

## Truly Random Operant Responding: Results and Reasons

GREG JENSEN, CLAIRE MILLER, AND ALLEN NEURINGER

In his influential book on evolutionary game theory, John Maynard Smith wrote, “I cannot see . . . (why) . . . animals do not have roulette wheels in their heads. . . . If it were selectively advantageous, a randomising device could surely evolve. . . .” (1982, p. 76).

We submit that such a device has evolved. In support, this chapter discusses evidence for three related claims: Response variability can be reinforced (i.e., it is an operant); reinforcers exert precise control over what, where, when, and how much to vary; and the resulting responses are at least sometimes stochastic (or random) in nature (i.e., they are emitted probabilistically and therefore unpredictably).

These are controversial claims, but their roots go back to the early history of Western thought. Epicurus suggested that random swerves of atoms help to explain novelty, creativity, and the initiation of action. He was objecting to Democritus’s deterministic philosophy. Many philosophers and psychologists since that time, including Gustav Fechner and William James, have posited random-like behaviors (see Neuringer, 2003, 2004). But an operant “randomizing device in the head” flies in the face of an assumption dear to most psychologists: namely, that psychological phenomena are determined by prior events—and therefore ultimately predictable. Until recently, there has been little direct evidence concerning the existence, characteristics, and functions of an operant variability-generating process. We describe current evidence, including three previously

unreported experiments, but first provide some definitions.

### DEFINITIONS

*Variability* Variability has many meanings. Sometimes, it implies ignorance of causal factors. Other times it is used in a statistical sense to indicate the spread of values, as in standard deviation and confidence intervals. Sometimes the term implies random-like outputs or a high degree of uncertainty. The term also refers to a dimension or continuum, ranging from repetition (and therefore high predictability) to random (and therefore maximal uncertainty). Context provides the appropriate meaning.

*Random* In lay terms, *random* often connotes “do anything” or “without reason.” That is not what we intend. Rather, we use *random* and *stochastic* interchangeably in their technical senses, to indicate that members of a specified set occur independently of prior events, and therefore that although a knowledgeable observer can predict relative frequencies (or probabilities, these terms also used interchangeably), particular instances cannot be predicted or explained at a more precise level than that of the probability statements.

Because stochasticity is often misunderstood, an example might help. Imagine a large, revolving

barrel filled with 1,000 blue (B) and 1,000 green (G) balls. A blindfolded individual picks balls one at a time and an observer notes the color before the ball is replaced in the barrel. The resulting sequence of Bs and Gs is stochastic, meaning that knowledge of past selections enables an observer to predict future relative frequencies—approximately equal numbers of Bs and Gs—but not the next color (or sequence of colors). The barrel could contain unequal numbers, e.g., 900 Bs and 100 Gs, and, under those conditions, B would be much more likely than G, but the resulting sequence would nevertheless be stochastic (probabilities of 0.9 and 0.1, respectively). This same analysis holds no matter how many different colors were in the barrel—two, four, eight, and so on—because blind selection would in each case result in stochastic outcomes.

*Operant* A rat pressing a lever for food pellets provides an example of the relationship between operant response and reinforcer. The lever press produces food and the food influences the action of pressing, both being necessary for the response to be defined as an operant and for the consequence to be defined as a reinforcer. Reinforcement shapes and maintains operant responses. Reinforcement also affects individual dimensions of responses, such as response force and speed. For example, if food depends on rapid responding, then high response rates may result.

## EVIDENCE FOR THE OPERANT NATURE OF VARIABILITY

### Reinforcement

Variability is influenced by reinforcers contingent on it. For example, when porpoises were reinforced for novel responses, they came eventually to emit behaviors not previously observed in any porpoise (Pryor, Haag, & O'Reilly, 1969). When pigeons were reinforced if their interresponse times—the intervals between consecutive pecks to a response key—were distributed in random-like fashion, pecks came to resemble the random emission of atomic particles (Blough, 1966). In another experiment, when pigeons were reinforced for infrequently occurring sequences of pecks across two response keys, the distribution of sequences came to match that expected from a random source (Machado,

1989). In yet another example, when high-school students were reinforced for random-like generation of sets of 100 responses across two computer keys, the students' performances came to approximate the random model (Neuringer, 1986). In many other experiments as well, animals and people have successfully been reinforced for generating highly variable behaviors (Barba & Hunziker, 2002; Machado, 1997; Neuringer, Deiss, & Olson, 2000; Neuringer & Huntley, 1992; Page & Neuringer, 1985).

There are many possible sources of variability, of course, including noise in the environment and withholding or decreasing reinforcement, and control procedures are required before we can conclude that variability is an operant. As one example, pigeons in an experimental condition were reinforced whenever a sequence of eight responses across left (L) and right (R) keys differed from each of the preceding 50 sequences, a contingency referred to as lag 50 (Page & Neuringer, 1985). In a "yoked" control condition, the same frequency and distribution of reinforcers were presented but now contingent only on the pigeon responding eight times per trial. Thus, in the experimental condition, trials terminated with food only if sequences varied, but under the yoked condition, the food did not depend on variability—the pigeons could vary or not without penalty. The important finding was that response variability was significantly higher when explicitly reinforced than not. This type of comparison provides strong evidence for the operant nature of response variability (see also Blough, 1966; Machado, 1989; Neuringer, 1986).

### Discriminative Cues

Another characteristic of an operant is influence by discriminative cues. For example, if response-contingent food is available only when a 1,000-Hz tone sounds, rats learn to respond when the tone is on and not in its absence. The same is true for response variability: For example, when rats were reinforced for varying in the presence of one stimulus and reinforced independently of variability in a different stimulus (the yoked control), levels of variability were significantly higher in the experimental period than in the yoked period (Denney & Neuringer, 1998). Other studies showed stimulus control over varying versus response repetitions (Cohen, Neuringer, & Rhodes, 1990; Page &

Neuringer, 1985). In summary, variability—of response topography, speed, and sequencing—is controlled in ways that are characteristic of operant responses.

### Response Dimensions

As suggested earlier, reinforcers also control individual dimensions of response, a phenomenon that applies to variability as well. For example, human participants were reinforced for drawing rectangles on a computer screen, and three attributes of the response were monitored—the area of the rectangle, its location on the screen, and its shape (Ross & Neuringer, 2002). Participants were reinforced for varying along two of the dimensions (e.g., shape and size), while simultaneously keeping the third constant (e.g., repeating the location of the rectangle). Participants learned quickly to satisfy the contingencies.

### Response Sets

Reinforcement also establishes the set of responses from which variations emerge. This fact is shown indirectly by all experiments on operant variability: Animals and people tend to limit their responses to the potentially reinforced ones, although others are certainly possible (see, for example, Neuringer, Kornell, & Olufs, 2001). A more direct demonstration was provided by Mook, Jeffrey, and Neuringer (1993): Rats were reinforced for varying sets of four responses across L and R levers, but, to be reinforced, sequences had to start with two L responses. The rats learned to limit their sequences while, at the same time, their responses varied among the potentially reinforced options. Another example: when variations among four-response sequences across L and R levers were reinforced in rats, except that one particular sequence, LLLL, was never reinforced, the rats learned to vary among all of the sequence other than LLLL (Neuringer, 1993). Thus, reinforcement not only engenders variations, but it helps to define the operative set.

### Levels of Variability

Just as particular response rates can be reinforced, or response forces, so, too, different minimum levels of variability can be required for reinforcement.

For example, Page and Neuringer (1985) used a lag schedule in which, across phases of the experiment, the current sequence of eight responses by pigeons had to differ from at least 1, 5, 10, 20, or 50 previous sequences (lag 1, 5, 10, 20, or 50); variability generally increased with the demands. Similar results were observed by Blough (1966), Machado (1989), and Grunow and Neuringer (2002). These studies leave unanswered, however, whether responding is sensitive to requirements for specific levels of variability (rather than the more permissive minimum levels required in previous experiments), and if so, whether variability can change rapidly in response to reinforcement demands. As described later, important theories of behavior, including matching theory and game theory, require such rapid sensitivity.

## EXPERIMENT 1: REINFORCEMENT OF STOCHASTIC DISTRIBUTIONS

Experiment 1 may best be introduced by returning to the balls-in-barrel example used above. Our question is whether behavioral allocations can change rapidly in a way analogous to what happens when the proportions of Gs and Bs change, such as from equal numbers of Bs and Gs to 3 Bs for 1 G. That is, does reinforcement exert precise and rapid control over distributions of stochastic responses?

A positive finding would be consistent with two influential theories of choice. Matching theory predicts that ratios of choices (analogous to Gs and Bs) will be an orderly function—a power function—of the ratios of reinforcements for those choices (Herrnstein, 1997). Matching theory is supported by much evidence (Davison & McCarthy, 1988), but the theory is silent with respect to how responses are allocated—stochastically or systematically. For example, if the rate of reinforcement for left responses were three times higher than for right, an animal would “match” responses to reinforcers by responding systematically with a 3:1 ratio (e.g., LLLRLLLRLLLR . . .) or, alternatively, by responding stochastically with the same 3:1 ratio (e.g., LRLLLRLRLLL . . .). Stochastic generation is predicted by a second theory, however, namely game theory, developed to explain choices when individuals compete with one another for resources. In such situations, it would be ineffective for one animal to permit an opponent to predict its choices—the opponent could

take countermeasures—whereas stochastic responding is functional. The Nash equilibrium combines matching and game theories in its prediction that, in many competitive situations, reinforcement will be optimized when relative frequencies of stochastic responses match relative frequencies of obtained reinforcers. The Nash equilibrium therefore predicts both matching and stochastic allocation. Glimcher (2003) found evidence for such stochastic matching: monkey and human subjects allocated choices in a way that was consistent with both matching and game theory (i.e., stochastic choices matched reinforcement proportions). This is an important result because it provides experimental support for applying the Nash formulation to operant choices. Experiment 1 used a different procedure to test whether reinforcement directly controls stochastic allocation of choice responses and, if so, to ascertain how rapidly that control is achieved.

### Procedure

College students were divided into Experimental ( $n = 116$ ) and Yoke control ( $n = 30$ ) groups, both gaining points for responding on two keys of a computer keyboard, to be referred to as L and R. The experiment consisted of five phases, each terminating after at least 150 responses and at least 25 points, with the different phases not cued and seamlessly joined in a single session lasting approximately 20 min. In most cases, 150 responses sufficed to gain the required 25 points per phase.

### Experimental Condition

Each of the five phases reinforced a different distribution of stochastic responses across the L and R keys: 0.25–0.75, 0.33–0.67, 0.40–0.60, 0.50–0.50, and back to 0.25–0.75, respectively. Thus, in the first phase, approximately 0.25 of responses were required on one key and 0.75 on the other. In the second phase, the required distribution was approximately 0.33 and 0.67, and so on, with Phase 5 repeating Phase 1. We say that responses had to “approximate” a given distribution because there was a delta window for each distribution such that if the response frequency fell within that window, the participant would be rewarded. Delta windows were created by assessing the performance of a stochastic model and using boundaries that resulted in the model being “reinforced,” according to the contingencies to be described later, on 80% of its trials.

Responses had to satisfy stochastic contingencies simultaneously at three levels of analysis, as described, based on separate response counters at each level. Level 1 analysis was based on overall percentages of L and R responses, their frequencies maintained in two associated counters. The particular keys were not specified; some participants responded more on the left key than the right and others did the opposite, but the response percentages had to be within required delta windows. Thus, in the first phase, for example, the relative frequencies of L and R responses were required to be approximately 0.25 and 0.75 on L and R keys or vice versa.

Level 2 analysis was based on the percentages of pairs of responses, with pair frequencies recorded in four counters: LL, LR, RL, and RR. If LRRLRL-LLR had just been emitted, then (in order of occurrence) one count would be added to the LR counter, one to the RR counter, one to the RL counter, and so on, with the most recent pair being indicated by the rightmost LR in the example. For a response to be reinforced, relative frequencies of the current “possible pairs” were required to fall within the defined delta windows. Referring back to the example, the most recent responses were L followed by R. Therefore, given the L response, LR and LL were the current “possible pairs,” because (again, given that an L response had occurred previously) only those two pairs were possible. Level 2 proportions were calculated by dividing the sum in one of the possible-pair counters (the pair that had actually been emitted) by the sum of the two possible-pair counters—in the example given by dividing the LR counter by the sum of LR + LL counters—and in all other ways treating the data as described for level 1.

Similarly, level 3 consisted of eight counters, LLL, LLR, . . . RRR, with, in the just given example, the terminal triplet being LLR. Level 3 analyses similarly required concordance with the stochastic model for response triplets and, in the example given, the number of LLR sequences was divided by the sum of LLR + LLL sequences—these being the only possible triplets given that LL had been emitted prior to the terminal response. In brief, in Phase 1 (where a 0.25–0.75 distribution was required), the last response in our example was reinforced if Ls and Rs had occurred with approximately the required 0.25–0.75 distribution; if LR and LL had occurred with the same approximate percentages; and similarly if approximately 0.25–0.75 distributions were obtained for LLR and LLL. Of course, the particular “possible” pairs and triplets changed with each response. Thus, for example, if following

emission of LLR, another R were to be emitted, then the current (i.e., prior to the last response) possible pairs would be RL and RR, and the current possible triplets, LRL and LRR. The beauty of the procedure lies in the fact that although increasingly demanding conditional probabilities are assessed, the required relative frequencies—0.25 and 0.75 in our example—remain constant.

One additional detail: In order to differentially weight current responses more heavily than those emitted earlier in a phase, each of the “possible” counters was multiplied by an amnesia coefficient of 0.95 following every response. Returning to our example, following the last R response, the level 1 L and R counters were each multiplied by 0.95 (because both L or R were possible), as well as the level 2 LR and LL counters and the level 3 LLR and LLL counters. Each phase began with all counters preset with a value of 1, and the data and calculations in each phase were independent of all others (see Miller, 2003, for additional procedural details).

Participants, naive to the nature of the task, received feedback that indicated how closely response percentages approximated the stochastic model at each of the three levels of analysis. This feedback consisted of a horizontal line across the center of the computer screen and three pairs of colored dots—red, green, and blue—moving symmetrically around the line. Response percentages at the three levels were normalized so that they could be represented by the single horizontal line and the distance of the dots from the line represented the difference between the participant's relative frequencies and the stochastic model at each of the three levels. A point was awarded when all dots were sufficiently close to the line, that is, within the delta windows, with cumulative points shown by a counter. The participant's task was to keep the balls as close to the horizontal line as possible. Rapid responding was discouraged by a “slow down” message appearing in the middle of the screen whenever response rates exceeded 4/s. In brief, human participants were rewarded for distributing responses in a way that matched (within delta windows) a stochastic model, with the model's response probabilities differing across the five phases.

### Yoked Condition

All aspects of the procedure were the same as for the Experimental participants, except that response distributions and variability had no influence on the displayed dots or the presentation of points. Rather,

for each Yoked participant individually, these depended on an arbitrarily paired Experimental player's responses and rewards; that is, feedback to each Yoked participant was identical to that received by the paired Experimental participant. Thus, when Yoked participants responded, they were reinforced at exactly the same rate and intermittency as the Experimentals. Experimental and Yoked participants were given minimal instructions and told only to try to keep the colored dots close to the horizontal line in order to maximize points.

### Measures

A common measure used to assess the stochasticity of responses is *U value*, which evaluates the distribution of relative frequencies of responses (Evans & Graham, 1980; Machado, 1989; Page & Neuringer, 1985; Stokes, 1995). If eight unique responses are possible,  $i = 1$  to 8, then the *U* value is given by:

$$U_{i=1 \text{ to } 8} = - [\text{RF}_i * \log_2(\text{RF}_i)] / \log_2(8)$$

where  $\text{RF}_i$  = relative frequencies (or percentages) of each of the eight responses. *U* values approach 1.0 when relative frequencies approximate one another, as would be expected over the long run from stochastic generation of equiprobable instances, and 0 when a single instance is repeated. *U* values can be applied as well to sequences of responses. Given a series of Ls and Rs, *U* values can be based on sets of three responses, (e.g., LLR, LLL), constructed from a moving window across the entire series. For example, if the emitted sequence had been LLRLR-RRLLRL . . . , then the sets of three would be LLR, LRL, RLR, LRR, and so on. Note that there are eight possible patterns of L and R taken three at a time, and thus, as in the just-described example, the *U* value would be based on eight possibilities. The *U* value thus is an index of the overall equality of members of a set of possible outcomes. Of course, when contingencies reinforce unequal distributions of responses (e.g., 0.25–0.75), *U* values are expected to be lower than when equality is reinforced (0.50–0.50). In the present study, participant-generated *U* values were compared to those from a stochastic model in each of the five phases.

Stochasticity can be evaluated in another way, namely by comparing proportions of participant-generated responses—L and R instances, pairs of instances, triplets, and so forth—to that expected from the stochastic model. In the balls-in-barrel

example, assuming equal numbers of blue and green balls, a stochastic process would yield not only equal relative frequencies of Bs andGs, but also equal pairs, i.e., blue followed by blue (BB) would approach 0.25, as would green followed by green (GG), blue followed by green (BG), and green followed by blue (GB). Similarly relative frequencies of BBB, BBG, BGB, etc., would each approach 0.125. More precisely, a stochastic process is indicated if the relative frequencies of pairs of instances is the multiplicative value of the first-order proportions (e.g., if  $B = 0.50$  and  $G = 0.50$ , then  $BB = 0.50 * 0.50 = 0.25$ ), and the same is true for GG, BG, and GB. Similarly,  $BBB = 0.50 * 0.50 * 0.50 = 0.125$ , and so on. When baseline proportions are unequal (e.g., if three Gs were selected for each B), then if the process were stochastic, the probability of BB would be  $0.25 * 0.25 = 0.0625$ , of GG,  $0.75 * 0.75 = 0.5625$ , and BG and GB =  $0.25 * 0.75 = 0.1875$ . The same holds for triplets. Thus, when first-order relative frequencies are unequal, a test for stochasticity involves predicting second order, third order, etc., proportions from the observed first-order values. The main questions asked in the present experiment were whether human participants would distribute their responses across two keys in a way predicted from a stochastic model, and whether they could learn to do so rapidly.

## Results

We first tested consistency with matching theory. Figure 24.1 plots, on log-log coordinates, the ratio of left to right responses (total Ls in a phase divided by total Rs in the phase) as a function of the ratio of left to right reinforcers, as commonly done in testing for matching. Each data point represents one participant under one phase, with all participants and all phases represented. Consistent with matching theory, response ratios were related to reinforcement ratios by a power function, with the least-squares best fitting line accounting for 68.6% of the variance. As is often the case in choice experiments, "under-matching" was observed (i.e., response distributions tended to be closer to 0.50-0.50 than the proportions of reinforcements), this indicated by the 0.583 value of the exponent. In the present case, undermatching may partly be explained by the fact that relatively few responses were collected in each phase (approximately 150), that all of these responses were

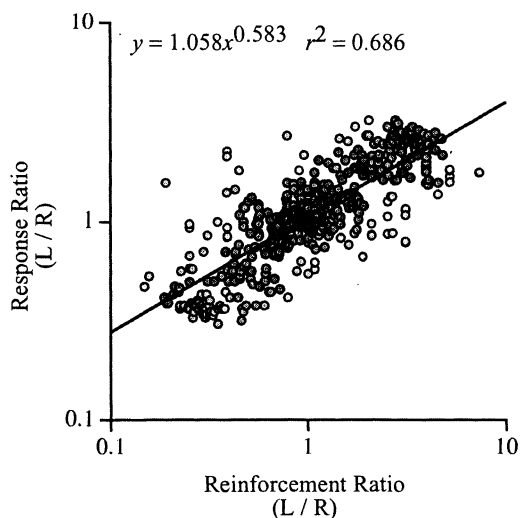


Figure 24.1. Log ratios of left to right responses (L/R) as a function of log ratios of obtained numbers of reinforcements (L/R) for all Experimental participants ( $n = 116$ ) in each of the five phases of Experiment 1. The least-squares best-fitting power function accounts for 68.6 percent of the variance.

included in the analyses, and that participants could, within a given phase, switch preferences from one key to the other. To a first order of approximation, however, the distributions of choices were consistent with predictions from matching theory.

Levels of variability were also influenced by the contingencies, this shown by the U values in figure 24.2. U values were calculated for each subject in each of the five phases based on proportions of response triplets, with group averages represented in figure 24.2. Experimental (filled circles) and Yoked participants (open circles) differed significantly,  $F(1, 144) = 96.119$ , and phase and interaction effects were also significant,  $F(4, 576) = 5.747$  and  $F(4, 576) = 37.278$ , respectively. Reinforcement contingencies therefore clearly influenced levels of response variability.

For comparison, figure 24.2 also shows performance of the stochastic model (Xs), programmed to generate L and R responses with the target probabilities and for the same number of responses as the human participants. Experimental participants' U values approximated those of the stochastic model, although the participants' U values were higher than the model's in Phases 1 and 5 (i.e., human participants distributed their responses more equally) and slightly lower in Phase 4.

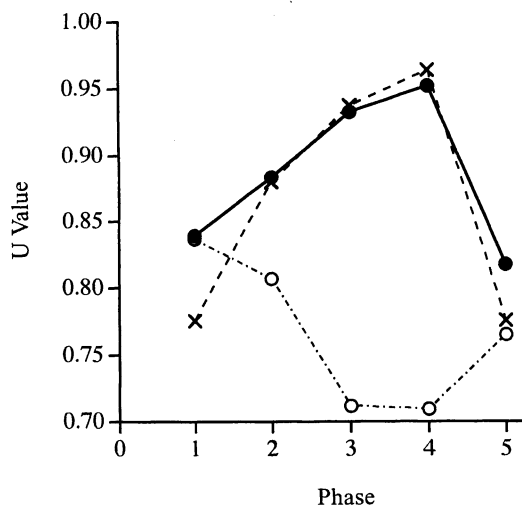


Figure 24.2. Filled circles show average U values for the Experimental participants, reinforced for responding with differing levels of stochastic allocations in the 5 phases of Experiment 1. Phase 1 required 0.75–0.25 distributions; Phase 2, 0.67–0.33; Phase 3, 0.60–0.40; Phase 4, 0.50–0.50; and Phase 5 was a replication of Phase 1. Xs represent a random model, programmed to respond with the five appropriate levels of stochastic allocation, and open circles represent Yoked control participants, whose reinforcement was presented independently of levels of variability.

Finer-grained analyses of pairs and triplets were also consistent with stochastic generation. Because we did not specify which key, L or R, had to be responded to more frequently and to facilitate the comparison between people and stochastic model, these analyses are based on Stays (Ss) and Changes (Cs)—the predictions for these being independent of whether L was preferred or R. A *Stay* was defined as two consecutive responses on the same key, LL or RR; a *Change* was defined as an alternation, LR or RL. Stochasticity was tested by ascertaining if S and C pairs (an S followed by another S, an S followed by a C, and so on) and triplets (SSS, SSC, and so on) could be predicted from the level 1 proportions of Ss and Cs, this being done for each participant individually and in exactly the same manner as described above in the B and G colored balls example. As described there, when response frequencies (or, in this case, Stay and Change frequencies) are unequal, stochasticity can be tested by calculating expected percentages of pairs and triplets based on level 1 frequencies. For example, if a participant emits relative frequencies of 0.90 Cs and 0.10 Ss, then if Cs and Ss were generated by a stochastic process, CC would equal 0.81, CCC, 0.729, and so on. We therefore used each participant's level 1 relative frequencies to predict that subject's pair (level 2) and triplet (level 3) proportions.

The results, shown in table 24.1, are based on the average of all Experimental participants across

Table 24.1 Relative frequencies of Stays (S) and Changes (C) given in five phases of Experiment 1

Type	Phase 1		Phase 2		Phase 3		Phase 4		Phase 5	
	Observed	Predicted	Observed	Predicted	Observed	Predicted	Observed	Predicted	Observed	Predicted
S	.671	.750	.641	.670	.577	.600	.553	.500	.660	.750
C	.329	.250	.358	.330	.424	.400	.447	.500	.340	.250
SS	.485	.451	.424	.412	.336	.333	.320	.306	.464	.436
SC	.186	.221	.218	.230	.241	.244	.233	.247	.196	.224
CS	.189	.221	.216	.230	.239	.244	.233	.247	.196	.224
CC	.141	.108	.142	.128	.184	.179	.214	.200	.144	.116
SSS	.372	.302	.294	.264	.204	.192	.191	.170	.346	.288
SSC	.112	.148	.130	.148	.131	.141	.130	.137	.118	.148
SCS	.115	.148	.150	.148	.169	.141	.154	.137	.122	.148
SCC	.072	.073	.067	.082	.072	.103	.079	.110	.074	.076
CSS	.113	.148	.128	.148	.129	.141	.130	.137	.117	.148
CSC	.075	.073	.089	.082	.111	.103	.103	.110	.079	.076
CCS	.072	.073	.067	.082	.072	.103	.078	.110	.075	.076
CCC	.069	.036	.075	.046	.113	.076	.136	.089	.069	.039

Note: The observed level-1 values (S and C) were used to generate the predicted level-2 (S, SC . . .) and level-3 (SSS, SSC . . .) values. The reinforced response distributions in the five phases were .75–.25, .67–.33, .60–.40, .50–.50, and .75–.25, respectively.

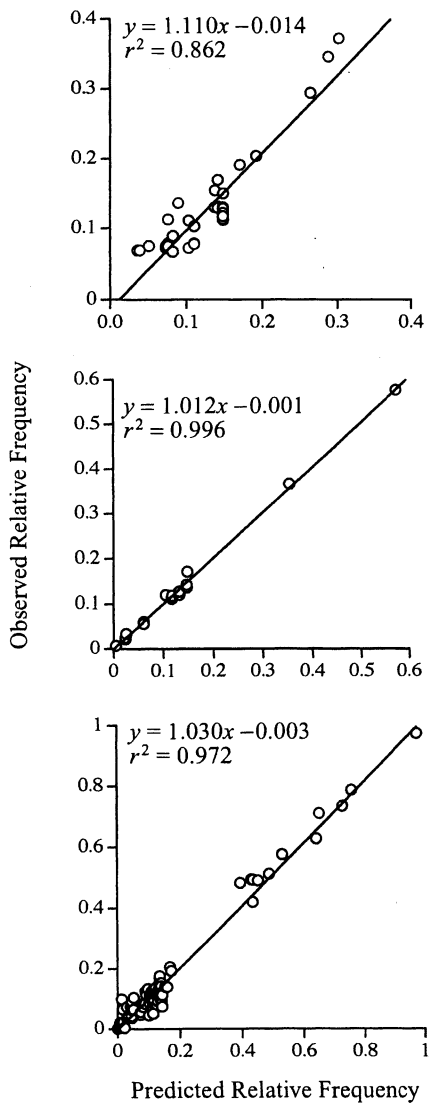


Figure 24.3. Percentages of Stays (S) and Change (C) triplets (SSS, SSC, SCS . . .) as a function of the predicted percentages based on a stochastic model. *Top*, Experiment 1 performances in which different distributions of stochastic responses were required for human participants responding across two keys. *Middle*, Data from Experiment 2a in which pigeons were reinforced for stochastic responding across different number of response operanda. *Bottom*, Human participants responding under the same conditions as Experiment 2a. To the extent that the least squares best fitting function has a slope of 1.0 and intercept of 0.0, the performances were predicted by a stochastic model.

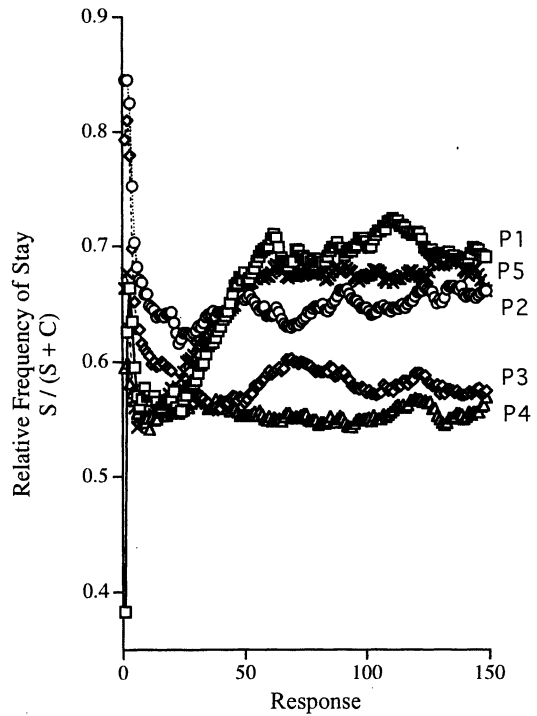


Figure 24.4. Relative frequencies of Stays (S/S + C) across each of the 150 responses in each of five phases, when Phase 1 (P1) required 0.75–0.25 distributions, P2 required 0.67–0.33, P3 required 0.60–0.40, P4 required 0.50–0.50, and P5 was a replication of P1. Each data point is an average across a moving window of 20 responses and across all Experimental participants.

the five phases, but these well represent individual performances. The participants' pairs and triplets values closely approximate those predicted from a stochastic model. A graphic representation of the goodness of the triplets' predictions is shown in figure 24.3, top: Observed percentages of triplets is plotted as a function of the values predicted by the stochastic model. The slope of the least-squares best-fitting line is close to 1, the intercept is close to 0, and the function accounts for 86% of the variance. Thus, the participants' responses—and levels of variability—were consistent with those generated by a stochastic model (table 24.1 and figure 24.3).

Furthermore, variability changed rapidly when the reinforcement contingencies were changed. Figure 24.4 shows probabilities of Stays (S/S + C) over a moving window of 20 responses in each of the five phases. The leftmost points in each phase



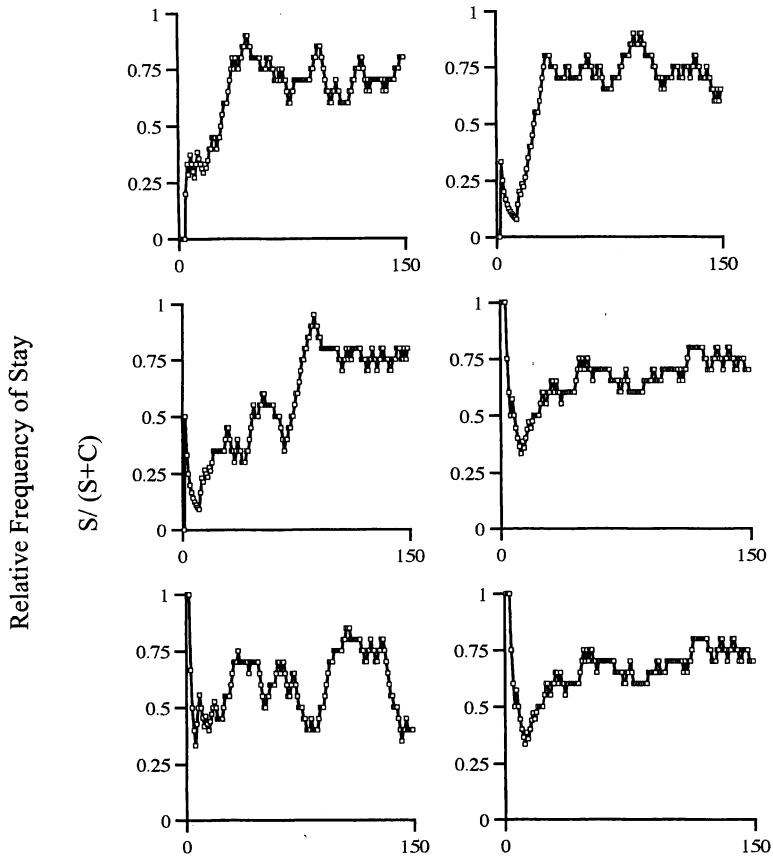


Figure 24.5. Relative frequencies of Stays ( $S/S + C$ ) in Component 1 when the required distribution was 0.75–0.25, for six Experimental subjects. Each data point is an average of a moving window of 20 responses.

indicate the average number of Ss and Cs when the second response in a phase had been emitted—two responses being the minimum necessary to define an S or C—the next point, the average over the first three responses, the next, first four, and so on until the 20th response, whereupon a constant 20-response moving window was represented. The distributions approximated asymptotic values generally within 75 responses. Again, these averages well represent individual performances, as seen in figure 24.5, which shows performances for six arbitrarily selected Experimental participants under Phase 1 conditions. These results are all consistent with stochastic allocation of responses and high sensitivity to reinforcement contingencies.

However, the “stochastic-generator hypothesis” can be challenged on at least two grounds. First, many studies have reported that people are unable to respond randomly when requested to do so, for example, to call out heads or tails randomly (Brugger, 1997; Nickerson, 2002). Later we attempt to explain why our results are consistent

with the stochastic hypothesis, whereas this large body of literature is not. But, first we consider a second objection: Namely, that an alternative hypothesis, one consistent with deterministic assumptions, can account for our results, as well as those from other experiments purporting to show stochastic-like responding (Blough, 1966; Machado, 1989; Neuringer, 1986). This alternative is that memory-based or deterministic computational processes can generate variable responses, including those that appear to match a stochastic model. We consider memory-based theories as an introduction to our second experiment.

#### MEMORY-BASED OPERANT VARIABILITY

Memorial processes can produce highly variable responses in a number of different ways. For example, a person can memorize a long list of “random” numbers and use these to generate responses.

Alternatively, an algorithm (e.g., an equation) that yields random-like sequences can be learned. Chaotic algorithms are one such class and these will be described below. Or one might attempt to remember previous responses or sequences and not repeat these, or remember frequencies of responses or sequences and attempt to equalize these. Each of these strategies relies on memory (explicit or implicit) for past responses or sequences to generate highly variable and, indeed, sometimes random-like outputs.

We first show that animals and people sometimes do manifest an “avoid repetition” or other memorial strategy when satisfying operant variability contingencies; and then we will turn to the question of how memory and stochastic hypotheses might be compared. Our claim is that there are multiple ways in which functionally variable responses can be generated, including memory based, and that responding stochastically is one such strategy.

As described, Pryor et al. (1969) reinforced responses for emitting “novel” responses, ones that had not previously been emitted in the same situation. Similar contingencies with human children resulted in novel drawings and block constructions (Holman, Goetz, & Baer, 1977). Other examples of memory-based “do not repeat” strategies include rats learning to avoid previously entered arms of a radial maze (Cook, Brown, & Riley, 1985); and pigeons, rats, or monkeys learning to choose a novel stimulus under a non-matching-to-sample paradigm.

Also, as indicated earlier, iterations of nonlinear dynamical or chaotic equations—one example is the logistic-difference equation—yield random-like outputs under some parameters. One important characteristic of such chaotic algorithms is that each instance is determined by prior instances and that the overall sequence may be “noisy.” In one experimental demonstration, human participants received feedback showing how closely their responses matched that expected from iterations of the logistic-difference function. The logistic-difference function can be represented as:

$$y_n = 4 * y_{n-1} * (1 - y_{n-1})$$

where  $y_n$  represents the value of the current response and  $y_{n-1}$  represents the value of the just-preceding response. The “4” in the equation is a parameter value, namely the value that generates the most chaotic, or “noisy” outcome, and the

process is initiated with some arbitrary value for  $y$  that is less than 1. The interested reader can reiterate the equation, always using the just-obtained value to seed the next iteration, and see that a highly variable series of outputs results. Neuringer and Voss (1993) showed that human participants could learn to generate responses that were increasingly like iterations of the chaotic model.

Memory insufficiencies have been posited to explain why people often fail to respond randomly, e.g., as found in the human random-generation literature (Brugger, 1997; Wagenaar, 1972). It is suggested that participants are unable to recall the number of responses in each category or the numbers of individual sequences and subsequences that, according to these memory-based theories, are necessary to satisfy randomness criteria (Wagenaar, 1972; Weiss, 1964, 1965). One particularly influential theory suggests that an executive monitor assesses the output of an internal response generator and, depending on prior responses and the individual’s criterion of random, inhibits those potential responses that fail to meet the criterion (Baddeley, 1966). As cognitive or memory loads increase, the executive monitor is hypothesized increasingly to fail.

Direct tests of the stochastic nature of operant variability. It is difficult to assess whether a memory-based or stochastic-based processes is involved, because both can generate identical sequences of responses. The iteration of pi is a case in point (the sequence of digits in pi being indistinguishable from random according to most tests), the iterated algorithms used by computer-based random-number generators is another, and chaotic algorithms a third. Each of these algorithms is memory based, in the sense that each instance is determined completely by the prior instance(s) while yielding sequences that cannot easily be distinguished from that of a stochastic source. There are, however, cases for which memory and stochastic theories make different predictions and these provide a way to test underlying processes. For example, memory-based theories predict that response interference—produced, for example, by interposing long pauses—should degrade approximations to a random model (because they degrade memory for prior responses), whereas a stochastic-generator hypothesis predicts no such effect.

As an example, if one were using memory to approximate a random sequence of B and G colored balls, then long pauses between each response, or

any other type of interference, would be expected to degrade performance. But, random selection of Bs and Gs from a large barrel would not be influenced by increasing the time between selections. Thus, if interresponse interference adversely affects the generation of variable responses, then memorial processes are likely to be involved. An absence of such effects would be consistent with an underlying stochastic process.

To test these opposing predictions, Neuringer (1991) reinforced two groups of rats: one for repeating a fixed, LLRR sequence of responses across two levers and the other for varying (under a lag contingency). The underlying assumption was that successful repetitions of LLRR depended in part on memorial processes, implicit or explicit, whereas the variability contingencies could be met by stochastic selection of Ls and Rs. Testing this possibility was the goal of the experiment. Following acquisition of the two types of behaviors to approximately equal levels of proficiency, pauses were interposed between consecutive responses, with pause lengths systematically increased across phases. The results were that LLRR performance was severely degraded as pause lengths increased, but performance under the variability contingencies was not, results consistent with stochastic generation of the variable responses.

In a related study, McElroy and Neuringer (1990) showed that administering alcohol caused performance decrements for a group of rats that was reinforced for repeating an LLRR sequence but had no effect on a variability-reinforced group. In another test, human participants were trained to alternate between chaotic-like sequences—thought to be memory based, as described earlier—and stochastic-like sequences—those that met a number of tests of randomness. When pauses or other interfering events were interposed between responses, only the chaotic-like sequences showed decrements (Neuringer, 2002). The evidence therefore supports a stochastic process being involved in at least some cases of operant variability.

However, there is conflicting evidence. Under human random-number-generation procedures (as described earlier), increasing the “memory load” is found to interfere with ability to vary. Memory load has been manipulated by increasing the number of possible responses in the to-be-varied set. For example, Rath (1966) compared random generation in human participants when the set of responses included the digits 0 through 9 versus

another case in which the responses were the letters A through Z. Greater deviations from a stochastic model were found with the latter task than the former, results interpreted as indicating that, as memory load increases (from 10 to 26), ability to vary decreases. In a more direct test, Wagenaar (1972) found that, as number of response alternatives increased from two to eight, approximations to an equiprobable model decreased (however, see also Towse, 1998). Consistent with these findings, when a competing task was concurrently presented with a random-generation task, decrements in ability to vary were again observed (Towse & Valentine, 1997). Also consistent with a memory hypothesis, Towse (1998) showed that when available responses were presented visually (e.g., the digits 1 through 10), better approximation to a stochastic model was observed than when the participants had to keep these options in memory.

Might these opposing results—memory interference leaves variability generation unaffected in some cases (e.g., Neuringer, 1991, 2002), whereas it adversely affects it in others (e.g., Rath, 1966; Wagenaar, 1972)—indicate that stochastic processes are involved in the one and memorial processes in the other? Almost all of the human random-generation experiments (with the exception of Neuringer, 1986) required relatively few responses (generally 100 total) and feedback was not provided: participants were simply asked to respond randomly. In most operant cases, including the Neuringer (1986) human random generation experiment, tens of thousands of responses were practiced with reinforcement contingent on high variability. A reasonable interpretation is that memorial strategies are invoked given few responses without feedback, whereas a stochastic-based strategy is employed given the long-term demands of reinforcement-of-variation contingencies.

To test the effects of memory load in an operant situation, Page and Neuringer (1985) studied pigeons responding under a lag 3 contingency, in which the current sequence had to differ from each of those in the preceding three trials. Memory load was varied by changing the required number of responses per trial: four, six, or eight responses per trial across different phases of the experiment. It was reasoned that if memory load had an effect, performance would be degraded when the number of responses increased (i.e., eight responses per trial requires subjects to remember more than four responses). The stochastic hypothesis predicts the

opposite result, as demonstrated by the following example. If responses were directed by the toss of a coin, then if trials consisted of only two responses (a small number used for the sake of this example), the probability of one trial repeating the previous trial is 0.25. (There are four 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 consisted of four responses, again directed by coin tosses, then the probability of a repetition by chance is 0.0625, or 1 in 16, and with eight responses per trial, the probability of repetition is .0039, or 1 in 256. So, if subjects used a stochastic process to generate Ls and Rs, performances should be more likely to satisfy a lag 3 contingency as responses per trial increased, but if subjects were trying to remember each of their sequences, then the opposite result might emerge. Results were precisely those predicted by the stochastic hypothesis and inconsistent with a memory strategy: increasing numbers of responses per trial, and therefore memory-load, resulted in an increased probability of meeting the contingencies. Stated differently, eight-response trials resulted in more frequent reinforcement than four-response trials, as would be expected if L and R responses were stochastically generated. These findings are all the more important because presumably the memory and cognitive capacities of pigeons are smaller than those of humans.

Two objections can be raised to the Page and Neuringer interpretation, however. First, because the number of possible sequences increased across the phases—given four responses per trial, there are a total of 16 unique sequences, but with eight responses per trial, there are 256—any “noise” in the generating system would more likely result in reinforcement under the eight than under the four phase (Peter Balsam & Pat Stokes, personal communication). Second, responses were distributed across two keys rather than the many distinct operanda or verbal responses used in the human random generation literature.

Thus, there is disagreement. Memory-based theories predict that as numbers of alternative responses increase, it should be more and more difficult to approximate random outputs, and data from the human random literature support that prediction. A stochastic-generator hypothesis predicts that approximations to random should be at least as readily obtained with many alternatives as with few (i.e., stochastic sequences are as likely when there are eight different colors in the barrel as two), and one

study, involving different numbers of responses per trial, reported data consistent with the stochastic prediction (Page & Neuringer, 1985).

In an attempt to disambiguate these different results, Experiment 2a varied the number of different response operanda (two, four, and eight), as is characteristically done in the human random-generation literature, but within an operant-conditioning context in which pigeons were given long-term practice and were provided with feedback to indicate successful variations. Because the data supported the stochastic hypothesis, a similar experiment was performed with human participants in Experiment 2b.

#### EXPERIMENT 2A: PROCEDURE

Five mature male Racing Homer pigeons were maintained at 85% of their normal body weights and reinforced for pecking small response squares (1 cm) projected on the monitor (33-cm Apple color monitor) of a touchscreen (Carroll Touch Smart-frames) located in a Gerbrands operant conditioning chamber. At a right angle to the touchscreen, a food hopper provided access to food pellet reinforcement. Additional details concerning this apparatus can be found in Vickrey and Neuringer (2000).

The procedure is diagramed at the top of figure 24.6. Each “trial” began with a 1-cm black square projected at the center of the screen—it will be referred to as the trial-initiating square—a response to which resulted in projection of two, four, or eight squares (depending on the condition), each of these 1 cm in size—to be referred to as the choice squares. Figure 24.6, bottom, shows how these choice squares were oriented—in the two-, four-, and eight-choice cases, respectively. (The numbers next to the squares were not shown during the experiment and are provided to facilitate description of the procedure.) A single peck to one of the choice squares resulted in a cue for reinforcement, if the variability contingency had been satisfied; otherwise, the trial-initiating square reappeared to indicate a new trial. The reinforcement cue was a green star, projected at the center of the screen, a single peck to which resulted—with a 0.25 probability—in access to food for 1.2 s, after which the trial-initiating square reappeared. Those pecks to the star that did not produce food (because of the 0.25 probability) led immediately to the reappearance of the trial-initiating square.

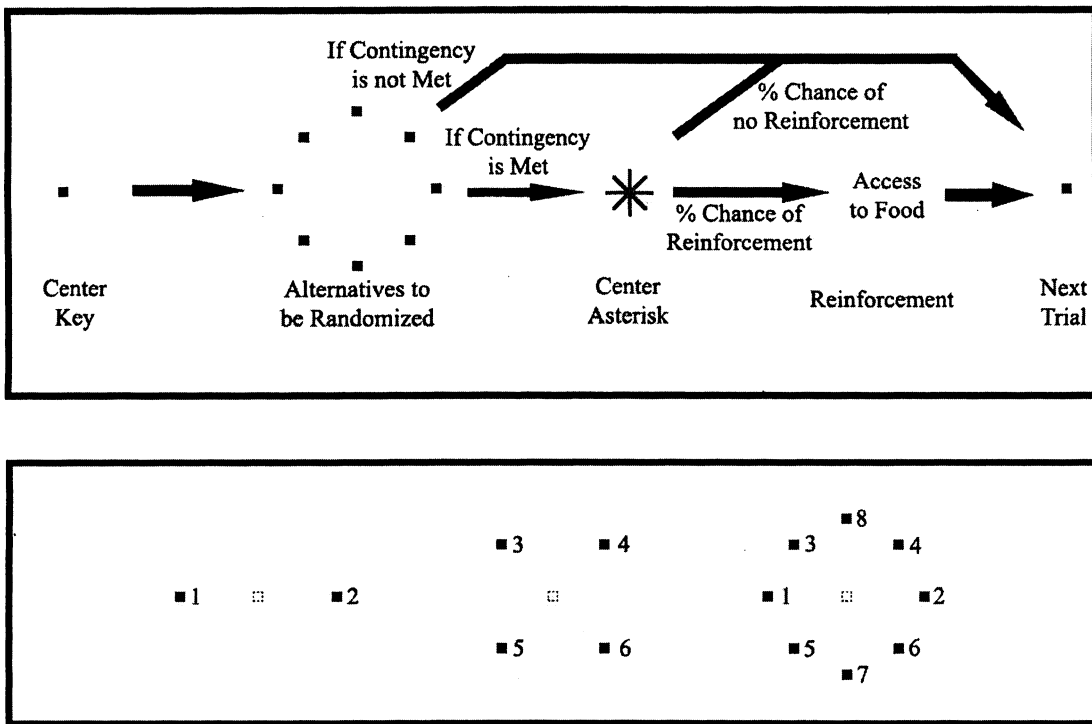


Figure 24.6. *Top*, Outline of the procedure in Experiment 2a. *Bottom*, Orientation of the response squares during the 2-choice condition (*left*), 4-choice condition (*middle*) and 8-choice condition (*right*). The squares are numbered for reference purposes.

Thus, a trial consisted of an initiating response to a center square followed by a response to one of the two-, four-, or eight-choice squares (depending on the condition in effect), and (if the contingency had been satisfied) a peck to a star that led to food 25% of the time and other times to initiation of the next trial. Response variability was reinforced according to the contingencies described later, and the main question was whether levels of variability would differ under the two-, four-, and eight-choice conditions.

Sessions cycled through the two, four, and eight conditions in that order, one session per condition, repeated over and over (two, four, eight, two, four, eight, etc.), with each session terminating after 2,000 responses. Approximately 75 sessions were provided (25 per condition), during the early portion of which parameters were modified to ensure stable responding (e.g., variations were made in the number of trials per session, amount of reinforcement, and probability of food access when the variability contingency was met).

### Variability Contingency

To be reinforced, a response had to complete a sequence—to be defined shortly—that had occurred infrequently, according to a variability contingency related to those used by Blough (1966) and by Denney and Neuringer (1998). For the contingency to be equivalently demanding across the two-, four-, and eight-choice conditions (and therefore avoid the objection to Page and Neuringer's interpretation), the number of possible sequences was kept constant at 64 in the following way. In the two-choice condition, each sequence was defined by 6 consecutive responses, with  $2^6 = 64$  possible sequences. (It is important to note that sequence length was a variable internal to the computer and that no external stimulus indicated length of trial.) As an example, assume that the responses were 1 and 0, and that the pigeon pecked 101000101101, with the rightmost digit indicating the most recent response. The current sequence was defined by the last six responses, or 101101, which (translating from binary to decimal) was

sequence number 45. The just-prior sequence was defined by moving to the left by one, or 010110 = 22, and the same procedure was used throughout the two-choice condition, with each sequence defined by a moving window consisting of 6 responses. In the four-choice case,  $4^3 = 64$  (with instances designated as zero, one, two, and three) and therefore three responses were used to define a "sequence." The eight-choice case involved sets of two responses,  $8^2 = 64$  (with instances designated as zero through seven). The response windows were therefore of different lengths (i.e., lengths of six in the two-choice condition, three in the four-choice condition, and two in eight-choice), in order to keep the amount of information constant (64, or 6 bits). Page and Neuringer held the window size constant (lag 3), while permitting information (number of possible sequences) to vary. It is not possible to hold both constant simultaneously. As will be seen, the consistency of results across these two procedures supports the conclusion that responses were generated stochastically.

Additional details are as follows. Each possible sequence was associated with a counter, for a total of 64 counters per condition. At the beginning of the experiment, all counters were initialized with a value of 20 units. Each response increased the value of its associated counter by 1 unit. In order for all responses to contribute equally, whether emitted early in a condition or later, a constant sum was maintained across all counters by subtracting  $\frac{1}{63}$  from each of the other 63 counters. To meet the variability contingency, the value of a sequence's counter had to be less than 21.6. These values, together with other parameters, were chosen so that a stochastic model would be reinforced on approximately 70% of trials. Each counter was multiplied by an amnesia coefficient (.984) following reinforcement for the same reasons as in Experiment 1. Counter values were maintained across sessions of a given condition (i.e., the values at the beginning of one session were the same as at the end of the previous session under the same condition) but were independent of the other conditions. Additional procedural details are given in Jensen (2003).

The final 18 sessions were used for all analyses, with 6 sessions per condition, 2,000 responses per session, and therefore 12,000 responses per condition. U values were computed for each pigeon, based on the 64 possible sequences in each condition.

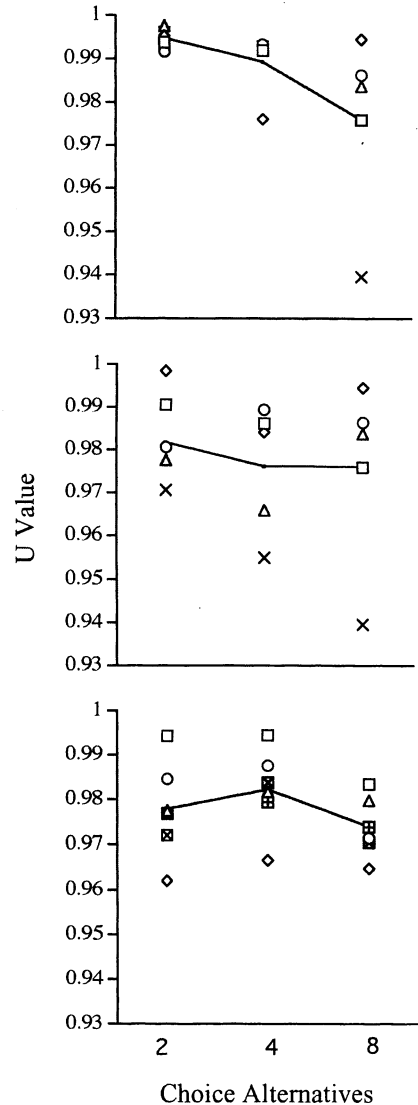


Figure 24.7. U values as a function of number of choice alternatives in Experiment 2a (*top and middle*), where pigeons served as subjects, and Experiment 2b (*bottom*), where human participants were studied. Each subject is represented with the solid line showing the average of all subjects.

## Results

Figure 24.7, top, shows U values for each of the pigeons as a function of number of choice alternatives. The line connects the group mean. Although U values decreased somewhat across conditions, the y-axis is highly magnified (beginning at .93) and a repeated measures analysis of

Table 24.2 Relative frequencies of Stays (S) and Changes (C) given two, four, or eight choice alternatives with pigeons

Type	Observed (2)	Predicted (2)	Observed (4)	Predicted (4)	Observed (8)	Predicted (8)
S	.529	.500	.293	.250	.170	.125
C	.477	.500	.707	.750	.830	.875
SS	.291	.280	.095	.086	.033	.029
SC	.238	.252	.198	.207	.137	.141
CS	.238	.252	.198	.207	.137	.141
CC	.233	.228	.509	.500	.693	.689
SSS	.171	.148	.034	.025	.007	.005
SSC	.128	.132	.061	.061	.026	.024
SCS	.120	.132	.056	.061	.022	.024
SCC	.118	.120	.140	.147	.116	.117
CSS	.128	.132	.061	.061	.026	.024
CSC	.114	.120	.136	.147	.112	.117
CCS	.118	.120	.142	.147	.112	.117
CCC	.120	.109	.366	.353	.577	.572

Note: The observed level-1 values (S and C) were used to generate the predicted level-2 (SS, SC . . .) and level-3 (SSS, SSC . . .) values.

variance (ANOVA) showed that effects were not statistically significant,  $F(2, 5) = 2.30$ .

The stochastic nature of these responses is indicated in table 24.2, which presents percentages of Stays and Changes, as in Experiment 1, together with the percentages of pairs and triplets, for each of the three conditions. As in Experiment 1, Stay was defined as two consecutive responses to the same square (e.g., in the eight-choice case, a response on square 4 followed by another response to 4 was a Stay, whereas a 4 response followed by one to any other square was a Change). In the two-choice condition, a stochastic generator is expected to repeat with a probability of .5, and the same for changes. For the four-choice case, a stochastic generator's expected Stay equals 0.25; and for the eight-choice case, expected Stays equal .125. As shown by the level 1 percentages in the top two lines of the table, the pigeons tended to repeat more than predicted. However, when, as in Experiment 1, percentages of pairs and triplets (SS, SC . . .; SSS, SSC . . .) were calculated from the subjects' own level 1 percentages, they closely matched the predictions from a stochastic model. Figure 24.3, middle, plots the pigeons' percentages of triplets on the y-axis as a function of the predicted values on the x-axis. The fit is almost perfect: Triplet percentages were precisely those predicted by a stochastic model.

In addition to the slight tendency to repeat, indicated by the level 1 percentages in table 24.2, some of the birds showed a bias for or against particular response locations. This bias was especially

noticeable in the eight-choice case where, for example, one subject responded to square 1, the top square, with a relative frequency less than 0.01, whereas the relative frequency of responses to square 6, on the bottom left was 0.22. The expected proportion (given an unbiased distribution) was 0.125 for each case. Other subjects showed similar, although less extreme, biases. We noticed that the smallest birds had the most difficulty responding to the top square (1), and when we measured birds' heights, a significant inverse correlation was found between height and percentage of responses to the top choice location ( $r^2 = 0.909$ , Fisher's  $r$  to  $z$  test ( $p < .04$ )).

Only the eight-choice condition contained the top choice location and that might have contributed to the marginally lower U values in the eight-choice case. To evaluate this possibility, we replicated the two- and four-choice conditions, but with different square locations from those in the initial phase of the study. In particular, we compared levels of variability in the two-choice condition when the locations were one and two; three and six; four and five; and seven and eight, respectively (shown in figure 24.6, bottom). We compared the four-choice conditions when the locations were three, four, five, and six versus when they were one, eight, two, and seven. Figure 24.8 indicates that locations in fact influenced levels of variability. Indeed, location exerted a much larger effect on U value than did number of choice alternatives. Figure 24.7, middle, shows the U values averaged across these repeated conditions together

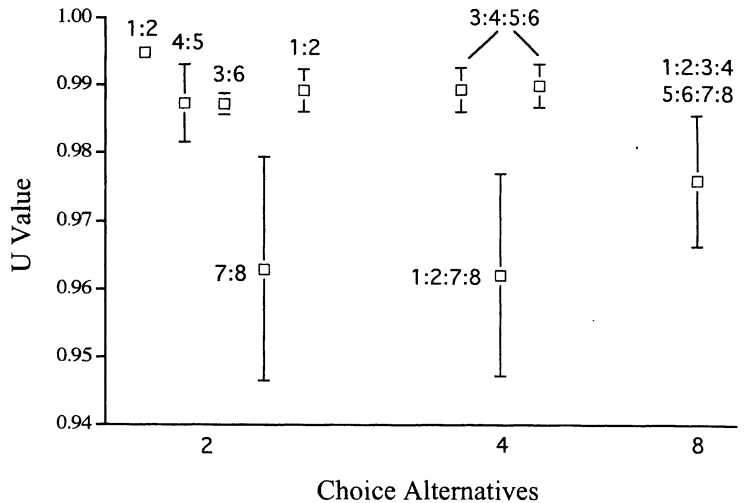


Figure 24.8. U values as a function of number of choice alternatives (two, four, or eight), when the particular response-square locations differed, these indicated by the numbers (e.g., 1:2; 4:5) above or to the left of the individual points. Bars show standard errors.

with the eight choices for comparison and clearly indicates that number of response alternatives had no significant impact on levels of variability. To the extent that number of choice alternatives appeared to affect response variability, biases for and against particular locations were responsible.

Figure 24.9, top, shows the one statistically significant effect of number of alternatives: namely the time between pecks to the trial-initiating center key and the response to one of the available choice alternatives,  $F(2, 8) = 14.26$ ,  $p < .003$ . More time was required to choose among eight alternatives than for four, and likewise for four than for two. This result is consistent with that reported by Baddeley (1966) for human participants.

The main finding was that number of operanda did not significantly influence response variability. These results differ from human random-generation experiments and support a stochastic-generator theory of operant variability, as did Page and Neuringer's earlier findings. However, both Page and Neuringer's study and the experiment just described studied pigeons, whereas the random generation literature is based exclusively on human participants. To test for species differences, we repeated Experiment 2a with human participants.

## EXPERIMENT 2B: PROCEDURE

Unless otherwise specified, the procedure was identical to that in Experiment 2a. Each of six college students was paid \$8 per hour for 5 hours of participation plus the possibility of two additional in-

centives. If the participant's lowest U value (in the two, four, or eight condition) was higher than the lowest U value of the pigeons in Experiment 2a, then an additional \$25 was awarded; and the participant whose lowest U value was higher than all other human participants' lowest values received an additional \$50.

Participants responded on the numeric keypad of an e-Mac computer. Keys 4 and 6 were used in the two-choice case; keys 1, 3, 7, and 9 in the four-choice condition; and keys 1 through 9, except 5, in the eight-choice condition. A visual representation of the active keys was shown on the screen, with the key's image illuminated after each response. The procedure differed from Experiment 2a in that center-key presses and responses to a star were not required, and participants were rewarded with points rather than food. Each session was divided into "blocks" consisting of 150 responses in two-, four-, or eight-choice conditions, with order of conditions randomized such that every set of three blocks contained each of the conditions. Additional feedback was provided at the end of each block in the form of percentage of reinforced responses, with a graph on the computer's screen showing all blocks of the condition just completed.

## Results

Each participant's final 3,600 responses (a number chosen because all participants emitted at least that number of responses per condition) provided the data for the analyses. Figure 24.7, bottom, shows U values as a function of number of choice alterna-



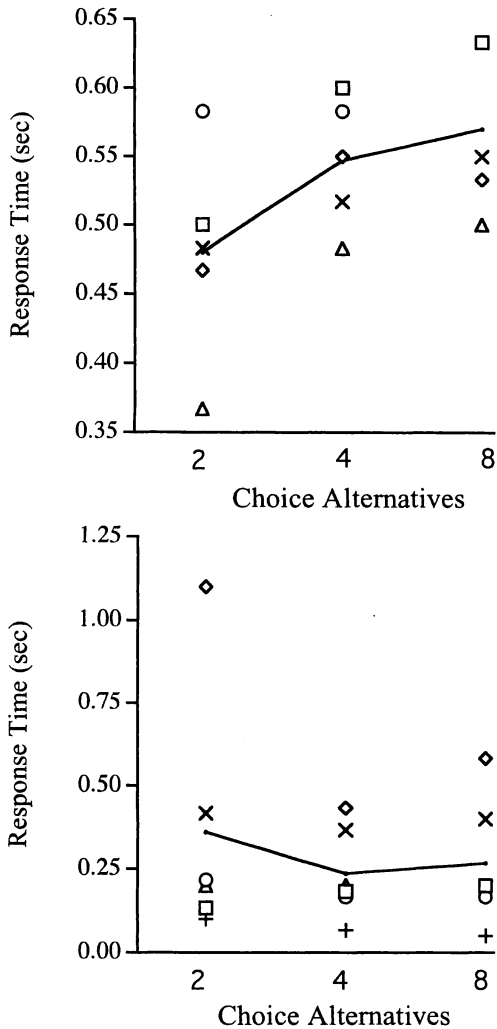


Figure 24.9. Average interresponse time under the 2-, 4- and 8-choice conditions of Experiments 2a (top) and 2b (bottom).

tives. A repeated-measures ANOVA indicated a significant effect of conditions,  $F(2, 6) = 6.26, p < .02$ , with the four-choice condition differing from the eight-choice (Fisher's PLSD). The two-alternative condition did not differ significantly from either of the other two conditions. Participants succeeded at achieving high levels of variability, and three of the six participants had higher lowest scores than the pigeons in Experiment 2a. Thus, although a significant difference emerged, the difference was small and a consistent decrease in variability across two-, four-, and eight-choice conditions was not observed.

Table 24.3 shows relative frequencies of Stays

and Changes, as in Experiment 2a. The human results were comparable to the pigeons' with percentages close to those predicted from a stochastic model. Figure 24.3, bottom, shows the high correlation between stochastic-based predictions and observed triplet values.

Figure 24.9, bottom, shows the mean of all participants' median response times. Apart from a higher mean in the two-choice condition (largely due to one participant responding especially slowly in that condition), no pattern emerged across conditions, and a repeated-measures ANOVA showed no significant effects. The difference between these findings and those from pigeons, where response latencies increased with number of choice alternatives, may partly be due to the location biases demonstrated by the pigeons, which were not seen with the human participants. Also, the procedures differed somewhat. In the pigeon case, latencies were measured from a trial initiating peck; in the human case, responding to the choice alternatives was continuous—there was no trial initiator in order to minimize the tedium of the experiment—and the datum therefore was average interresponse time rather than latency. That the response time data differ also from those previously reported in the human randomness literature (e.g., Baddeley, 1966) may be due to the fact that, in the present case, many thousands of responses were emitted, under control of reinforcing feedback, whereas, as noted above, few responses and the absence of feedback characterize the previous human random-generation research.

The main finding was that, for both pigeons and people, levels of variability were not consistently affected by number of choice alternatives, a result predicted by stochastic theory. We conclude that there are multiple sources of operant variability. We know that people can memorize random sequences (Ericsson & Chase, 1982) and can respond in chaotic-like ways (Neuringer & Voss, 1993). Furthermore, when people are asked to respond randomly, but few responses are required and feedback is not provided, memory for instances and subsequences may well be involved. However, when reinforcement is contingent on approximations to a stochastic model and extended practice is provided, the evidence supports a stochastic-generating process (Experiments 2a and 2b), one that is highly sensitive to changing contingencies (Experiment 1).

We turn now to the general question of why ability to respond unpredictably, whether based on memorial or stochastic processes, might be adaptive.

Table 24.3 Relative frequencies of Stays (S) and Changes (C) given two, four, or eight choice alternatives with people

Type	Observed (2)	Predicted (2)	Observed (4)	Predicted (4)	Observed (8)	Predicted (8)
S	.409	.500	.240	.250	.110	.125
C	.501	.500	.760	.750	.890	.875
SS	.257	.167	.060	.058	.018	.012
SC	.233	.205	.177	.182	.093	.098
CS	.233	.205	.177	.182	.093	.098
CC	.277	.251	.585	.577	.797	.792
SSS	.117	.068	.010	.014	.001	.001
SSC	.141	.085	.051	.044	.016	.011
SCS	.116	.085	.074	.044	.030	.011
SCC	.116	.103	.103	.139	.062	.087
CSS	.141	.085	.051	.044	.016	.011
CSC	.092	.103	.128	.139	.076	.087
CCS	.116	.103	.103	.139	.062	.087
CCC	.161	.126	.482	.439	.727	.705

Note: The observed level-1 values (S and C) were used to generate the predicted level-2 (SS, SC . . .) and Level-3 (SSS, SSC . . .) values.

## REASONS TO RESPOND UNPREDICTABLY

### Protection

Driver and Humphries (1988) describe "protean behavior . . . (that) is sufficiently unsystematic in appearance to prevent a reactor predicting in detail the position or actions of the actor" (p. 36). Stochastic-like protean responses have evolved in many species as a means of protection from attack or predation. Examples include the unsystematic zigzag flights of butterflies and similar movements by other species—mosquitoes, stickleback fish, ptarmigan, squirrels, rabbits, antelopes—in response to the threat of attack by a predator. Mobbing behaviors are similarly protean in nature, such as unpredictable attacks by gulls against a potential predator or the aerial mobbing by hawks of starlings. Driver and Humphries note that, besides providing immediate protection, protean behaviors interfere with the ability of an opponent to learn to anticipate antipredation responses. They also note that protean behavior is "not so random as to be formless; it is a structured system within which predictability is reduced to a minimum" (p. 157). This point parallels one made above: Selection pressures, whether phylogenetic or ontogenetic, help to establish the set of possibly functional responses from which instances emerge stochastically.

### Attraction

Habituation is basic: Repeated stimuli tend to be ignored and unexpected variations attract attention. "Variations attract" describes mating preferences in some species, such as songbirds (Catchpole & Slater, 1995). Female mockingbirds, for example, prefer males who sing complex songs; female sparrows display sexually more to song variety than stereotypy; and great tits demonstrate sexual interest in males with the largest song repertoires. Implied by these studies is that females can discriminate levels of stochasticity (or entropy); evidence supporting this conjecture comes from a series of studies showing that pigeons and people can discriminate among levels of entropy in visual displays (Young & Wasserman, 2001). That males respond to female preferences for variability was shown by Searcy and Yasukawa (1990), who observed that when male red-winged blackbirds were presented with a female dummy, song variability increased. Evidence of direct sensitivity of song complexity to reinforcement contingencies was shown by Manabe, Staddon, and Cleaveland (1997): The variability of budgerigar songs was directly reinforced with food under a lag schedule and song complexity was sensitive to the value of the lag. Thus, attention and attraction by conspecifics may reinforce variations.

## Competition

When animals compete for such resources as food, shelter, and mates, predictable responding may be disadvantageous, as a competitor might thereby take countermeasures. As indicated above, game theory shows that, in such competitive situations, optimization of reinforcement may depend on stochastic allocation of choices; in particular, the Nash equilibrium predicts that animals should match relative allocations of stochastic responses to relative frequencies of reward.

## Exploration

We use exploration in its generic sense to imply the exploration of a problem space—geographic, artistic, scientific, or personal. The goal may be to discover new resources, as in exploration of spatial locations or new lands (Peterson, 1996; Viswanathan et al., 1996), or the discovery of solutions to some problem, as in scientific exploration or solving mechanical problems (Beveridge, 1950; Maier, 1933), or exploration with an aesthetic goal, as in artistic and literary creativity (Campbell, 1960). Stochastic responding within circumscribed limits may be functional in these cases, because it avoids overreliance on previous patterns of response that may no longer be effective. As with all of the other cases described in this chapter, it is important for a functional set to be circumscribed or defined, a point emphasized by Stokes with respect to creativity (Stokes, 2001; Stokes & Harrison, 2002).

## Knowledge and Skill Acquisition

Choosing varied over repetitive stimulation (Fiske & Maddi, 1961) enables acquisition of knowledge. Variations in behaviors and strategies also facilitate the acquisition of cognitive and motor skills. For example, Siegler (1996) has shown that children who vary their strategies are most successful in acquiring mathematical skills. Varying practice routines facilitates acquisition of motor skills (Manoel & Connolly, 1997; Mechner, 1992; Schmidt & Lee, 1999). And, acquisition of difficult-to-learn operant sequences is facilitated by reinforcement of response variations (Neuringer et al., 2000; Seymour, 2003).

## Volition

The operant may be conceptualized as a class of responses, with instances emerging stochastically (Skinner, 1974). In an example of rat's lever-pressing given above, the class may comprise pressing with left paw, right paw, or mouth; with high or low force; and so on. Discriminative cues may increase the likelihood of an operant class, but the particular instances cannot be predicted, either in terms of time of occurrence or topography; the individual responses emerge stochastically. This view of the operant provides a model of voluntary action because voluntary actions are both functional—they are goal-directed actions that can be explained, at least in part, by reinforcement—and potentially unpredictable or stochastic (see Neuringer, 2002). Operant variability manifests both of these characteristics—functionality and stochasticity—and therefore may play an important role any explanation of voluntary action.

In each of these cases—protection, attraction, competition, exploration, acquisition, and volition—controlled variability appears to be functional. In some instances, variability is a species-typical, evolved response to a stimulating situation (e.g., varied bird songs). In other cases, variability is not selected by evolutionary pressures but rather by reinforcing consequences experienced by the individual organism. Thus, many normal, ongoing activities involve controlled, stochastic-like emission of functional responses—for both phylogenetic and ontogenic reasons. Absence of such controlled variability may characterize some psychopathologies. For example, levels of variability in those diagnosed with depression tend to be lower than in nondepressed individuals (Channon & Baker, 1996; Hopkinson & Neuringer, 2003), and the same is true for individuals with autism (Baron-Cohen, 1992; Lee, McComas, & Jawor, 2002; Miller & Neuringer, 2000). Individuals with attention deficit-hyperactivity disorder may manifest the opposite, that is, abnormally high (and uncontrolled) levels of variability.

We again emphasize the importance of controlled variability, because “doing anything” may have low probability of success or be deleterious. By “controlled,” we mean that operant reinforcement and phylogenetic survival shape the set from which possible responses emerge, establish when and

where variations are adaptive, establish within-class probability distributions, and, no doubt, determine when it is adaptive to use memorial processes to behave variably and when to use stochastic ones. Operant stochasticity therefore combines two views: one in which behavior is determined by genes and experiences and the other in which some behaviors are unpredictable or indeterminate, even assuming high-quality knowledge of prior experiences and genetic contributions. There is determination of response classes and potential indetermination of within-class instances. Stated differently, operant variability is a process of stochastic-like emissions from a defined set of possible instances.

### Why Stochastic?

But why is the emission process stochastic, at least under some circumstances? At present, we can only speculate. Stochastic variability may be adaptive in protection, attraction, and competition for reasons indicated by game theory: namely, to counter prediction by another animal or person. Stochastic variability maximizes unpredictability. Stochastic variability may be generated in nonsocial cases as a way to avoid behavioral traps and to produce reinforcing effects. For example, in human creativity, stochastic behaviors may result in effects that are surprising to the creator himself or herself, and therefore are autoreinforcing. Additionally, randomizing heuristics may have evolved because memory for instances and sequences, as well as iterative computations based on chaotic-like functions, require more "computational power" than stochastic emission.

*Acknowledgments* The research and preparation of this manuscript were in part supported by National Institutes of Health Grant MH068259. Parts of the research were submitted as senior theses to Reed College by G. J and C. M.

### References

- Baddeley, A. D. (1966). The capacity for generating information by randomization. *Quarterly Journal of Experimental Psychology*, *18*, 119–129.
- Barba, L. S., & Hunziker, M. H. (2002). Variabilidade comportamental produzida por dois

esquemas de reforçamento (Behavioral variability produced by two reinforcement schedules). *Acta Comportamentalia*, *10*, 5–22.

Baron-Cohen, S. (1992). Out of sight or out of mind? Another look at deception in autism. *Journal of Child Psychology and Psychiatry*, *33*, 1141–1155.

Beveridge, W. I. B. (1950). *The art of scientific investigation*. New York: Vintage Books.

Blough, D. S. (1966). The reinforcement of least frequent interresponse times. *Journal of the Experimental Analysis of Behavior*, *9*, 581–591.

Brugger, P. (1997). Variables that influence the generation of random sequences: An update. *Perceptual and Motor Skills*, *84*, 627–661.

Campbell, D. T. (1960). Blind variation and selective retention in creative thought as in other knowledge processes. *Psychological Review*, *67*, 380–400.

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

Channon, S., & Baker, J. E. (1996). Depression and problem-solving performance on a fault-diagnosis task. *Applied Cognitive Psychology*, *10*, 327–336.

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.

Cook, R. G., Brown, M. F., & Riley, D. A. (1985). Flexible memory processing by rats: use of prospective and retrospective information in the radial maze. *Journal of Experimental Psychology: Animal Behavior Processes*, *11*, 453–469.

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 and Behavior*, *26*, 154–162.

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

Ericsson, K. A., & Chase, W. G. (1982). Exceptional memory. *American Scientist*, *70*, 607–615.

Evans, F. J., & Graham, C. (1980). Subjective random number generation and attention deployment during acquisition and overlearning of a motor skill. *Bulletin of the Psychonomic Society*, *15*, 391–394.

Fiske, D. W., & Maddi, S. R. (1961). *Functions of varied experience*. Homewood, IL: Dorsey Press.

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

- Grunow, A., & Neuringer, A. (2002). Learning to vary and varying to learn. *Psychonomic Bulletin & Review*, 9, 250–258.
- Herrnstein, R. J. (1997). *The matching law*. Cambridge, MA: Harvard University Press.
- Holman, J., Goetz, E. M., & Baer, D. M. (1977). The training of creativity as an operant and an examination of its generalization characteristics. In B. Etzel, J. LeBland, & D. Baer (Eds.), *New development in behavior research: Theory, method and application* (pp. 441–471). Hillsdale, NJ: Erlbaum.
- Hopkinson, J., & Neuringer, A. (2003). Modifying behavioral variability in moderately depressed students. *Behavior Modification*, 27, 251–264.
- Jensen, G. (2003). *Mechanisms and variables involved in aptitude for random generation*. Unpublished undergraduate thesis, Reed College.
- 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.
- Machado, A. (1989). Operant conditioning of behavioral variability using a percentile reinforcement schedule. *Journal of the Experimental Analysis of Behavior*, 52, 155–166.
- 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.
- Maier, N. R. F. (1933). An aspect of human reasoning. *British Journal of Psychology*, 24, 144–155.
- 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.
- Manoel, E. J., & Connolly, K. J. (1997). Variability and stability in the development of skilled actions (pp. 286–318). In K. J. Connolly & H. Forssberg (Eds.), *Neurophysiology and neuropsychology of motor development*. London: MacKeith Press.
- McElroy, E., & Neuringer, A. (1990). Effects of alcohol on reinforced repetitions and reinforced variations in rats. *Psychopharmacology*, 102, 49–55.
- Mechner, F. (1992). *Learning and practicing skilled performance*. New York: Behavior Science Applications.
- Miller, C. (2003). *Metavariability: Interrelations with creativity, depression, and SAT/GPA*. Unpublished undergraduate thesis, Reed College.
- Miller, N., & Neuringer, A. (2000). Reinforcing variability in adolescents with autism. *Journal of Applied Behavior Analysis*, 33, 151–165.
- Mook, D. M., Jeffrey, J., & Neuringer, A. (1993). Spontaneously hypertensive rats (SHR) readily learn to vary but not to repeat instrumental responses. *Behavioral and Neural Biology*, 59, 126–135.
- Neuringer, A. (1986). Can people behave “randomly?”: The role of feedback. *Journal of Experimental Psychology: General*, 115, 62–75.
- Neuringer, A. (1991). Operant variability and repetition as functions of interresponse time. *Journal of Experimental Psychology: Animal Behavior Processes*, 17, 3–12.
- Neuringer, A. (1993). Reinforced variation and selection. *Animal Learning & Behavior*, 21, 83–91.
- Neuringer, A. (2002). Operant variability: evidence, functions, and theory. *Psychonomic Bulletin & Review*, 9, 672–705.
- Neuringer, A. (2003). Creativity and reinforced variability. In K. A. Lattal & P. N. Chase (Eds.), *Behavior theory and philosophy* (pp. 323–338). New York: Kluwer Academic/Plenum.
- Neuringer, A. (2004). Reinforced variability in animals and people. *American Psychologist*, 59, 891–906.
- Neuringer, A., Deiss, C., & Olson, G. (2000). Reinforced variability and operant learning. *Journal of Experimental Psychology: Animal Behavior Processes*, 26, 98–111.
- Neuringer, A., & Huntley, R. (1992). Reinforced variability in rats: Effects of gender, age and contingency. *Physiology & Behavior*, 51, 145–149.
- Neuringer, A., Kornell, N., & Olufs, M. (2001). Stability and variability in extinction. *Journal of Experimental Psychology: Animal Behavior Processes*, 27, 79–94.
- Neuringer, A., & Voss, C. (1993). Approximating chaotic behavior. *Psychological Science*, 4, 113–119.
- Nickerson, R. A. (2002). The production and perception of randomness. *Psychological Review*, 109, 330–357.
- Page, S., & Neuringer, A. (1985). Variability is an operant. *Journal of Experimental Psychology: Animal Behavior Processes*, 11, 429–452.
- Peterson, I. (1996). Trails of the wandering albatross. *Science News*, 150, 104–105.
- 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.
- Rath, G. J. (1966). Randomization by humans. *American Journal of Psychology*, 79, 97–103.
- Ross, C., & Neuringer, A. (2002). Reinforcement of variations and repetitions along three independent response dimensions. *Behavioural Processes*, 57, 199–209.

- Schmidt, R. A., & Lee, T. D. (1999). *Motor control and learning: A behavioral emphasis*. Champaign, IL: Human Kinetics.
- 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.
- Seymour, K. H. (2003, May). The effects of reinforcing operant variability on task acquisition. Presented at the Association for Behavior Analysis, San Francisco.
- Siegler, R. S. (1996). *Emerging minds: The process of change in children's thinking*. New York: Oxford University Press.
- Skinner, B. F. (1974). *About behaviorism*. New York: Knopf.
- Smith, J. M. (1982). *Evolution and the theory of games*. Cambridge: Cambridge University Press.
- Stokes, P. D. (1995). Learned variability. *Animal Learning & Behavior*, 23, 164-176.
- Stokes, P. D. (2001). Variability, constraints, and creativity: Shedding light on Claude Monet. *American Psychologist*, 36, 355-359.
- Stokes, P. D., & Harrison, H. M. (2002). Constraints have different concurrent effects and aftereffects on variability. *Journal of Experimental Psychology: General*, 131, 552-566.
- Towse, J. N. (1998). On random generation and the central executive of working memory. *British Journal of Psychology*, 89, 77-101.
- Towse, J. N., & Valentine, J. D. (1997). Random generation of numbers: A search for underlying processes. *European Journal of Cognitive Psychology*, 9, 381-400.
- Vickrey, C., & Neuringer, A. (2000). Pigeon reaction time, Hick's law, and intelligence. *Psychonomic Bulletin & Review*, 7, 284-291.
- Viswanathan, G. M., Afanasyev, V., Buldyrev, S. V., Murphy, E. J., Prince, P. A., & Stanley, H. E. (1996). Levy flight search patterns of wandering albatrosses. *Nature*, 381, 413-415.
- Wagenaar, W. A. (1972). Generation of random sequences by human subjects: A critical survey of literature. *Psychological Bulletin*, 77, 65-72.
- Weiss, R. L. (1964). On producing random responses. *Psychological Reports*, 14, 931-941.
- Weiss, R. L. (1965). "Variables that influence random-generation": An alternative hypothesis. *Perceptual and Motor Skills*, 20, 307-310.
- Young, M. E., & Wasserman, E. A. (2001). Entropy and variability discrimination. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, 27, 278-293.

COMPARATIVE COGNITION

*Experimental Explorations of Animal Intelligence*

---

*Edited by*

Edward A. Wasserman

Thomas R. Zentall

**OXFORD**  
UNIVERSITY PRESS

2006