maintain variability-generation in the genome with consequent selection of instances leading to evolved changes. A question is whether direct reinforcement of variability (selection of variability) would contribute to acquisition of operant responses (selection of instances).

Neuringer, Deiss, and Olson (2000, Exp 2) reinforced three groups of rats for a difficult-to-learn 5-response target sequence, RLLRL. A Control group was reinforced *only* for the target, with all other 5-response sequences leading to brief timeouts. A Var group was reinforced for varying 5-response sequences as well as being reinforced for RLLRL whenever it occurred. Var reinforcers were limited to no more than one per min (VI 1 min). A yoke group, referred to as Any, received the same VI 1 min reinforcers as the Var group, but they were given following completion of any sequence and these animals were not required to vary. As in the other two conditions, Any animals were always reinforced for the target sequence. The main result was that only the Var group learned the target sequence (Figure 8). The added VI reinforcers enabled both Var and Any rats to respond throughout the experiment whereas the absence of these reinforcers caused responding to extinguish in most members of the Control group. But only the Var group maintained high levels of variability until the target sequence was learned. The experiment was replicated with a different target sequence, LLRRL, and again, only the Var group learned, this shown in the bottom of Figure 8 (see also Neuringer, 1993). It was hypothesized that reinforcement of variability provided the requisite baseline for contact to be made between the target sequence and reinforcers. Since shaping of new responses always depends upon such contact, the concurrent reinforcement of variability may facilitate shaping whenever baseline variability is low.
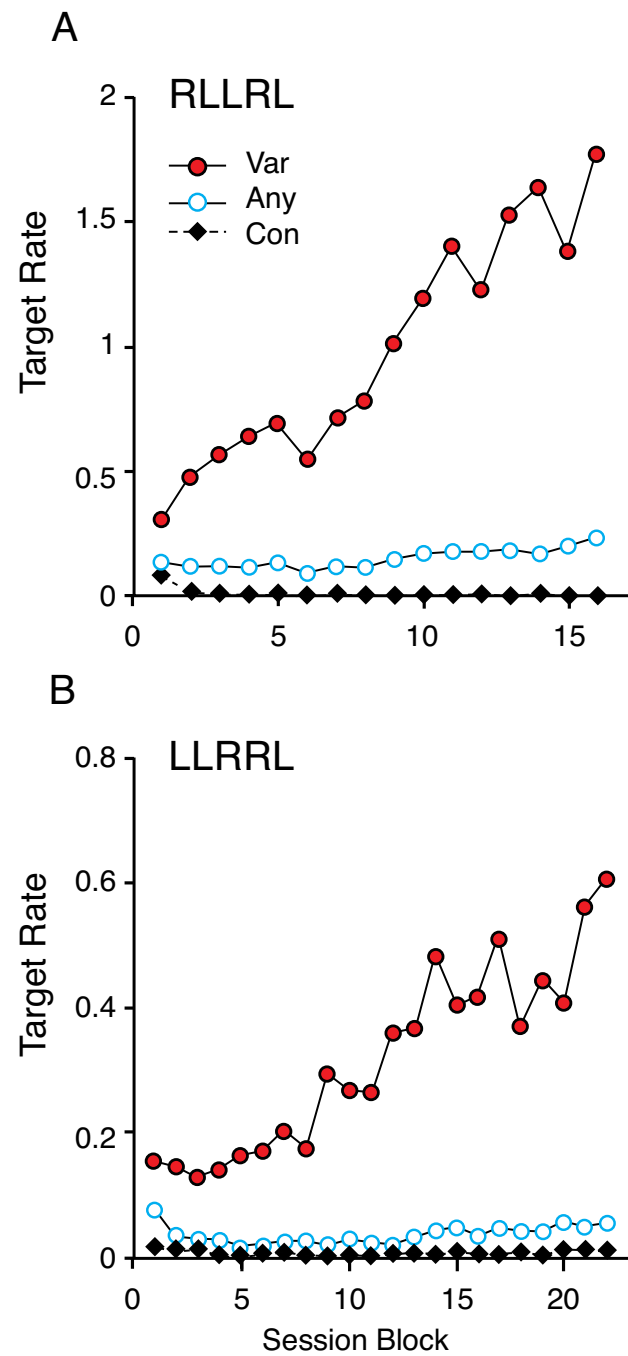
Figure 8. The y-axes show the rates at which a target sequence was emitted when the target was RLLRL (top) and LLRRL (bottom). Group averages are shown for three groups of rats (10 rats each): Var group was reinforced on the average of once per min for variable sequences plus reinforced whenever the target sequence was emitted; Any group was reinforced on the average of once per min for any sequence plus whenever the target was emitted; Con group was reinforced only whenever the target was emitted. Session blocks are shown on the x-axis with each point an average across 5 sessions. (Adapted with permission from Neuringer, A., Deiss, C., & Olson, G. (2000) Reinforced variability and operant learning. Journal of Experimental Psychology: Animal Behavior Processes, 26, 98-111.)

As part of the Grunow and Neuringer (2002) experiment, described above, levels of reinforced variability were shown also to contribute to the facilitative effects: the higher the variability, the more likely that a difficult-to-learn target sequence was acquired. However, human participants, playing a computer game analogous to these rat experiments, did not learn target sequences faster when variability was reinforced (Bizo & Doolan, 2008; Maes and van der Goot, 2006). The human control groups (who were reinforced only for emitting the target and not for variable sequences) were the only ones to learn. Neuringer (2009) discusses possible reason for this species difference, but as yet there is no clear explanation. We note, however, that the motivations of a deprived animal working for needed food, and doing so over the course of one hundred or more sessions, differs appreciably from that of a human participant spending a brief time at a computer in a psychology experiment.

Problem solving. Arnesen (2000; see also Neuringer, 2004) studied whether rewarding rats for variable interactions with objects would facilitate their ability to explore a novel space and discover food hidden within and under novel objects. Rats in a Variability group were reinforced with food pellets for varying object interactions, such as touching a soup can, pushing it, climbing on it, poking nose in it, etc., with the overall goal being to reinforce variable responses to the object. A new object was provided for each of 10 sessions with variable responses reinforced throughout. A Yoke control group experienced the same objects for the same time periods but food pellets were given without regard to the rats' interactions with the objects. A second control group was simply handled. Following these experiences, each rat was placed alone in a 6 ft by 8 ft room, on the floor of which were 30 objects (e.g., a toy truck, metal plumbing pipes, a hair brush, a doll's chest-of-drawers), chosen arbitrarily but as different as possible from those used during the training phase. Hidden within or under each object was a food pellet and the hungry rats were permitted to explore freely for 20 min. The Variability group discovered and consumed significantly more pellets than either of the control groups, which did not differ from one another. The Variability rats also explored more (and seemingly more boldly) than the controls, many of whom showed signs of fear such as hovering along the wall of the room and freezing should they accidentally cause an object to move or produce a noise. Thus, learning to interact variably with objects facilitated exploration and discovery of reinforcers in a novel, foraging-type environment. The advantages incurred by variations are discussed in the human problem-solving literature, such as brainstorming, but there have been few tests of direct reinforcement-of-variability procedures for problem solving more generally.

Autism and depression. Low levels of variability are characteristic of some pathologies. A question of applied interest is whether direct reinforcement of variability can influence those levels. Miller and Neuringer (2000) reinforced five individuals diagnosed with autism. Such individuals often demonstrate stereotypic, highly repetitive behaviors. These five, plus nine control participants (children and adults), were reinforced independently of response variability during a baseline phase (Prob) of a simple computer-game procedure. This was followed by a Var phase in which sequence variations were required for reinforcement and then a return to Prob. The participants with autism behaved less variably than the normal controls throughout the experiment but when directly reinforced, variability increased in both those with autism and the controls (Figure 9). The important point is that response variability increased in individuals with autism, an outcome often difficult to obtain.
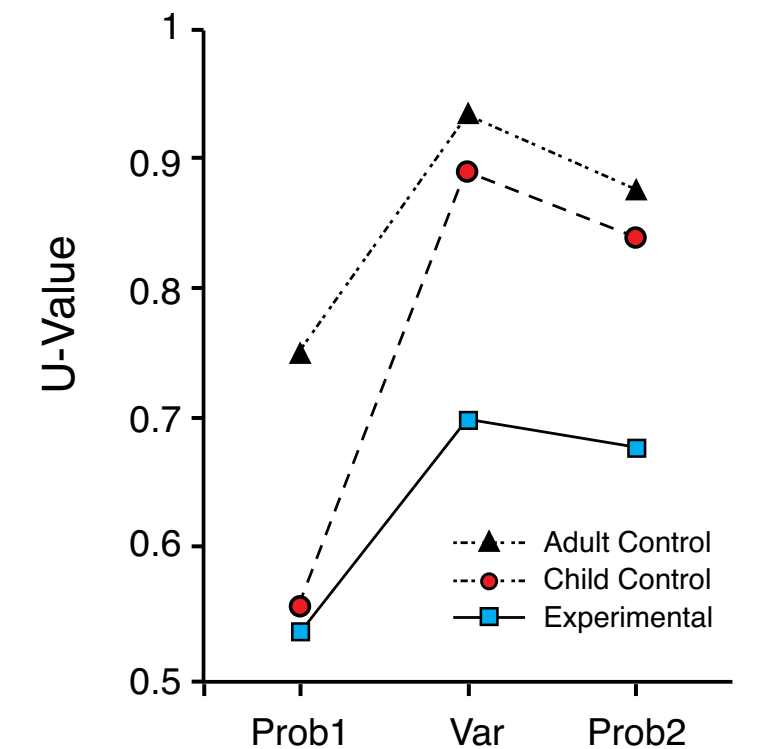
Figure 9. U values based on the 16 possible sequences for each of three groups of human participants. During the Prob phases, reinforcers were provided independently of sequence variability. During the Var phase, reinforcers depended upon variability. The Experimental group consisted of 5 individuals who had been diagnosed with autism and were in a residential treatment program. The Adult control participants were 5 college students. Four children, ages 4 to 9, made up the Child control group. (Adapted with permission from Miller, N. & Neuringer, A. (2000). Reinforcing variability in adolescents with autism. Journal of Applied Behavior Analysis, 33, 151-165.)

Ronald Lee and co-workers extended this work by reinforcing individuals with autism for varying appropriate verbal responses to questions (Lee, McComas, & Jawor, 2002;

Lee & Sturmey, 2006). The efficacy of direct reinforcement was also shown by Newman, Reinecke, and Meinberg (2000): two of three young children diagnosed with autism learned to self-administer reinforcers contingent upon their own increasingly varied responses. Thus, although not extensive, the experimental evidence indicates that the abnormally low levels of variability characteristic of individuals with autism may be influenced by contingencies of reinforcement directed at variability. Because operant behaviors generally manifest some level of within-class variability, and because the variability is normally consequence-controlled, an important step in helping to change autistic behaviors in the direction of normalcy may be explicitly to reinforce for varying levels of variability, levels that range from unpredictable to repetitive.

In an experiment similar to the work with autism, Hopkinson and Neuringer (2003) asked whether low behavioral variability associated with depression (Channon & Baker, 1996; Horne, Evans, & Orne, 1982) could be increased by direct reinforcement. Based on their scores from the Center for Epidemiological Studies Depression Scale (Radloff,
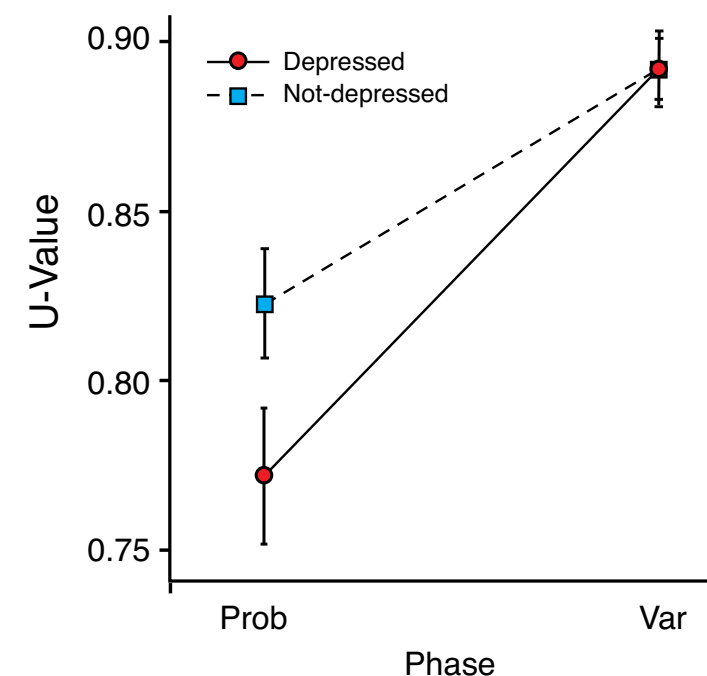
Figure 10. *U values as a function of reinforcement contingencies in a computer game. During Prob, reinforcers were delivered independently of variability levels. During Var, reinforcers depended on variability in the same game. Seventy-five undergraduate students were divided into moderately depressed (36 participants) and not depressed (39 participants) based on a self-evaluation scale. Error bars indicate standard errors. (Adapted with permission from Hopkinson, J., & Neuringer, A. (2003). Modifying behavioral variability in moderately depressed students. Behavior Modification, 27, 251-264.)*

1991), college students were divided into mildly depressed and not depressed groups. Each participant then played a computer game in which sequences of responses were first reinforced independently of variability (Prob), followed by direct reinforcement of variable sequences (Var). Figure 10 shows that, under Prob, the depressed students' variability (average U value) was significantly lower than the non-depressed. When variability was explicitly reinforced, however, levels of variability increased in both groups to the same high levels. This result, if general, is important since it indicates that variability can be explicitly reinforced in those manifesting mild depression (see also Beck, 1976).

Many other cases demonstrate functionality of variability, both induced and reinforced. In competitive situations, such as war or some games, unpredictability is a way to thwart an opponent and attain a goal. Variability increases attention and counteracts habituation. And variations in variability/predictability are found throughout the arts.

Determinism and Stochasticity

It is impossible to *prove* that a given output, no matter how long the series or observation period, is generated by a stochastic process. Any finite[2] stochastic-seeming sequence could be the result of a deterministic process. For example, computer-based random number generators use non-linear deterministic algorithms to generate random numbers. Outputs may appear to be unpredictable and, indeed, pass many tests of randomness, but if one knows the algorithm, each instance can be predicted. Computer algorithms periodically reseed themselves using the system clock to keep their predictability from becoming a security liability. As will be seen, human participants can learn to generate stochastic-like but completely predictable outputs, in ways similar to that of a computer.

However, it is equally the case that determinism can't be proved: Variability is always observed at *some level of analysis*, and proving a thesis to be universally true is impossible by empirical means. These points may be obvious, but need be stated because many individuals, scientists as well as others, take firm positions: In simplified terms, some argue that the universe is determined while others insist that the universe contains indeterminate events.

The determinism versus indeterminism debate has roots in early Greek times – Democritus versus Epicurus – and continues to this day in philosophy (Kane, 2002). In science, the issue was intensely argued in late 19th- and early 20th-century physics, with Ludwig Boltzmann being the first to definitively undermine the prevailing paradigm of the clockwork universe. In Boltzmann's view, atomic particles could be understood as inherently probabilistic. The insight

that irreducible uncertainty underlies physical phenomena was profoundly disruptive and ran contrary to the assumptions of many, but also initiated the revolutionary work that laid the foundation for modern theoretical physics, including quantum mechanics (Lindley, 2001).

The conflict between a clockwork universe and one in which there exists fundamental uncertainty is by no means merely historical. It underlies a clash between *frequentist* and *Bayesian* statistical approaches today (Bland & Altman, 1998). In most of the sub-disciplines within psychology, researchers are implicit adherents of the determinist position; and mainstream psychology in the 20th century has relied heavily on frequentist analyses, such as analysis of variance (ANOVA). Frequentist statistical analyses assume that sets of measurements provide approximations of a single true value. Put another way, those who rely on ANOVAs and related statistical procedures implicitly assume that variance is caused by errors in measurement, and that a sufficiently large number of measurements will permit the true value to be identified. Bayesian approaches, by contrast, treat the world as inherently probabilistic, and assume that measured variations in values characterize true distributions. However, the assumption that events are inherently probabilistic, while widely accepted in contemporary physics, is unacceptable to most psychologists, and this has contributed to the view that response variability is merely error to be reduced and factored out (see, also, Bayarri & Berger, 2004).

With respect to operant variability, there is support for both deterministic and indeterministic processes playing a role, often in conjunction with one another. In some cases, generation of highly variable responses relies chiefly on memory for past events. In others, the evidence is consistent with a primarily stochastic process. We will consider evidence for both positions.

Memory-Based Operant Variability

Memory-based theories of operant variability posit that individual events – stimuli, reinforcers, and previous responses – can be identified that enable exact predictions of future responses, even when they appear to be stochastically or randomly generated

Radial arm maze experiments. Responses of a rat in a radial-arm maze are partly based on the rat's memory for its previous experiences during a session (Olton, Collison, & Werz, 1977; Olton & Samuelson, 1976). The rat is free to explore a maze that consists of 8 (or more) arms radiating from a central platform with a pellet of food located at the end of each arm. Once eaten, the pellets are not replaced. Thus, it is advantageous for the rat to avoid previously visited arms. Rats are quite good at the task and (after some

experience) make few repetition errors, although they often visit the remaining arms in unpredictable order. We see in these experiments memory (for visited arms) combined with possibly stochastic selection among the remaining arms. Additional evidence for the involvement of memory in this task is seen in the rat's errors when timeout periods are imposed between arm entries. For example, if the rat is removed from the maze after half of the pellets had been consumed, longer intervals prior to returning to the maze result in a larger number of reentry errors (Beatty & Shavalia, 1980). Thus, radial arm maze performance combines memory-based non-repetitions with stochastic-like choices among as-yet unentered arms. Consistent with this conclusion are findings that alcohol administration increases number of repeated arm entries, presumably because memory is degraded (Devenport, Merriman, & Devenport, 1983). The increase in errors yields sequences that more closely resemble stochastic choices, i.e., moves the rats from memory-based choice allocations to stochastic allocation (McElroy & Neuringer, 1990).

Lag schedules. As indicated above, under Lag 50 schedules, where the current response sequence must differ from each of the previous 50 sequences, pigeons respond in a stochastic-like manner (Page & Neuringer, 1985). Consistent with this claim is that both pigeons and stochastic simulations generate more repetition errors than would be obtained from a purely memory-based strategy, e.g., cycling repeatedly across 50 different sequences. However, memory-based strategies are often observed under Lag 1 or 2 schedules, with the animals in fact cycling through 2 or 3 sequences. The result of such cycling is reinforcement for every sequence, which is a better return rate than would result from stochastic choices. Machado (1993) showed that cycling occurs only when the memory demands are within the subject's capacity. Using a frequency-dependent variability contingency, Machado found that when the contingencies differentially reinforced the least frequent individual response (given the possibility of L or R), the pigeons responded LRLRLR…, the optimal solution. When the contingencies reinforced the least frequent of pairs of responses (LL, LR, RL, or RR), the birds again developed memory-based patterns, e.g., repeating RRLLRRLL. However, when triads were the unit, the birds apparently could not develop the optimal fixed pattern of RRRLRLLL… but, instead, reverted to "random-like behavior" (Machado, 1993, p.103).

Similar results were obtained from song birds when variable songs were reinforced (Manabe et al., 1997). Under Lag 1, the birds tended to generate two songs, using a win-stay, lose-switch strategy; under Lag 2, they generated three songs. But when the Lag was increased to 3, multiple strategies emerged, some of which were highly stochastic. Thus, a memory-based strategy was employed when possible, but

when the memory demands became too high, stochastic allocation emerged. Reinforcement of vocal variations has been extended to other species including walruses (Schusterman & Reichmuth, 2008) and people (Lee et al., 2002).

Chaotic responding. Chaos theory describes phenomena that appear to be random and unpredictable on their surface but are in fact generated by non-linear, deterministic processes. Chaotic behavioral strategies can result in highly variable outputs but do so in a manner in which each output is precisely controlled by prior events (Hoyert, 1992; Mosekilde, Larsen, & Sterman, 1991; Townsend, 1992). Neuringer and Voss (1993) showed that human participants could learn to generate chaotic-like sequences: individual responses appeared to be unpredictable, but were based on (and predictable from) the logistic difference function:

$$R_n = t \cdot R_{n-1} \cdot \left(1 - R_{n-1}\right) \qquad \text{[Eq. 3]}$$

Here, R_n refers to the nth iteration in a series, each R is a value between 0.0 and 1.0, and t is a constant between 1.0 and 4.0. As t approaches 4.0, outputs appear increasingly to be unpredictable despite being strictly determined. The participants were shown the difference between each of their responses and that of the iterated logistic-difference equation with t=4.0. With training, they learned to generate highly variable responses (top panel in Figure 11), but when responses in the current trial (n) were plotted as a function of responses in the just-prior trial ($n-1$), the data were closely fit by a parabolic function (bottom panel). Because each iteration of the equation is completely determined by the prior output (given the one multiplicative constant), it is reasonable to assume that responses were based on memory for the previous response, with the participants having learned (or memorized) a long series of "if the previous response was value A, then the current response must be value B" pairs (Metzger, 1994; Ward & West, 1994). Neuringer and Voss (described in Neuringer, 2002) provided evidence for this hypothesis. Introducing delays or otherwise interfering with ongoing responding resulted in degrading of the target chaotic sequence. Thus very high levels of surface variability can be memory based, but as discussed next, experimental results are also consistent with stochastic generating processes in which responses do not depend upon memory for prior responses or stimuli.

Stochastic Variability

A number of studies report that responses approximate those expected from a stochastic process. For example, in one, human participants learned to satisfy 10 statistical tests of randomness (Neuringer, 1986) and, in a self experiment performed by the senior author, 30 statistical tests of ran-
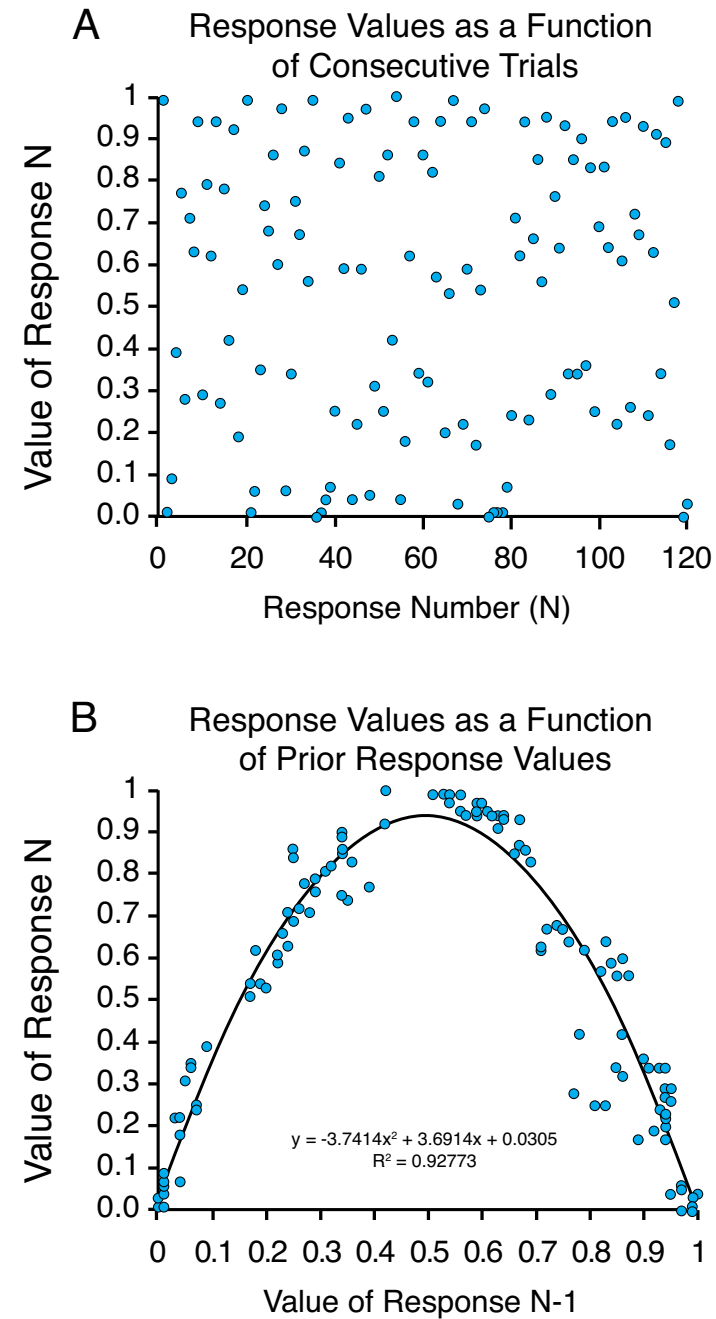
Figure 11. Generation of chaotic-like sequences by one human participant. The top panel shows the values of individual responses (y-axis) across consecutive responses (x-axis). The bottom panel shows the values of response n (y-axis) as a function of the values of response n-1 (x-axis). These panels show that highly "noisy" responding (top) was generated by a highly orderly deterministic process (bottom). (Adapted with permission from Neuringer, A. & Voss, C. (1993). Approximating chaotic behavior. Psychological Science, 4, 113-119.)

domness were satisfied (Roberts & Neuringer, 1998). These results differed, however, from many previous studies in which participants failed to produce equiprobable, random-like responses (Brugger, 1997; Wagenaar, 1972).
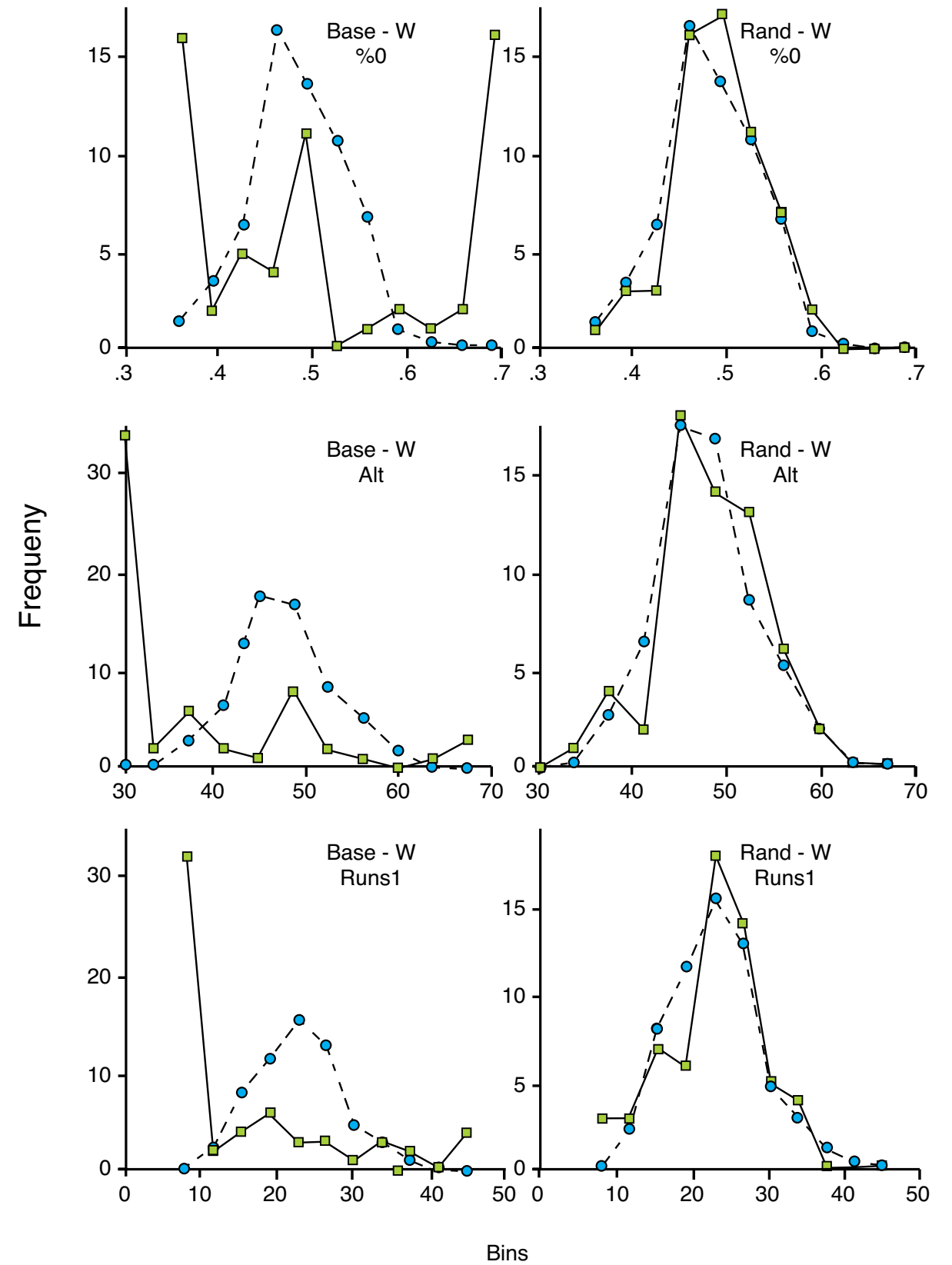
Figure 12. Frequency distributions for one participant on three statistical evaluations of response randomness. The left column shows performances during baseline conditions when responses produced no feedback. The right column shows performances after extended reinforcement for approximating random outputs. The solid line shows the participant's performance and the dotted line shows comparable data generated by a computer-based random number generator. (Adapted with permission from Neuringer, A. (1986). Can people behave "randomly?": The role of feedback. Journal of Experimental Psychology: General, 115, 62-75.)

Procedural details indicate why the results (and conclusions) differed. In Neuringer (1986), students generated sequences of 1's and 2's on a computer keyboard, with each set of 100 responses constituting one trial. The students were instructed to respond as randomly as possible, as if they were tossing a coin and calling heads or tails. A baseline phase lasted for 60 such trials for a total of 6000 responses. The only feedback following each trial (set of 100 responses) was to indicate that the trial had been completed. As in all previous studies in this area, the participants' responses differed significantly from a stochastic model. During the training phase that followed, the participants received feedback at the end of each trial, enabling them to compare their performances to a stochastic model, first according to one statistical test, then another, until following each trial, feedback was provided from 10 different statistics. The distributions of the 10 statistics, which differed from the stochastic model at the beginning of training, came to approximate the model at the end (Figure 12). That is, according to 10 tests, the participants learned to approximate a random model.

The Neuringer (1986) study differed from previous ones in a number of ways. It was the first to *explicitly reinforce* equiprobable, random-like responding in human participants; in previous cases, feedback was not provided. Participants in the Neuringer study generated tens of thousands of responses, whereas previous experiments often asked for as few as 100 responses. The feedback and extensive training enabled Neuringer's participants to learn to avoid the biases (e.g., short runs) found in previous research. In all studies of human random generation, participants indeed vary their responses but practice and reinforcing feedback may be necessary to approximate unbiased, equiprobable, random outputs.

Let's look critically at the evidence just described. When trying to decide whether a particular response stream had been stochastically generated, the best a researcher can do is to estimate the *probability* that a stochastic process was involved. For example, if the first 100 selections were green from a well-mixed urn containing 500 red and 500 green balls, it would be unlikely *but not impossible*, that the balls were selected randomly. *Any* sub-sequence of any length is possible, and every *particular sequence* is exactly as likely as any other of equal length (see Lopes, 1982). These considerations indicate the impossibility of proving that a particular finite sequence *deviates* from random: The observed sequence *may have* been selected from an infinite random series (see Chaitin, 1975). However, the *probability* of approximating 100 green balls in a row is extremely low; whereas approximating a 50-50 spit between green and red balls is much more likely – there are many more sequences that yield a 50-50 split than 100-0. Thus, one can specify the likelihood that a given output matches the characteristics of

a stochastic model whose outputs are of the same length[3]. A second problem is that seemingly stochastic outputs may be generated by non-stochastic processes, such as iterations of the logistic difference equation or the digits of π. Thus, behavioral outputs can be highly variable and at the same time predictable and consistent with a determinist model.

There may be experimental ways, however, to assess whether variable behavior is generated by a stochastic or a deterministic process, and these involve interfering events. We will focus on comparing a stochastic-generation hypothesis with the most likely deterministic process, one that involves memory. A memory-based response, by definition, depends upon control by prior events, either stimuli or responses, and if an interfering event is interposed between the controlling event and the behavior in question, then memory might be degraded and the outcome suffer. On the other hand, stochastically generated outcomes do not depend upon (nor can they be predicted with knowledge of) prior stimuli or responses and thus interfering events should not affect stochastic outputs. In short, if an interfering event degrades operantly reinforced variations, then that provides evidence consistent with a memory and *against* a stochastic generation process. Absence of memory interference provides evidence *consistent with* stochastic generation. In cases where interference produces a partial reduction in operant variation, it is reasonable to assume that both play a role.

In the Neuringer and Voss experiment described above (Neuringer, 2002), interposed timeout periods between responses interfered with chaotic-type outputs, thereby implicating memory. In another phase of the same experiment, however, participants received statistical feedback, as in the Neuringer (1986) study, leading to stochastic-like responding. Now when the timeouts were interposed, there were little or no detrimental effects, supporting the stochastic claim. Thus, participants responded chaotically in one part of the experiment and stochastically in another – memory dependent in the former and independent of memory in the latter. The operative contingencies controlled which strategy was employed.

As with human stochastic generation, memory interference leaves operant variability intact in rats. For example, Neuringer (1991) reinforced two groups of rats for L and R lever presses, four responses per trial. One group (Var) was trained under a lag 5 contingency, and the other (Rep) for repeating a single sequence, LLRR. It was assumed that accurate Rep performance depends upon working memory. After performances had stabilized, timeouts were inserted between each response, from 0.5 s to 20 s in different phases of the experiment. During these forced pauses, the chamber was dark and responses were not counted. If the animals used the previous response(s) as cues for the current

response, then, it was reasoned, these interpolated pauses should degrade accuracy.

The results from the Rep group were as expected: as pauses increased from 0 to 6 s, percentages of correct sequences were relatively unaffected but beyond 6 s, probability of a correct LLRR sequence indeed fell sharply. The rats appeared able to remember their past responses for up to 6 sec, but not beyond that. The Var group's results were quite different: as pauses increased between 0- and 6-s, the rats were *increasingly likely* to satisfy the variability contingencies, and with pauses greater than 6 s, percent correct remained at a high asymptotic level. Thus, the Var group's results were opposite to the Rep group's and opposite to that predicted from a memory-based strategy. Similar results – showing that operant variability increases or is maintained as responding is slowed – have been reported by Morris (1987) with pigeons and Baddeley (1966) and others with random response generation by human subjects.

Memory for prior responses may not have controlled the Var group's performance but why did variability *increase*? Each of three answers is consistent with a stochastic-generator interpretation. Weiss (1964; 1965) hypothesized that voluntary random generation requires current responses to be independent of previous ones and memory for, or control by, prior responses would interfere with such independence. A second possibility is that at short IRTs, animals tend to repeat responses, or respond twice quickly on the same operandum. Blough (1966) found this in pigeons and excluded such double pecks from his analyses because the double pecks appeared not to be under the control of the reinforcement schedule. Morris (1987) also found a tendency for birds to repeat in the absence of brief timeouts to separate responses. A third hypothesis is that there were two contributors to the observed variability. One was a reinforcement-controlled stochastic-generation process; the other was pause-induced (or elicited) variability. The induced effect – slowed responding generates high variability – is a general phenomenon, supported in many other cases. According to this interpretation, operant variability in the Var group was governed by a stochastic process, operant repetition in the Rep group by a memory-based process, and pauses elicited variability under both contingencies. The result was that Rep performance was interfered with while Var was facilitated. Each of these hypotheses is consistent with a conclusion that memory for (or discriminative control by) prior responses does not contribute to, and possibly interferes with, variable responding when an organism is reinforced for variability, and this, in turn, is consistent with a stochastic-generator hypothesis.

Also cosistent are the effects of alcohol on rats' repetitions (Rep component of a multiple schedule in which LLRR is

reinforced) versus variations (Var component containing lag contingencies) (Cohen et al., 1990). With rats responding under the multiple schedule, injections of ethanol degraded Rep performance but did not affect performance in Var. Similar results were obtained from pigeons when d-amphetamine was administered as well as ethanol (Ward, Bailey, & Odum, 2006; see also Abreu-Rodrigues et al., 2004). We conclude that Rep and Var performances were controlled by different underlying processes, primarily memory-based in Rep and primarily stochastic in Var. When repetitions are reinforced, responses appear to be more sensitive to disruption – by drugs as well as other stimulus and contingency influences – than when variability is reinforced (Doughty & Lattal, 2001).

Other methods have been employed to test the stochastic-generator hypothesis. For example, Page & Neuringer (1985) systematically manipulated number of responses per trial while maintaining a constant lag 3 contingency. In separate phases of the experiment with pigeons, trials consisted of 4, 6, or 8 responses. If previous responses served as cues (e.g., for what *not* to do), then it was hypothesized that performance should be degraded as number of responses per trial increased: 8 responses per trial require subjects to remember more than 4 responses. The stochastic hypothesis predicts the opposite, as demonstrated by the following example. Assume that the lag value was 1, i.e., a sequence was required simply not to repeat the previous trial's sequence. If each trial were 2 responses in length and *stochastically generated*, then the probability that a given trial would repeat the previous one is .25. (There are 4 possible sequences in the first trial – RR, RL, LR, and LL. Thus, the second trial has a 1 in 4 chance of matching the first.) If a trial consists of 4 responses, the probability of a repetition by chance is .0625, or 1 in 16. Thus, if subjects used a stochastic (coin-toss-like) process to generate Ls and Rs, performances should *improve* with increasing responses per trial – reinforcements would be more frequent because repetitions were less frequent by chance. The results from the pigeons were exactly as predicted by the stochastic hypothesis: The probability of satisfying the variability contingency increased as responses per trial increased and the pigeons were reinforced increasingly. It appeared that 8-response trials were easier for pigeons than 4. (For a similar perceptual effect, see Wasserman, Young, & Cook, 2004.) A follow-up study by Jensen, Miller, and Neuringer (2006) confirmed and expanded these results with pigeons and people.

To this point, we've provided evidence that variability is an operant – it can be reinforced – and that a stochastic generating process may be responsible. Without explicit training, the process is often biased, e.g., in terms of response preferences and patterns. With training, responses can approximate a random model. We will now extend the discussion

to three related areas: operant responses generally, choices, and voluntary behaviors. In each of these, we suggest that reinforced variability plays a central, indeed, defining role. Operant behavior is only briefly considered since much of the preceding discussion has been about the stochastic nature of operant classes.

The Stochastic Operant

Many attempts have been made to distinguish between emitted and elicited responses, e.g., between operants and Pavlovian responses. These attempts span identifying different physiological systems to different behavioral contingencies. We cannot resolve the issue but suggest one important attribute of all operant behaviors, namely independent control of variability by contingencies of reinforcement. *Without some degree of selective control over variability, manifest or potential, a behavior is not an operant.* Skinner defined the operant by relationships in a 3-term contingency: In the presence of a discriminative stimulus, if a response produces a reinforcer, and the response is affected, e.g., increases in frequency, then the response is an operant. We revise that view in two ways. First, we define the response in terms of its dimensions rather than proximal outcomes, with reinforcement contingencies affecting one or more dimensions, such as topography, location, rate, force, frequency, or the like. Second, we suggest that, for all operant responses, variability/predictability along at least one of these dimensions must be sensitive to reinforcement contingencies. Stated simply, variability is a reinforceable dimension of emitted behaviors.

All behaviors vary to some degree, of course, and, as indicated above, levels of variability/predictability change as a function of eliciting environmental events and response magnitudes. Thus, neither variability alone nor changing levels of variability indicate operant response. Necessary is *independent* contingency-of-reinforcement control, i.e., independent of elicited or induced effects and independent of average values. Paeye & Madelain (2011) provide an excellent example. The variability of saccadic eye movements was thought to be due exclusively to neural noise until these researchers showed that the variability is independently controlled by consequences. In human participants, reinforcing feedback increased or decreased saccade variability, depending upon the contingency, while average amplitude was unchanged. In a yoked condition, the variability was unaffected by the same reinforcers. Thus, saccade variability was shown to be an operant.

The degree to which reinforcers control variability differs across species, individuals, response types, motivational and drug states, and so on. Sizes of operant classes differ, number of classes that can be demonstrated by a given organism or

by a species differ, within-class probability distributions differ, e.g., whether the probabilities of each possible response are equal, normally distributed, skewed, and so on. And, most importantly, these variables differ in their sensitivity to control by contingencies of reinforcement. That is to say, we hypothesize a continuum of operant control over behavioral variability. Variability is more-or-less operant in nature, and this will characterize species differences (variability is more operant in humans than in drosophila), differences across response domains (verbal variability is more operant than saccade) and differences resulting from intra-organism states, such as age, drugs, and psychopathologies. For example, whether variability is reinforced or not, SHR rats, a proposed model of human Attention Deficit Hyperactivity Disorder (ADHD), vary their response sequences more than do control WKY rats (Mook et al., 1993). The SHRs, however, are less able to alter their levels of variability, e.g., when repetitions of a single sequence are reinforced. Some animals cannot vary their levels (or degrees) of variability as well as other animals. This might parallel the tendency of individuals with ADHD to vary without regard to the contextual demands, highlighting that variability per se does not indicate adaptive operant behavior. Selective, bounded, functionally changing and reinforced variability is the sign of an operant. At the other end of the continuum and as described above, individuals with autism also have great difficulty in modifying levels of behavioral variability, but in this case, they are constrained to the repetition end. Thus, operants can be classified in terms of the potential range of reinforced variations and the ease with which different levels can be generated. Stated differently, the operant nature of a response is closely related to operant variability. The most skilled operants are characterized by highest levels of feedback-controlled variability, with responses readily moving from predictable, repetitive habits to novel, unpredictable, creations.

Stochastic Choice

Behavioral evidence. De Villiers and Herrnstein (1976) conceptualized all operant responses as choices, e.g., between activating an operandum and doing anything else. We take that point, and suggest that just as operant variability is an essential part of all operant behavior, the same holds for all choices. Choices are sometimes repetitive and highly predictable. A person who likes chocolate ice cream will choose it with high probability. A rat that is reinforced only for choices of the left arm of a T-maze will quickly learn to choose that arm predictably. However, given a change in contingencies, choices can change, both in terms of what they are and their predictability. Thus, as with operant responses generally, choices can sometimes be predicted but reinforcement contingencies and other factors can modify that predictability.
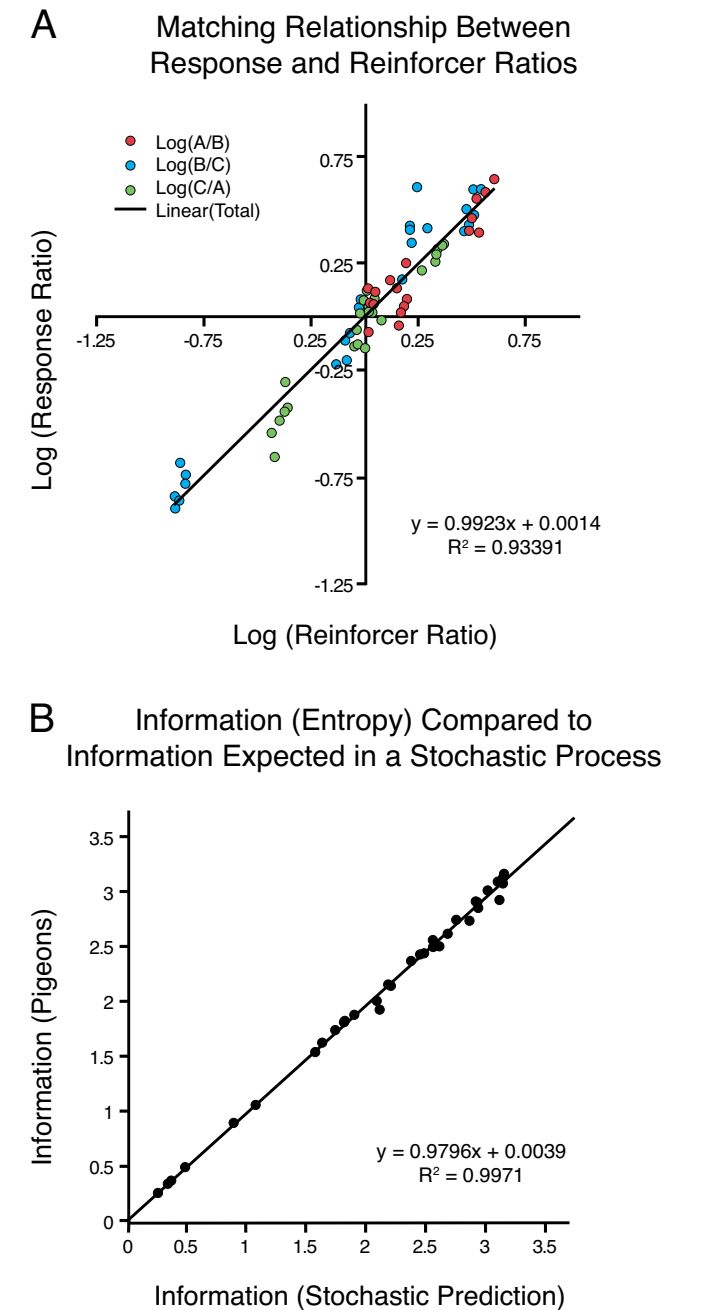
For example, when reinforcers are uncertain – in terms of their location, availability, magnitude, or quality – individual choices are often unpredictable. A commonly studied procedure is the concurrent VI schedule. Under concurrent VI's, reinforcers are independently programmed for two (or sometimes more) options, and subjects choose freely among them. In a VI 1 min : VI 3 min procedure, for example, reinforcers are made available unpredictably on one operandum once per min on average, and on the other once every 3 min on average. The two programming schedules are independent (i.e., reinforcer availability on one operandum has no influence on the other). Also, once a reinforcer becomes available (has set up), it remains available until the next response to that operandum – as when a letter is delivered to a mailbox and remains available until retrieved.

When VI values are systematically varied across phases of an experiment, overall ratios of left-to-right choices are found to be functionally related to ratios of left-to-right obtained reinforcers, a relationship commonly described as a power function and referred to as the *generalized matching law* (Baum, 1974):

$$\frac{C_X}{C_Y} = \left(\frac{k_X}{k_Y}\right) \bullet \left(\frac{R_X}{R_Y}\right)^s \qquad \text{[Eq. 4]}$$

Here, C_X refers to observed choices of alternative X, and R_X corresponds to delivered reinforcers (C_Y and R_Y correspond to alternative Y, accordingly). The parameter k_X refers to bias for X, such as due to side preferences. The s parameter refers to the sensitivity of choice ratios to reinforcement ratios. When $s = 1.0$, choice ratios exactly match (or equal) reinforcement ratios, as was originally described by Herrnstein (1961). In some cases, however, $s < 1.0$ and choice ratios are not as extreme as the ratio of reinforcers; in other cases, $s > 1.0$ and choice ratios are more extreme than reinforcers (see below for further discussion).

The generalized matching law describes molar (i.e.,overall) distributions of choices as a function of obtained reinforcers (Davison & McCarthy, 1988) and is found to hold in many cases of uncertain reinforcements. Molar choices are measured, for example, by the total number of responses on L and R operanda during a session. These totals appear to be generated by stochastic processes: Choice sequences indicate stochasticity (Glimcher, 2003, 2005; Jensen & Neuringer, 2008; Nevin, 1969; see also Silberberg, Hamilton, Ziriax, & Casey, 1978, for an alternative view). Thus, *individual* choices cannot be predicted with substantially greater accuracy than provided by knowledge of their molar distributions. Some biases are observed, (e.g., subjects switch more frequently than predicted from first-order response prob-

A Matching Relationship Between Response and Reinforcer Ratios

Log(A/B)
Log(B/C)
Log(C/A)
Linear(Total)

y = 0.9923x + 0.0014
R² = 0.93391

Log (Response Ratio)
Log (Reinforcer Ratio)

B Information (Entropy) Compared to Information Expected in a Stochastic Process

y = 0.9796x + 0.0039
R² = 0.9971

Information (Pigeons)
Information (Stochastic Prediction)

Figure 13. The upper graph shows logs of response ratios (left key/center key; center key/right key; right key/left key) with individual points representing individual pigeons (6 subjects in the experiment) during individual phases (where ratios of reinforcers were varied). The line shows the least-squares, best fitting function. The lower graph compares the pigeons' distributions of response dyads (LL, LC, LR, CL…), or Information, to those expected from a stochastic generator. To the extent that the data conform to a straight line with slope = 1.0, the pigeons' performances were similar to the stochastic model. (Adapted with permission from Jensen, G. & Neuringer, A. (2008). Choice as a function of reinforcer "hold": From probability learning to concurrent reinforcement. Journal of Experimental Psychology: Animal Behavior Processes, 34, 437-460.)

abilities), but these may be related to the particular contingencies and physical aspects of the testing environment (see below).

Stochastic matching was documented when pigeons chose among three concurrently available sources of uncertain reinforcements (Jensen & Neuringer, 2008). Figure 13 top shows matching of response proportions to obtained reinforcer proportions. The bottom of the figure shows the extent to which choice uncertainty or variability matched that expected from a stochastic source. (The stochastic model, shown on the x-axis, predicted relative frequencies of pairs of the pigeons' responses, shown on the y-axis, based on first-order relative frequencies.) That is, the pigeons' choices were consistent with the stochastic model – selection of colored balls from an urn – described earlier in this paper. In different phases of the experiment, reinforcement frequencies differed (the relative numbers of each color differed), but responses were stochastically generated throughout (selection was blind). Thus, while overall choice proportions could readily be predicted, individual choices could not.

Although often unpredictable, choices become predictable when cues for reward are available. These cues may be external stimuli, such as a change in key color indicating availability and location of a reward, or intrinsic to the schedule of reinforcement. As an example of the latter, under concurrent VI schedules, reinforcers are generally more likely for alternations than for long runs of responses, and therefore switching between operanda is generally higher than would be predicted by a stochastic model (see Jensen & Neuringer, 2008). Changeover delays (i.e., withholding reinforcer availability for periods of time following each switch) are imposed to overcome the high levels of alternation. Similarly, in competitive situations, where two individuals compete for rewards, contextually stochastic response strategies are most effective (Nash, 1951) except when one can predict the opponent's choices, at which point predictable strategies become functional (Dorris & Glimcher, 2004; Lee et al., 2004; Lee, McGreevy, & Barraclough, 2005). Thus, when reinforcer availability is uncertain, choices are emitted stochastically but when an organism can discriminate that a reinforcer is more probable for one of the alternatives, choices are governed by that fact. A combination of stochastic and deterministic strategies best describes choices. As with operant responses generally, choices are governed by functionality, that is, by the contingencies between the choices and reinforcers. And again as with operant responses, stochastic distributions (and therefore predictability of responses) are highly sensitive, and change rapidly, to changes in reinforcer demands (Gallistel, Mark, King, & Latham, 2001).

Physiological evidence. Doris and Glimcher (2004) recorded from single cells in rhesus monkey lateral intraparietal cortex (LIP). These cells receive information from retinal receptive fields and are involved in the control of saccades to those areas. In the experiment, the monkeys were rewarded for looking left or right, the saccades constituting the behavioral choices. Amounts and probabilities of reinforcement for each of the responses were systematically varied across phases of the experiment. In one phase, only one response option was provided at any given time. Firing rates of the LIP neurons tracked reinforcer values, e.g., if left looks were reinforced more frequently than right, the LIP neuron associated with left saccades fired more rapidly than those associated with right. More generally, the LIP neurons responded to the relative values of anticipated reinforcers contingent upon the saccade movements.

In a second phase, the monkeys played a competitive game, with the computer serving as the opponent. Programmed probabilities and amounts of reinforcement were again systematically varied for looking left versus right but in this case the monkey could freely *choose* to look left or right. Furthermore, because the contingencies were those of a competitive game, the monkey was rewarded only if the computer did not correctly predict its choices: The monkey had to outwit the computer. Left/right choice proportions were found to be related to reinforcers by the generalized matching function (Eq. 4) and, at the same time, individual responses were highly unpredictable (although not random) – as necessary to fool the computer (see also Louie & Glimcher, 2010). These saccade responses constituted the behavioral side of the choices. What about the physiological results? The left and right LIP firing rates were found to be approximately equal. This is a wow result because it supports both the Nash equilibrium theory applied to concurrent choices and an explanation of why stochastic choices in fact match reinforcer distributions: they do so in order to equalize subjective values of left and right choices, as indicated by single cells in the cortex. Stated differently, the matching of choices to reinforcement frequencies provided an equilibrium point where values for each of the choices were equal, these values being represented by relative LIP firing rates. (See also Barraclough, Conroy, and Lee, 2004, for related findings.)

Glimcher (2005) outlined evidence showing stochasticity throughout the central nervous system. He writes that whereas the average firing rates by neurons in the visual cortex were precisely controlled by visual stimuli, "…the exact pattern of firing that gave rise to this average rate seemed to be almost completely unpredictable. The time at which a spike occurred could be described as a fully stochastic process …" (p. 46). Glimcher went on to suggest that the source of this randomness is release of neurotransmitters by synaptic vesicles. "Vesicular release seems to be an appar-

ently indeterminate process" (p. 48). To take this one step further, changes in post-synaptic neuron membrane potentials are:

> …a product of interactions at the atomic level, many of which are governed by quantum physics and thus are truly indeterminate events. Because of the tiny scale at which these processes operate, interactions between action potential and transmitter release as well as interactions between transmitter molecules and postsynaptic receptors … seem likely to be fundamentally indeterminate. (p. 49)

But it is within-class indeterminism, since the *possibility* of synaptic vesicular release and post-synaptic activation depends upon an action potential having occurred in the presynaptic nerves, thus activating a class of indeterminate events. We see here a combination of causal determination and indetermination at the level of the single nerve in a way that parallels the determinate influences on operant classes and the indeterminate generation of within-class instances, and, similarly, determinate and indeterminate influences on DNA and gamete formation.

Voluntary Behavior

Operant variability is equally important when attempting to characterize volition. A major issue in philosophical discussions of volition is how to combine the functional, goal-directed, intentional, or rational aspects of voluntary actions with their apparent independence from environmental determination. The first of these implies that a knowledgeable observer should be able to predict behavior, the second that voluntary behavior is unpredictable. These discussions have been ongoing for thousands of years and continue to present times (Kane, 2002). We suggest that voluntary behavior is functional (or intended to be so) and sometimes highly predictable, other times unpredictable, with predictability governed by the same relationships with consequences as for all operants. That is, we hypothesize that *a critical characteristic of the voluntary act is ability to vary levels of predictability under the feedback influence of consequences* (see, also, Brembs, 2011).

Using a psychophysical procedure, Neuringer, Jensen, and Piff (2007) tested this conjecture. Human participants judged that *virtual actors* (dots moving on the screen of a computer) represented voluntary human behavior when the actors' choices (i) matched obtained relative frequencies of reinforcement and (ii) did so by stochastic generation of those choices. Here are some details. The participants observed 6 different actors (on 6 different computer screens) as each made thousands of choices (represented by movements of the dots). Each of the actors chose repeatedly among three options in what was said to be a gambling

game. Reinforcers were programmed by concurrent reinforcement schedules (reinforcement shown by color changes on the screen). The actors' choices and the reinforcers were programmed by, and under control of, the computer. Participants, who were told nothing about the contingencies, played no active role and only observed. Across different phases of the observation period, frequencies of reinforcement for the three choice options were systematically manipulated. The actors' choices were generated by iteration of the generalized matching power function (Eq. 4) extended to a three-alternative situation[4]. But the actors differed in terms of their choice strategies (as given by their s parameters). Some actors chose approximately equally among the three alternatives – manifesting maximum unpredictability – no matter the distributions of reinforcers (low s value). Some chose predominantly the highest payoff (high s value), despite the fact that additional reinforcers could be obtained from the two other alternatives. And one actor matched response probabilities to obtained reinforcers (s=1.0), thereby varying both distributions of responses and levels of predictability.

Following the observation periods, the participants judged how well the actors' choices represented *voluntary choices made by a real human player*. Figure 14 shows estimates by the participants (in two experiments) of how well the ac-
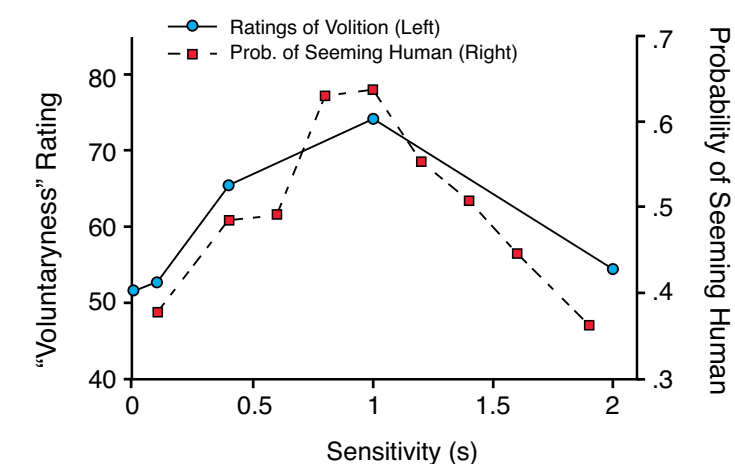
Figure 14. Ratings of how well individual actors represented voluntary human choices (left axis) and, in a separate experiment, the probabilities of identifying an actor as a "voluntarily choosing human player." The x-axis shows different actors, from low s values, indicating an actor who responded maximally unpredictably under all conditions, to high s values, indicating an actor who repeated choices predictably much of the time. An s value of 1.0 indicated a stochastically matching actor whose levels of predictability changed with the distributions of reinforcers. (Adapted with permission from Neuringer, A., Jensen, G., & Piff, P. (2007). Stochastic matching and the voluntary nature of choice. Journal of the Experimental Analysis of Behavior, 88, 1-28.)

tors represented volitional choices. The $s=1.0$ actor, whose choice distributions most closely matched the reinforcer distributions, was rated as most similar to a person who was making voluntary choices.

A series of control experiments evaluated alternative explanations. For example, rates of reinforcement were overall slightly higher (across different phases) for the $s = 1.0$ matcher than for any of the other actors and one control showed that overall reinforcement rates were not responsible for the volitional judgments. The most important control procedure tested whether matching of responses to reinforcers alone implied volition or whether variations in levels of predictability were important. Stated differently, were the volitional judgments governed by changes in response distributions or response predictability or both? To answer this questions, a different set of participants compared two actors, both of whom exactly matched choice proportions to reinforcer proportions; however, one actor matched by stochastically allocating its choices whereas the other allocated its choices in an easily predictable fashion. The stochastic matcher responded as follow: if reinforcers were programmed for the three choice alternatives in a ratio of 5:3:2, the stochastic matcher responded to the left alternative with a .5 probability, the center a .3 probability and the right with a .2 probability. When reinforcers were equal across the three choices – .33 : .33 : .33 – predicting the next choice was exceedingly difficult but when reinforcers were predominantly obtained from one of the alternatives – .9 : .05 : .05 – predictions could easily be made. Thus, although the stochastic matcher indeed responded stochastically throughout the experiment, its choices were more-or-less predictable, depending upon the reinforcement distributions. By contrast, the patterned matcher also matched exactly, but did so in a patterned and therefore readily predictable manner throughout. For example, it would respond LLLLLCCCCRR, cycling through the same 5:3:2 strings of responses, when the reinforcers were programmed with those same ratios. The patterned matcher similarly repeated patterns for all reinforcement distributions. When reinforcers were equal – .33 : .33 : .33 – the patterned matcher responded LCRL-CRLCR… Because both actors matched, they were reinforced equally. The results were clear. Participants judged the stochastic matcher to represent a voluntary human player significantly better than the patterned one, showing that both functionality (matching, in this case) and stochasticity were jointly necessary for the highest ratings of volition. Thus, a combination of functional choice distributions (matching) and choice variability (more or less predictability) provided the discriminative cues to indicate voluntary behavior.

Experiments on operant variability show that levels, or degrees, of behavioral (un)predictability are guided by environmental consequences. We propose that the same is true for voluntary actions. Voluntary behaviors are sometimes readily predictable, sometimes less predictable, and sometimes quite unpredictable. In all cases, reasons for the general response can be identified (given sufficient knowledge) – but the precise behaviors may still remain unpredictable. For example, under some circumstances, the response to "What are you doing tonight?" can readily be predicted for a given acquaintance. Even when the situation warrants unpredictable responding, some veridical predictions can be made: that the response will be verbal, that it will contain particular parts of speech, and so on. The functionality of variability implies a *degree of predictability* in the resulting behaviors that is related to the activated class. That is, the class can often be predicted based on knowledge of the organism and environmental conditions. But the within-class instance may be difficult or impossible to predict, especially when large response classes are activated.

Unpredictability, real or potential, is emphasized in many discussions of volition. Indeed, the size of the activated set can be exceedingly large – and functionally so – for if someone were attempting to prove that she is a free agent, the set of possibilities might consist of all responses in her repertoire (see Scriven, 1965). But we return to the fact that voluntary behaviors can be predictable as well as not. The most important characteristic is *functionality of variability*, or ability to change levels of predictability in response to environmental demands. Equally this is an identifying characteristic of operant behavior and of choice, where responses are functional and stochastically emitted. Thus, with Skinner, we combine 'voluntary' and 'operant' in a single phrase, but research now indicates why that is appropriate. Operant responses are voluntary precisely because they combine functionality with levels of predictability.

References

Abreu-Rodrigues, J., Hanna, E. S., de Mello Cruz, A. P, Matos, R., & Delabrida, Z. (2004). Differential effects of midazolam and pentylenetetrazole on behavioral repetition and variation. *Behavioural Pharmacology, 15*, 535-543. doi.org/10.1097/00008877-200412000-00002 PMid:15577452

Amabile, T. M. (1983). *The social psychology of creativity*. New York: Springer-Verlag. doi.org/10.1007/978-1-4612-5533-8

Antonitis, J. J. (1951). Response variability in the white rat during conditioning, extinction, and reconditioning. *Journal of Experimental Psychology, 42*, 273-281. doi.org/10.1037/h0060407 PMid:14888834.

Arnesen, E. M. (2000). *Reinforcement of object manipulation increases discovery*. Unpublished undergraduate the-
sis, Reed College, Portland, OR.

Akins, C. K., Domjan, M., & Gutiérrez, G. (1994). Topography of sexually conditioned behavior in male Japanese quail (*Coturnix japonica*) depends on the CS-US interval. *Journal of Experimental Psychology: Animal Behavior Processes, 20*, 199-209. doi.org/10.1037/0097-7403.20.2.199 PMid:8189188

Baddeley, A. D. (1966). The capacity for generating information by randomization. *Quarterly Journal of Experimental Psychology, 18*, 119-129. doi.org/10.1080/14640746608400019   PMid:5935121

Barraclough, D. J., Conroy, M. L., & Lee, D. (2004). Prefrontal cortex and decision making in a mixed-strategy game. *Nature Neuroscience, 7*, 404-410. doi.org/10.1038/nn1209   PMid:15004564

Barsalou, L. W. (1987). The instability of graded structures in concepts. In U. Neisser, ed. *Concepts and conceptual development: Ecological and intellectual factors in categorization* (pp. 101-140). New York: Cambridge University Press.

Baum, W. M. (1974). On two types of deviation from the matching law: Bias and undermatching. *Journal of the Experimental Analysis of Behavior, 22*, 231-242. doi.org/10.1901/jeab.1974.22-231   PMid:16811782 PMCid:1333261

Baum, W. M. (1994). *Understanding behaviorism*. New York: Harper Collins.

Bayarri, M. J. & Berger, J. O. (2004). The interplay of Bayesian and frequentist analysis. *Statistical Science, 19*, 58-80. doi.org/10.1214/088342304000000116

Beatty, W. W., & Shavalia, D. A. (1980). Spatial memory in rats: Time course of working memory and effect of anesthetics. *Behavioral and Neural Biology, 28*, 454-462. doi.org/10.1016/S0163-1047(80)91806-3

Beck, A. (1976). *Cognitive therapy and the emotional disorders*. New York: The New American Library.

Bizo, L. A. & Doolan, K. (May, 2008). Reinforced behavioural variability in humans. Paper presented at the Association for Behavior Analysis Meeting, Chicago.

Bland, J. M. & Altman, D. G. (1998). Bayesians and frequentists. *British Medical Journal, 317*, 1151. doi.org/10.1136/bmj.317.7166.1151 PMid:9784463   PMCid:1114120

Blough, D. S. (1966). The reinforcement of least-frequent interresponse times. J*ournal of the Experimental Analysis of Behavior, 9*, 581-591. doi.org/10.1901/jeab.1966.9-581 PMid:5964515   PMCid:1338218

Bouton, M. E. (1994). Conditioning, remembering, and forgetting. *Journal of Experimental Psychology: Animal Behavior Processes, 20*, 219-231. doi.org/10.1037/0097-7403.20.3.219

Brembs, B. (2011). Towards a scientific concept of free will as a biological trait: spontaneous action and decision-making in invertebrates. *Proceedings of the Royal Society, B, 278*, 930-939. doi.org/10.1098/rspb.2010.2325 PMid:21159679   PMCid:3049057

Brugger, P. (1997). Variables that influence the generation of random sequences: An update. *Perceptual & Motor Skills, 84*, 627-661. doi.org/10.2466/pms.1997.84.2.627 PMid:9106859

Caporale, L. H. (1999). Chance favors the prepared genome (1999) *Annals of the New York Academy of Sciences, 870*, 1-21.   doi.org/10.1111/j.1749-6632.1999.tb08860.x PMid:10415469

Catania, A. C. (1995). Selection in biology and behavior. In J. T. Todd & E. K. Morris (Eds.). *Modern perspectives on B. F. Skinner and contemporary behaviorism* (pp. 185-194). Westport, CT: Greenwood Press.

Catchpole, C. K. & Slater, P. J. (1995). *Bird song: Biological themes and variations*. Cambridge: Cambridge University Press.

Chaitin, G. J. (1975). Randomness and mathematical proof. Scientific American, 232, 47-52. doi.org/10.1038/scientificamerican0575-47

Chance, M. R. A. (1957). The role of convulsions in behavior. *Behavioral Science, 2*, 30-45. doi.org/10.1002/bs.3830020104

Channon, S. & Baker, J. E. (1996). Depression and problem-solving performance on a fault-diagnosis task. Applies Cognitive Psychology, 10, 327-336. doi.org/10.1002/(SICI)1099-0720(199608)10:4<327::AID-ACP384>3.0.CO;2-O

Cherot, C., Jones, A., & Neuringer, A. (1996). Reinforced variability decreases with approach to reinforcers. *Journal of Experimental Psychology: Animal Behavior Processes, 22*, 497-508. doi.org/10.1037/0097-7403.22.4.497 PMid:8865615

Cohen, L., Neuringer, A., & Rhodes, D. (1990). Effects of ethanol on reinforced variations and repetitions by rats under a multiple schedule. *Journal of the Experimental Analysis of Behavior, 54*, 1-12. doi.org/10.1901/jeab.1990.54-1   PMid:2398323 PMCid:1322941

Craig, W. (1918). Appetites and aversions as constituents of instincts. *Biological Bulletin, 34*, 91-107. doi.org/10.2307/1536346

Davison, M., & McCarthy, D. (1988). *The matching law*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Denney, J. & Neuringer, A. (1998). Behavioral variability is controlled by discriminative stimuli. *Animal Learning & Behavior, 26*, 154-162. doi.org/10.3758/BF03199208

Devenport, L. D., Merriman, V. J, & Devenport, J. A. (1983). Effects of ethanol on enforced spatial variability in the 8-arm radial maze. *Pharmacology, Biochemistry, & Behavior, 18*, 55-59. doi.org/10.1016/0091-3057(83)90251-4

De Villiers, P. A. & Herrnstein, R. J. (1976). Toward a law of response strength. *Psychological Bulletin, 83*, 1131-1153. doi.org/10.1037/0033-2909.83.6.1131

Dorris, M. C. & Glimcher, P. W. (2004) Activity in posterior parietal cortex is correlated with the relative subjective desirability of action. *Neuron, 44*, 365-378. doi.org/10.1016/j.neuron.2004.09.009 PMid:15473973

Doughty, A. H. & Lattal, K. A., (2001). Resistance to change of operant variation and repetition. *Journal of the Experimental Analysis of Behavior, 76*, 195-215. doi.org/10.1901/jeab.2001.76-195 PMid:11599639   PMCid:1284834

Driver, P. M. & Humphries, D. A. (1988). *Protean behavior: The biology of unpredictability.* Oxford: Oxford University Press.

Eckerman, D. A. & Lanson, R. N. (1969). Variability of response location for pigeons responding under continuous reinforcement, intermittent reinforcement, and extinction. *Journal of the Experimental Analysis of Behavior, 12*, 73-80.   doi.org/10.1901/jeab.1969.12-73 PMid:16811341   PMCid:1338576

Galbicka, G. (1994) Shaping in the 21st century: Moving percentile schedules into applied settings. *Journal of Applied Behavior Analysis, 27*, 739-760. doi.org/10.1901/jaba.1994.27-739 PMid:16795849   PMCid:1297861

Gallistel, C. R., Mark, T. A., King, A. P., & Latham, P. E. (2001). The rat approximates an ideal detector of changes in rates of reward: Implications for the law of effect. *Journal of Experimental Psychology: Animal Behavior Processes, 27*, 354-372. doi.org/10.1037/0097-7403.27.4.354 PMid:11676086

Gharib, A., Gade, C., & Roberts, S. (2004). Control of variation by reward probability. *Journal of Experimental Psychology: Animal Behavior Processes, 30*, 271-282. doi.org/10.1037/0097-7403.30.4.271 PMid:15506853

Glimcher, P. W. (2003). *Decisions, uncertainty, and the brain.* Cambridge, MA: MIT Press.

Glimcher, P. W. (2005). Indeterminacy in brain and behavior. *Annual Review of Psychology, 56*, 25-56. doi.org/10.1146/annurev.psych.55.090902.141429 PMid:15709928

Goetz, E. M. & Baer, D. M. (1973). Social control of form diversity and emergence of new forms in children's block-building. *Journal of Applied Behavior Analysis, 6*, 209-217.   doi.org/10.1901/jaba.1973.6-209 PMid:16795402   PMCid:1310828

Grunow, A. & Neuringer, A. (2002). Learning to vary and varying to learn. *Psychonomic Bulletin & Review, 9*, 250-258.   doi.org/10.3758/BF03196279

Guthrie, E. R. & Horton, G. P. (1946). Cats in a Puzzle Box.

New York: Rinehart.

Herrnstein, R. J. (1961). Relative and absolute strength of response as a function of frequency of reinforcement. *Journal of the Experimental Analysis of Behavior, 4*, 267-272. doi.org/10.1901/jeab.1961.4-267 PMid:13713775   PMCid:1404074

Hopkinson, J., & Neuringer, A. (2003). Modifying behavioral variability in moderately depressed students. *Behavior Modification, 27*, 251-264. doi.org/10.1177/0145445503251605 PMid:12705108

Horne, R. L., Evans, F. J., & Orne, M. T. (1982). Random number generation, psychopathology, and therapeutic change. *Archives of General Psychiatry, 39*, 680-683. doi.org/10.1001/archpsyc.1982.04290060042008 PMid:7092501

Hoyert, M. S. (1992). Order and chaos in fixed-interval schedules of reinforcement. *Journal of the Experimental Analysis of Behavior, 57*, 339-363. doi.org/10.1901/jeab.1992.57-339 PMid:16812657   PMCid:1323234

Hull, D. L., Langman, R. E., Glenn, S. S. (2001). A general account of selection: Biology, immunology, and behavior. *Behavioral and Brain Sciences, 24*, 511-573. PMid:11682800

Humphries, D. A. & Driver, P. M. (1970). Protean defence by prey animals. *Oecologia, 5*, 285-302. doi.org/10.1007/BF00815496

Hunziker, M. H. L., Saldana, R. L., & Neuringer, A. (1996). Behavioral variability in SHR and WKY rats as a function of rearing environment and reinforcement contingency. *Journal of the Experimental Analysis of Behavior, 65*, 129-144.   doi.org/10.1901/jeab.1996.65-129 PMid:8583193   PMCid:1350067

Jensen, G., Miller, C., & Neuringer, A. (2006). Truly random operant responding: Results and reasons. In E. A. Wasserman & T. R. Zentall, eds. *Comparative cognition: Experimental explorations of animal intelligence* (pp. 459-480). Oxford: Oxford University Press.

Jensen, G. & Neuringer, A. (2008). Choice as a function of reinforcer "hold": From probability learning to concurrent reinforcement. *Journal of Experimental Psychology: Animal Behavior Processes, 34*, 437-460. doi.org/10.1037/0097-7403.34.4.437 PMid:18954229   PMCid:2673116

Jensen, G. & Neuringer, A. (2009). Barycentric extension of generalized matching. *Journal of the Experimental Analysis of Behavior, 92*, 139-159. doi.org/10.1901/jeab.2009.92-139 PMid:20354596   PMCid:2732327

Kane, R., ed. (2002). *The Oxford handbook of free will.* Oxford: Oxford University Press.

Knuth, D. E. (1969). *The art of computer programming.*

Reading, MA: Addison-Wesley.

Lee, D., Conroy, M. L, McGreevy, B. P., & Barraclough, D. J. (2004). Reinforcement learning and decision making in monkeys during a competitive game. *Cognitive Brain Research, 22*, 45-58. doi.org/10.1016/j.cogbrainres.2004.07.007 PMid:15561500

Lee, D., McGreevy, B. P., & Barraclough, D. J. (2005). Learning and decision making in monkeys during a rock-paper-scissors game. *Cognitive Brain Research, 25*, 416-430.   doi.org/10.1016/j.cogbrainres.2005.07.003 PMid:16095886

Lee, R., McComas, J. J., & Jawor, J. (2002). The effects of differential and lag reinforcement schedules on varied verbal responding by individuals with autism. *Journal of Applied Behavior Analysis, 35*, 391-402. doi.org/10.1901/jaba.2002.35-391 PMid:12555910   PMCid:1284401

Lee, R. & Sturmey, P. (2006). The effects of lag schedules and preferred materials on variable responding in students with autism. *Journal of Autism and Developmental Disorders, 36*, 421-428. doi.org/10.1007/s10803-006-0080-7 PMid:16568357

Lee, R., Sturmey, P., & Fields, L. (2007). Schedule-induced and operant mechanisms that influence response variability: A review and implications for future investigations. *The Psychological Record, 57*, 429-455.

Lindley, D. (2001). *Boltzmann's atom: The great debate that launched a revolution in physics.* New York: The Free Press.

Lopes, L. L. (1982). Doing the impossible: A note on induction and the experience of randomness. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 8*, 626-636.   doi.org/10.1037/0278-7393.8.6.626

Louie, K. & Glimcher, P. W. (2010). Separating value from choice: Delay discounting activity in the lateral intraparietal area. *Journal of Neuroscience, 30*, 5498-5507. doi.org/10.1523/JNEUROSCI.5742-09.2010 PMid:20410103   PMCid:2898568

Machado, A. (1989). Operant conditioning of behavioral variability using a percentile reinforcement schedules. *Journal of the Experimental Analysis of Behavior, 52*, 155-166   doi.org/10.1901/jeab.1989.52-155 PMid:2794841   PMCid:1338957

Machado, A. (1992). Behavioral variability and frequency-dependent selection. *Journal of the Experimental Analysis of Behavior, 58*, 241-263. doi.org/10.1901/jeab.1992.58-241 PMid:1402601   PMCid:1322058

Machado, A. (1993). Learning variable and stereotypical sequences of responses: Some data and a new model. *Behavioural Processes, 30*, 103-130.

doi.org/10.1016/0376-6357(93)90002-9

Machado, A. (1994). Polymorphic response patterns under frequency-dependent selection. *Animal Learning & Behavior, 22*, 53-71.   doi.org/10.3758/BF03199956

Machado, A. (1997). Increasing the variability of response sequences in pigeons by adjusting the frequency of switching between two keys. *Journal of the Experimental Analysis of Behavior, 68*, 1-25. doi.org/10.1901/jeab.1997.68-1 PMid:9241860   PMCid:1284614

Macnab, R. M. & Koshland, D E., Jr. (1972). The gradient-sensing mechanism in bacterial chemotaxis. *Proceedings of the National Academy of Sciences of the United States of America, 69*, 2509-2512. doi.org/10.1073/pnas.69.9.2509

Madelain, L, Chaprenaut, L, & Chauvin, A. (2007). Control of sensorimotor variability by consequences. *Journal of Neurophysiology, 98*, 2255-2265. doi.org/10.1152/jn.01286.2006 PMid:17699687

Maes, J. H. R. (2003). Response stability and variability induced in humans by different feedback contingencies. *Learning & Behavior, 31*, 332-348. doi.org/10.3758/BF03195995

Maes, J. H. R. & van der Goot, M. (2006). Human operant learning under concurrent reinforcement of response variability. *Learning and Motivation, 37*, 79-92. doi.org/10.1016/j.lmot.2005.03.003

Manabe, K., Staddon, J. E. R., & Cleaveland, J. M. (1997). Control of vocal repertoire by reward in Budgerigars (Melopsittacus undulatus). *Journal of Comparative Psychology, 111*, 50–62. doi.org/10.1037/0735-7036.111.1.50

McElroy, E. & Neuringer, A. (1990). Effects of alcohol on reinforced repetitions and reinforced variations in rats. *Psychopharmacology, 102*, 49-55. doi.org/10.1007/BF02245743   PMid:2392507

Mechner, F. (1958). Probability relations within response sequences under ratio reinforcement. *Journal of the Experimental Analysis of Behavior, 1*, 109-121. doi.org/10.1901/jeab.1958.1-109 PMid:16811206   PMCid:1403928

Metzger, M. A. (1994). Have subjects been shown to generate chaotic numbers? Commentary on Neuringer and Voss. *Psychological Science, 5*, 111-114. doi.org/10.1111/j.1467-9280.1994.tb00641.x

Miller, G. F. (1997). Mate choice: From sexual cues to cognitive adaptations. In Cardew, G. (ed.), *Characterizing human psychological adaptations* (Ciba foundation Symposium 208), New York: John Wiley, 71-87.

Miller, N. & Neuringer, A. (2000). Reinforcing variability in adolescents with autism. *Journal of Applied Behavior Analysis, 33*, 151-165. doi.org/10.1901/jaba.2000.33-151

PMid:10885524   PMCid:1284235

Mook, D. M., Jeffrey, J., & Neuringer, A. (1993). Spontaneously hypertensive rats (SHR) readily learn to vary but not to repeat instrumental responses. *Behavioral & Neural Biology, 59*, 126-135.
doi.org/10.1016/0163-1047(93)90847-B

Morris, C. J. (1987). The operant conditioning of response variability: Free-operant versus discrete-response procedures. *Journal of the Experimental Analysis of Behavior, 47*, 273-277.   doi.org/10.1901/jeab.1987.47-273
PMid:16812481   PMCid:1348310

Mosekilde, E., Larsen, E. R., & Sterman, J. D. (1991). Coping with complexity: Deterministic chaos in human decision-making behavior. In Casti, J. L. & Karlqvist, A. (eds.), *Beyond belief: Randomness, prediction, and explanation in science*, Boston: CRC Press, 199-229.

Murphy, G. L. (2002). *The big book of concepts*, Cambridge: MIT Press.

Nash, J. (1951). Non-cooperative games. *Annals of Mathematics, 54*, 286-295.   doi.org/10.2307/1969529

Neuringer, A. (1986). Can people behave "randomly?": The role of feedback. *Journal of Experimental Psychology: General, 115*, 62-75.
doi.org/10.1037/0096-3445.115.1.62

Neuringer, A. (1991). Operant variability and repetition as functions of interresponse time. *Journal of Experimental Psychology: Animal Behavior Processes, 17*, 3-12.
doi.org/10.1037/0097-7403.17.1.3

Neuringer, A. (1992). Choosing to vary or repeat. *Psychological Science, 3*, 246-250.
doi.org/10.1111/j.1467-9280.1992.tb00037.x

Neuringer, A. (1993). Reinforced variation and selection. *Animal Learning & Behavior, 21*, 83-91.
doi.org/10.3758/BF03213386

Neuringer, A. (2002). Operant variability: Evidence, functions, and theory. *Psychonomic Bulletin & Review, 9*, 672-705.   doi.org/10.3758/BF03196324

Neuringer, A. (2003). Creativity and reinforced variability. In K. A. Lattal and P. N. Chase (Eds.) *Behavior theory and philosophy* (pp. 323-338). New York: Plenum Publishing.

Neuringer, A. (2004). Reinforced variability in animals and people. *American Psychologist, 59*, 891–906.
doi.org/10.1037/0003-066X.59.9.891
PMid:15584823

Neuringer, A. (2009). Operant variability and the power of reinforcement. *The Behavior Analyst Today, 10*, 319-343. Retrieved Oct. 30, 2011 from http://baojournal.com/BAT%20Journal/BAT%2010-2.pdf

Neuringer, A., Deiss, C., & Olson, G. (2000). Reinforced variability and operant learning. *Journal of Experimental Psychology: Animal Behavior Processes, 26*, 98-111.
doi.org/10.1037/0097-7403.26.1.98
PMid:10650547

Neuringer, A. & Jensen, G. (2010). Operant variability and voluntary action. *Psychological Review, 117*, 972-993.
doi.org/10.1037/a0019499   PMid:20658860

Neuringer, A., Jensen, G., & Piff, P. (2007). Stochastic matching and the voluntary nature of choice. *Journal of the Experimental Analysis of Behavior, 88*, 1-28.
doi.org/10.1901/jeab.2007.65-06
PMid:17725049   PMCid:1918082

Neuringer, A., Kornell, N., & Olufs, M. (2001). Stability and variability in extinction. *Journal of Experimental Psychology: Animal Behavior Processes, 27*, 79-94.
doi.org/10.1037/0097-7403.27.1.79
PMid:11199517

Neuringer, A. & Voss, C. (1993). Approximating chaotic behavior. *Psychological Science, 4*, 113-119.
doi.org/10.1111/j.1467-9280.1993.tb00471.x

Nevin, J. A. (1969). Interval reinforcement of choice behavior in discrete trials. *Journal of the Experimental Analysis of Behavior, 12*, 875-885.
doi.org/10.1901/jeab.1969.12-875
PMid:16811416   PMCid:1338697

Newman, B., Reinecke, D. R., & Meinberg, D. L. (2000). Self-management of varied responding in three students with autism. *Behavioral Interventions, 15*, 145-151.
doi.org/10.1002/(SICI)1099-078X(200004/06)15:2<145::AID-BIN50>3.0.CO;2-3

Nickerson, R. A. (2002). The production and perception of randomness. *Psychological Review, 109*, 330-357.
doi.org/10.1037/0033-295X.109.2.330
PMid:11990321

Notterman, J. M. & Mintz, D. E. (1965). *Dynamics of response*. New York: Wiley.

Odum, A. L., Ward, R. D., Barnes, C. A., & Burke, K. A. (2006). The effects of delayed reinforcement on variability and repetition of response sequences. *Journal of the Experimental Analysis of Behavior, 86*, 159-179.
doi.org/10.1901/jeab.2006.58-05
PMid:17002225   PMCid:1592357

Olton, D. S., Collison, C., & Werz, M. A. (1977). Spatial memory and radial arm maze performance of rats. *Learning & Behavior, 8*, 289-314.
doi.org/10.1016/0023-9690(77)90054-6

Olton, D. S. & Samuelson, R. J. (1976). Remembrance of places passed: spatial memory in rats. *Journal of Experimental Psychology: Animal Behavior Processes, 2*, 97-116.   doi.org/10.1037/0097-7403.2.2.97

Page, S. & Neuringer, A. (1985). Variability is an operant. *Journal of Experimental Psychology: Animal Behavior Processes, 11*, 429-452.
doi.org/10.1037/0097-7403.11.3.429

Paeye, C. & Madelain, L. (2011). Reinforcing saccadic amplitude variability. *Journal of the Experimental Analysis of Behavior, 95*, 149-162.

Neuringer, A. & Jensen, G. (2010). Operant variability and voluntary action.

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.   doi.org/10.1901/jeab.1985.44-217
PMid:16812432   PMCid:1348179

Pennisi, E. (1998). How the genome readies itself for evolution. *Science, 281*, 1131-1134.
doi.org/10.1126/science.281.5380.1131   PMid:9735027

Pesek-Cotton, E. R., Johnson, J. E., & Newland, M. C. (2011). Reinforcing behavioral variability: An analysis of dopamine-receptor subtypes and intermittent reinforcement. *Pharmacology, Biochemistry and Behavior, 97*, 551-559.   doi.org/10.1016/j.pbb.2010.10.011
PMid:21050870

Platt, M. L. & Glimcher, P. W. (1999). Neural correlates of decision variables in parietal cortex. *Nature, 400*, 233-238.   doi.org/10.1038/22268   PMid:10421364

Pryor, K. W., Haag, R., & O'Reilly, J. (1969). The creative porpoise: Training for novel behavior. *Journal of the Experimental Analysis of Behavior, 12*, 653-661.
doi.org/10.1901/jeab.1969.12-653
PMid:16811388   PMCid:1338662

Radloff, L. S. (1991) The use of the Center for Epidemiological Studies Depression Scale in adolescents and young adults. *Journal of Youth and Adolescence, 20*, 149-166.
doi.org/10.1007/BF01537606

Roberts, S. & Neuringer, A. (1998). Self-experimentation. In K. A. Lattal and M. Perone (Eds.) *Handbook of Research Methods in Human Operant Behavior* (pp. 619-655). New York: Plenum Press.

Rosch, E. H. (1978). Principles of categorization. In E. H. Rosch & B. Lloyd, eds. *Cognition and categorization* (pp. 27-48). Hillsdale: Erlbaum Associates.

Ross, C. & Neuringer, A. (2002). Reinforcement of variations and repetitions along three independent response dimensions. *Behavioural Processes, 57*, 199-209.
doi.org/10.1016/S0376-6357(02)00014-1

Sakata, J. T., Hampton, C. M., & Brainard, M. S. (2008). Social modulation of sequence and syllable variability in adult birdsong. *Journal of Neurophysiology, 99*, 1700-1711.   doi.org/10.1152/jn.01296.2007
PMid:18216221

Schoenfeld, W. N., Harris, A. H., Farmer, J. (1966) Conditioning response variability. *Psychological Reports, 19*, 551-557.   doi.org/10.2466/pr0.1966.19.2.551
PMid:5912818

Schusterman, R. J. & Reichmuth, C. (2008). Novel sound production through contingency learning in the Pacific walrus (*Odobenus rosmarus divergens*). *Animal Cognition, 11*, 319-327.
doi.org/10.1007/s10071-007-0120-5

PMid:18038276

Schwartz, B. (1982). Failure to produce response variability with reinforcement. *Journal of the Experimental Analysis of Behavior, 37*, 171-181.
doi.org/10.1901/jeab.1982.37-171
PMid:16812263   PMCid:1333132

Schwartz, B. & Lacey, H. (1982). *Behaviorism, science, and human nature, 2nd ed.*, New York: W. W. Norton.

Scriven, M. (1965). An essential unpredictability in human behavior. In Wolman, B., ed. *Scientific psychology* (pp. 411-425). New York: Basic Books.

Searcy, W. A., & Yasukawa, K. (1990). Use of song repertoire in intersexual and intrasexual contexts by male red-winged blackbirds. *Behavioral & Ecological Sociobiology, 27*, 123-128.   doi.org/10.1007/BF00168455

Shannon, C. E. (1948). A mathematical theory of communication. *The Bell Systems Technical Journal, 27*, 379-423, 623-656.
http://cm.bell-labs.com/cm/ms/what/shannonday/paper.html

Shimp, C. P. (1967). Reinforcement of least-frequent sequences of choices. *Journal of the Experimental Analysis of Behavior, 10*, 57-65.   doi.org/10.1901/jeab.1967.10-57
PMid:16811306   PMCid:1338318

Silberberg, A., Hamilton, B., Ziriax, J. M., & Casey, J. (1978). The structure of choice. *Journal of Experimental Psychology: Animal Behavior Processes, 4*, 368-398.
doi.org/10.1037/0097-7403.4.4.368

Skinner, B. F. (1938). *The behavior of organisms*. New York: Appleton-Century-Crofts.

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

Souza, A. S., Abreu-Rodrigues, J., & Baumann, A. A. (2010). History effects on induced and operant variability. *Learning & Behavior, 38*, 426-437.
doi.org/10.3758/LB.38.4.426

Staddon, J. E. R. & Simmelhag, V. L. (1971). The "superstition" experiment: A reexamination of its implications for the principles of adaptive behavior. *Psychological Review, 78*, 3-43.   doi.org/10.1037/h0030305

Stokes, P. D. (1995) Learned variability. *Animal Learning & Behavior, 23*, 164-176.   doi.org/10.3758/BF03199931

Thorndike, E. L. (1911). *Animal intelligence*, New York: Macmillan.

Townsend, J. T. (1992). Chaos theory: A brief tutorial and discussion. In Healy, A. F., Kosslyn, S. M., & Shiffrin, R. M. (eds.), *From Learning Theory to Connectionist Theory: Essays in Honor of William K. Estes*, Hillsdale: Erlbaum, Vol. 1, 65-96.

Wagenaar, W. A. (1972). Generation of random sequences by human subjects: A critical survey of literature. *Psychological Bulletin, 77*, 65-72.   doi.org/10.1037/h0032060

Wagner, K. & Neuringer, A. (2006). Operant variability

when reinforcement is delayed. *Learning & Behavior, 34*, 111-123.
doi.org/10.3758/BF03193187

Ward, L. M. & West, R. L. (1994). On chaotic behavior. *Psychological Science, 5*, 232-236.
doi.org/10.1111/j.1467-9280.1994.tb00506.x

Ward, R. D., Bailey, E. Ml. & Odum, A. LO. (2006). Effects of d-amphetamine and ethanol on variable and repetitive key-peck sequences in pigeons. *Journal of the Experimental Analysis of Behavior, 86*, 285-305.
doi.org/10.1901/jeab.2006.17-06
PMid:17191754   PMCid:1679968

Ward, R. D., Kynaston, A. D., Bailey, E. M., & Odum, A. L. (2008). Discriminative control of variability: Effects of successive stimulus reversals. *Behavioural Processes, 78*, 17-24.   doi.org/10.1016/j.beproc.2007.11.007
PMid:18191503

Wasserman, E. A., Young, M. E., & Cook, R. G. (2004). Variability discrimination in humans and animals: Implications for adaptive action. *American Psychologist, 59*, 879-890.   doi.org/10.1037/0003-066X.59.9.879
PMid:15584822

Weiss, R. L. (1964). On producing random responses. *Psychological Reports, 14*, 931-941.
doi.org/10.2466/pr0.1964.14.3.931

Weiss, R. L. (1965). "Variables that influence random-generation": An alternative hypothesis. *Perceptual & Motor Skills, 20*, 307-310.   doi.org/10.2466/pms.1965.20.1.307
PMid:14286544

Young, M. E., & Wasserman, E. A. (2001). Entropy and variability discrimination. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 27*, 278-293.
doi.org/10.1037/0278-7393.27.1.278
PMid:11204103

1. The hazard function associated with an exponential distribution is "flat" because regardless of how many seconds have already elapsed, an observer still only knows that the next upcoming second has a 0.5 probability of ending the interval.

2. Or infinite, as with the digits of π.

3. Formally, this often entails calculating the likelihoods of various models, based on the observed evidence, and comparing them to see which model better explains that evidence.

4. See Jensen & Neuringer (2009) for a generalization of Eq. 4.