Operant Variability and the Power of Reinforcement

Allen Neuringer

Abstract

Reinforcers contingent on response variability exert powerful and precise control over levels of variability, from stereotypy to stochasticity. This paper reviews how variability-contingent reinforcers interact with non-contingent, eliciting events to influence the variability of operant responses. Relationships to stimulus control, choice, acquisition of new responses, voluntary action, autism, and ADHD are discussed.

Keywords: Behavioral Variability, Reinforcement Contingencies, Autism, Attention Deficit Hyperactivity Disorder (ADHD), Operants, Stimulus Control, Volition.

"Since Darwin, the central project of evolutionary biology has been to explain the origin of biodiversity – to determine how novel species and their characteristics have evolved" (Thorton, 2006, p. 157). Operant conditioning can be described in similar terms: Since Skinner, the central project has been to explain how operant behaviors originate and change. To explain biodiversity, on the one hand, and behavioral diversity, on the other, we must consider the variations from which each emerges. This paper is about the causes, consequences, and possible applications of variability, but I begin by noting some parallels in the area of genetics.

Continual variation in genetic material provides the bases of all evolved forms of life. Lewis Thomas said this in a more evocative way: "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). Genetic variability, due to mutations in DNA, have many causes including errors during normal replication, insults from chemicals or radiation, jumps or transpositions of genetic materials, and other "spontaneous" changes. In sexually reproducing organisms, another source of continual variation occurs during gamete formation. 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 are said to occur "randomly," that is, without regard to the current "needs" of an organism or changes that result. However, "random" does not mean without influence or boundaries. The processes that permit and maintain genetic variability have themselves evolved under selection pressures. "Chance favors the prepared genome…Evolutionary strategies evolve, under the pressure of natural selection; this makes the process of evolution more efficient…(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. 1 & 15). A combination of variation and selection at work *within the genome itself* may best be described as *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, similar bounded stochasticity is an attribute of operant behavior as well.

A word about terminology. "Stochastic" and "random" will be used synonymously in the present paper, both to indicate a sequence of instances from a defined set, with prediction of the next instance not possible at a level greater than the relative frequencies of members of the set. Imagine, for example, a well-mixed tub filled with 200 red balls, 200 blue balls, and 200 green balls. Balls are selected one at a time with replacement and continual mixing. Prediction of the next color will be no better than 1/3 on average, (200/600), the relative frequencies of each color. If there were 100 red, 200 blue, and 300 green balls, then prediction of green would be correct no more than ½ the time on average,

(300/600), and so on. Similarly, one can imagine a biased pair of die: Although both dice come up sixes more often than expected (say each does so 1/3 of the time), the outcome of each die roll remains stochastic because having just rolled a six neither raises nor lowers the chances of rolling another six in the future. Although selection may occur stochastically, the emergent instances are limited to members of the set and influenced by their relative proportions or "strengths." As will be shown, there are many types of pressures at work to define and influence the set from which operant instances emerge.

All behaviors vary, of course, as do all things physical. But variability plays an especially important role for operant behaviors. Operants are actions that produce reinforcing changes (both positive and negative), changes for which an animal or person will normally work to gain or to avoid. Informally, operants are behaviors directed at obtaining future goals and are sometimes referred to as instrumental responses or voluntary actions. To be successful, operants must be sensitive to continual changes, both in physical environment and in contingencies of reinforcement, and thus operant responses must continually vary. The variability is often stochastic in nature, but as just noted, within the defining limits of a currently operative set. Note that, under some circumstances, the set can be extremely small, leading to easy predictions of repetitive responses. In other circumstances, however, the set may be immense, consisting of all currently possible members of one's behavioral repertoire, e.g., when asked to do something completely unpredictable (Scriven, 1965).

The ability of operants to vary, depending upon moment-to-moment stimulus and contingency changes, can be contrasted with another basic form of learning, Pavlovian conditioning, where highly constrained, species-typical responses are the rule. For example, when food is contingent upon the presentation of a red light (light then food, repeated over and over), the light will come to elicit in many mammals species-typical conditioned responses (CRs), including salivation and gastric secretion. In Pavlovian conditioning, the stimuli can vary over a wide range, but the CR is highly constrained or determined. In the operant case, both discriminative stimuli and operant responses may vary over wide ranges. If the contingency is modified even slightly, quite different responses can be engendered. The important point here is that an extremely wide variety of operant responses can be conditioned to a consequence such as food whereas these same events, when contingent upon a conditioned *stimulus*, result in a highly stereotyped CR. The "openness" of the operant response, its modifiability, depends to a large extent on the ability of the response to vary in a way that is influenced by contingent reinforcers.

Shaping. Shaping provides a salient example of how the generation of a new operant depends upon an underlying milieu of reinforced variations. For example, Deich, Allan, and Zeigler (1988) measured the opening of a pigeon's beak, referred to as gape size, when the bird was pecking at a response key. When it pecks at food, a pigeon opens its beak, and it does the same when pecking at a key to obtain food. During baseline, when pecks were occasionally reinforced independently of size, gape size was found to vary – these were baseline or "operant level" variations. Such baseline variability is characteristic of all operants. The authors then utilized the variability to selectively reinforce either larger or smaller gape sizes and thereby successfully shaped gape sizes that the birds had never before emitted. Gape sizes continued to vary throughout the shaping process, but the distributions diverged, depending upon the reinforcement contingencies. Shaping depends upon differential selection of *sets of variations.* That is, when an operant "response" is reinforced, changes are seen in a set of variations, or in the size, distribution, or component members of operant "classes."

Levels of variability can change, even after repeated reinforcement of a tightly constrained operant response. One may throw a ball to a friend in a relatively fixed and repetitive way, but it is possible to change the velocity or direction, or throw (if so requested) in an unpredictable manner. The same is true for the words we speak (in terms of word choice, loudness, speed, etc.), the pace of our walk,

and all other operant acts. Identify any common operant response in a normal intact animal or person, and variability is present or possible. Variability may be low in some circumstances, such as answering the phone when it rings, and therefore be predictable, while in other circumstances variability changes continually, for example, when problem solving or creating a work of art, but variability *can* change in all of these circumstances, given a demand for change. The existence of and potential for variation is an essential characteristic of the operant. Evidence for these assertions will be provided below.

Direct reinforcement of operant variability. As with the maintenance of genetic variations, there are multiple sources that engender and maintain variability in the operant, and in both areas these suggest evolutionary selection pressures. In a complex and ever-changing world, such as our own, genes, species, and operant behaviors that vary will be most likely to support survival, success, and procreation. Of course survival also depends upon learning to respond in consistent or repeatable ways to certain environmental conditions, and an organism must therefore learn when to vary, when to repeat, and adaptive levels of variations between these two. In the case of the operant, direct reinforcement of *levels of variability* has been demonstrated, and some of the relevant research will be described.

Donald Blough's (1966) tour-de-force research on the reinforcement of stochastic pecking in pigeons initiated on the study of reinforced variability. Blough required pigeons to peck a response key with a distribution of inter-peck intervals that approximated those that would be expected from a stochastic source, for example, from emissions of electrons as measured by a Geiger counter. To accomplish this, Blough established 16 interval categories and reinforced a peck only if it fell within a category currently containing the fewest peck instances. The durations associated with the categories were defined in such a way that pecks were expected to fall into the 16 categories equally often, as would be expected from an atomic emitter. Although the birds showed a non-random tendency to emit double pecks (two pecks with a very short time between them), in most other ways they were able to satisfy the "random peck" requirement. This provided the first strong indication that animals could explicitly be reinforced for varying their responses and, indeed, levels of variability could approach that of a true random generator. A number of confirming studies followed, and these, in turn, led to experimental analyses.

Karen Pryor's work with porpoises gained much attention when she reported that she was able to reinforce the porpoises for engaging in novel behaviors – flips, swims, turns (Pryor, Haag & O'Reilly, 1969). The result of this reinforcement-of-novel-behavior procedure was to engender types of behaviors that experienced trainers had never before observed in any porpoise. Pryor's work was followed by Goetz and Baer (1973), who rewarded preschool children for block constructions that differed from all that they had previously constructed during one session. As training proceeded, the children constructed increasingly different forms, including ones never before emitted by the child. When reinforcement was later made contingent upon construction of a single form, the children satisfied the contingency and repeatedly constructed one form.

The Pryor and Blough papers suggested that variability is a dimension of behavior similar to other "operant dimensions," such as topography, location, speed, and force. These are all controlled by reinforcers contingent upon particular values of the dimension. Is it possible that variability can be dynamically modified by reinforcers contingent upon particular levels? This is a radical question, since in behavioral analysis, as well as evolutionary theorizing, it was initially believed that processes causing random variations and processes resulting in functional selections were separate and independent, with selection working on "blindly" generated variations.

The issue is sufficiently important that, before we conclude that response variability can directly be reinforced, alternative explanations must be ruled out. The most likely alternative is that variability is somehow elicited by the situation – the way a stumble is elicited by a misplaced rock on the walk – rather than reinforced – the way a jump is reinforced by avoidance of a puddle. The main evidence regarding this possibility comes from comparisons of two conditions, one in which animals were reinforced for varying sequences of responses (VAR) and a comparison condition in which equally frequent reinforcers were presented contingent on responding but independent of variability (YOKE), as will be described in more detail.

Suzanne Page, in her senior thesis as an undergraduate at Reed College, studied pigeons that were reinforced for distributing their pecks across two keys in a way that met a variability contingency, that is, by responding unpredictably (Page & Neuringer, 1985). The experimental session was divided into many trials, each consisting of 8 pecks. A trial ended with food if the variability criterion was met and otherwise with a brief timeout. A Lag contingency was used in which the current pattern of 8 responses across the two keys had to differ from the patterns in some number of previous trials, this number defined by the Lag value. Thus, under a Lag 5 contingency, if the current pattern differed from each of the patterns in the previous 5 trials (within a moving window), then the trial ended in reinforcement; otherwise it ended with a timeout. To give just one example, if the current pattern was LLRLRRRL, with L indicating a left peck and R a right one, and that particular sequence had not be emitted during any of the previous 5 trials, food would be provided. The pigeons readily le arned to meet the schedule requirement, even when the lag value was raised to as high as 50, where the current pattern had to differ from those in each of the previous 50 trials. Additional experiments showed that the way the birds satisfied the Lag contingency was not by "remembering" the previous sequences but rather by responding in a stochastic -like manner, that is, as if they were flipping a coin and pecking L when it landed with heads up and R with tails.

But perhaps the most important part of the Page and Neuringer (19856) study was a comparison phase that followed reinforcement of variability. Here the birds continued to be reinforced at the end of some of the eight-response trials, but now reinforcement did not depend upon the bird's variability. Instead, reinforcements during this phase were yoked to the earlier pattern of reinforcement delivered during the VAR phase. Thus, if an individual bird had been reinforced during the nth trial in a given VAR session, then the analogous trial in the YOKE phase would also end with reinforcement, regardless of whether the lag contingency had been met. By self-yoking reinforcements in the two phases, we assured that frequencies and patterns of reinforcements would be identical for a given bird and that the only difference was that during VAR, variable response sequences were *required* whereas during YOKE, response sequences were *permitted* to vary, but not required.

The question, of course, was whether response variability would be higher during VAR than YOKE and the answer was clearly yes. When directly reinforced for varying (VAR), the birds varied. When reinforced independently of their variability (YOKE), the birds tended to repeat pecks on one or the other key. This demonstrated direct reinforcement of variability together with necessary controls and, indeed, was the first such demonstration in the literature. The conclusion was that variability was, in fact, an operant dimension of behavior, a conclusion supported by much additional evidence, some of which is described below.

Discriminative Stimulus control. Skinner described operants in terms of a three-term relationship: A discriminative stimulus sets the occasion for emission of operant responses that produce, at least occasionally, reinforcing consequences. If variability is an operant dimension, then it too should come under the control of discriminative cues, and a number of experiments have shown this to be the

case. For example, under a multiple schedule of reinforcement, in the presence of one stimulus, rats learned to emit variable four-response sequences across L and R levers, but in the presence of a different stimulus to repeat a single pattern, LLRR (Cohen, Neuringer, & Rhodes, 1990). The stimuli exerted excellent control over varying versus repeating. As part of the experiment, the rats were next injected with doses of ethanol, from control doses (0 gm ethanol per kg), through low (1.25 g/kg and 1.75 g/kg) to high doses (2.25 g/kg). Figure 1 shows that administering the ethanol, so as to make the animals increasingly drunk across these phases, had no detrimental effect on their ability to meet the variability contingency (shown on the left) but severely impacted their ability to repeat LLRR (on the right).
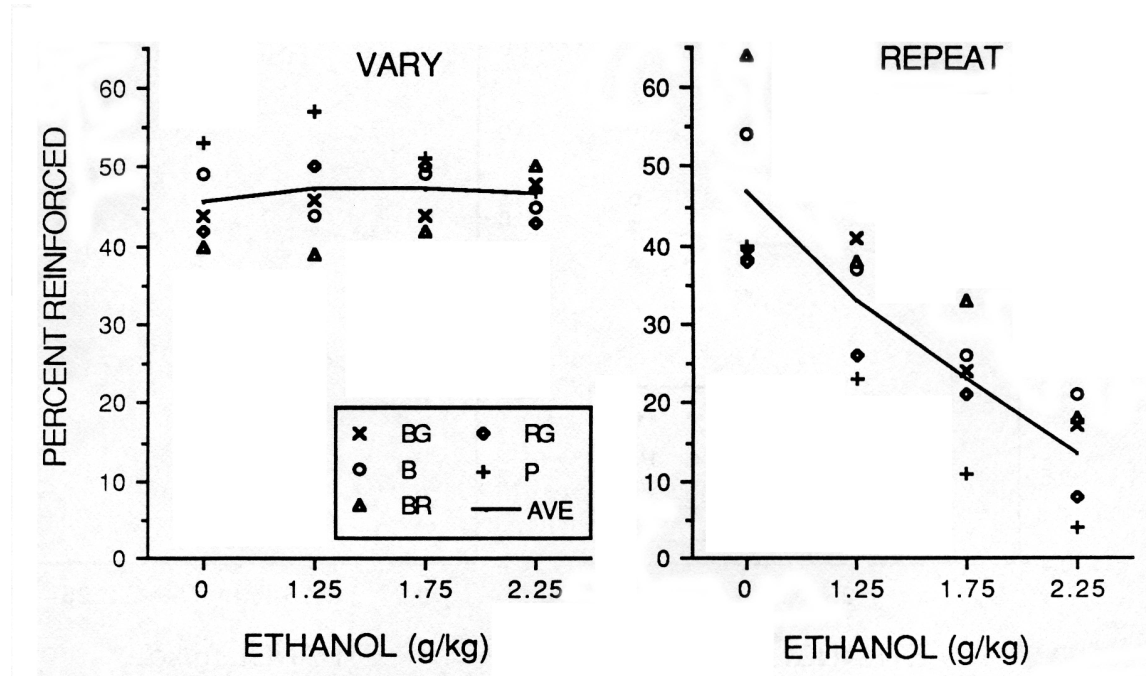

Figure 1. Percentage of trials that met variability contingencies (on left) and repetition contingencies (on right) under a multiple schedule of reinforcement when ethanol was administered via ip injections at the doses shown along the x-axis. The individual symbols represent individual rats and the solid lines connect the means.

Indeed, at the highest dose, the rats were rarely reinforced for repeating since errors were so common. This finding that operant variability is resistant to interference has been reported in other studies (Doughty & Lattal, 2001). In a more stringent test of control by discriminative cues (Denney & Neuringer, 1998), rats were reinforced for varying in one stimulus condition whereas reinforcement was yoked under a comparison stimulus, the two stimuli alternating within each session. Under yoke conditions the rats were free to vary or not since reinforcement did not depend upon levels of variability, but the cues came to exert strong control over variability levels, with significantly higher variability during the VAR stimulus than the YOKE. These results suggest that in one context, an individual may vary, behave "loosely," possibly be original or creative, whereas in a different context, the same individual will respond in a more pedestrian, repetitive, predictable, and unchanging manner. The results also suggest that operant variability may be influenced by drugs and other insults differently than operant repetitions, a conclusion to which we will return.

Stochastic responding. As initially suggested by Blough (1966), research shows that given appropriate reinforcement, levels of variability can approach that of a random, or stochastic, generator.

People, pigeons, and rats have been shown to emit highly variable sequences of responses under a number of different contingencies that reinforce variability and, under the most stringent contingencies, approximate stochastic models (Neuringer, 2002). The results from these studies contrast with a large number of other studies on "random number generation" in which human participants were simply asked to respond randomly (Brugger, 1997; Neuringer, 1986). The almost universal findings from that literature were that merely requesting randomness, without providing feedback about the quality of the behavior, failed to produce random-like behavior. Thus, direct reinforcement appears to enable people to maximally approximate stochastic behavior, but simply asking for such behavior does not suffice. Reinforcement works in a similar manner for non-human animals as well.

Many interesting implications follow from demonstrations of reinforced stochastic responding. Some of these are theoretical, for example, implications for determinism versus "free will" that will be discussed below, and others are practical. For example, in many competitive situations, from games to wars, there are times when truly unpredictable actions – within the limits of a set of adaptive or functional choices – are the best strategy (Maynard Smith, 1982). An example from the *Art of War*: "Any military operation takes deception as its basic quality... [list of ways to deceive and be unpredictable follows]… All the above-mentioned is the key to military victory, but it is never possible to formulate a fixed plan beforehand." The influences of reinforcement contingencies may help us to understand why and when unpredictable behaviors emerge.

Variability continuum. Reinforcers exert more precise control over the variability-of-response dimension than a dichotomous "respond stochastically" versus "respond predictably." To understand such influence, variability must be viewed not as a "thing" or "end point" but rather as a continuum, with repetition and high predictability at one end and stochastic unpredictability on the other. A number of studies have shown that *levels* of variation are controlled by reinforcers contingent upon those levels, again from repetitions to approximations to stochastic distribution. For example, Grunow and Neuringer (2002) reinforced rats for emitting three-response sequences across three different operanda, namely by pressing left (L) and right (R) levers and pushing a center key (K). There are 27 possible patterns, for example, LLL, LLR, LLK, and so on. The rats were divided into four groups that differed in terms of levels of variability required for reinforcement. One group was reinforced for very high levels. A computer kept track of the frequencies of all 27 possible patterns and reinforced a sequence only if its relative frequency fell below a very low threshold level, that is, if the pattern had occurred less than 4% of the time. A truly random generator would be expected to distribute its responses equally across the three operanda such that each possible pattern would occur, over the long run, approximately 3.7 percent of the time (or relative frequency of .037). Thus, the "high variability" group was reinforced for approximating a random generator. The other three groups had less demanding thresholds: 0.055, 0.074, and 0.37, this last being ten times more permissive than the "high variability" group. The results were that the high-variability group responded most variably and the other three groups fell in line, with the .37 group often repeating sequences and responding not very variably at all.

Definition of response classes. Reinforcement does more than control the level of response variability. It also selects the set of responses from which variations emerge. For example, rats were rewarded for emitting four-response sequences that met a Lag variability requirement (Mook & Neuringer, 1994). The rats were initially reinforced for emitting the 16 possible sequences – RRRR, RRRL, RRLR… LLLR, LLLL – approximately equally. In a later phase of the experiment, the contingencies required both that a sequence begin with two right responses, RR, and that it also differ from the just-preceding sequence (Lag 1). Here the appropriate set was RRRR, RRRL, RRLR, and RRLL, and the rats responded appropriately. More informally, in a chamber containing three keys, when variability of responses across only two of the keys is reinforced, responses vary across the two "on" keys

and are not emitted (or very rarely so) on the third "off" key. Thus, reinforcement simultaneously defines the acceptable response set and the level of variations required from within that set.

The control exerted by reinforcers over the variability continuum is even more diverse and powerful than has yet been described. Every operant has many dimensions, for example, in the case of lever-press, there are rates of response, topographies, forces, durations, and so on. Reinforcers may independently and simultaneously influence each of these dimensions – and that is true for variability. In an experiment by Chris Ross (Ross & Neuringer, 2002), individual dimensions were independently and simultaneously reinforced for varying and repeating. Here is the experiment. College students drew rectangles on the screen of a computer in order to gain points. Although they were not informed about the underlying contingencies, one group of students was reinforced for drawing rectangles whose sizes were approximately the same (within a certain "delta" window), trial after trial, while simultaneously varying both the locations of the rectangles on the screen and the forms of the rectangle (whether it was square, or long in the horizontal or vertical direction, etc.). Thus, reinforcement depended on repeating size and simultaneously varying location and shape. Only when the rectangle met all three requirements simultaneously was a point awarded. The participants learned to do what was required, as shown at the left of Figure 2. A second group had to repeat location while varying size and form (center
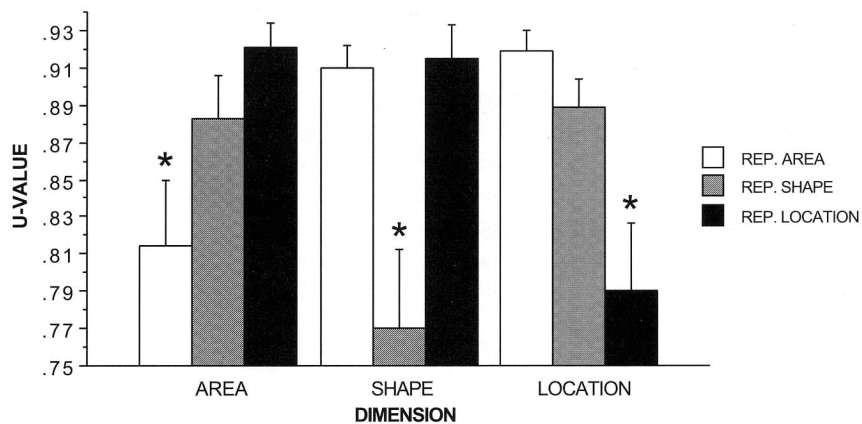

Figure 2. U-value, a measure of response variability – the higher the U, the closer the distribution of responses to a stochastic model—in three groups of human participants. The "area" group (on the left) was reinforced for drawing rectangles on the screen of a computer whose areas repeated (within a delta window) while simultaneously varying the shapes and locations of the rectangles. The "shape" group (center) was required to repeat shape, while varying area and location. The "location" group (right) was required to draw rectangles in approximately the same location, while varying shape and area. Error bars indicate standard errors and stars indicate statistically significant difference from the other two dimensions.

of the figure). A third group to repeat form while varying size and location (right). All three groups significantly improved their rates of reinforcement but, interestingly, many of the participants were unable to describe accurately what they had to do in order to be reinforced. Reinforcers powerfully and precisely controlled the variability and stereotypy of multiple dimensions of a response, and did so concurrently and independently.

The Ross study highlights the power of reinforcement. Operant behaviors are conglomerations of a multiplicity of potentially independent responses, for example, by different parts of the body, and dimensions of responses, for example, force, speed, topography, location. It is astounding to me that a

binary event – reinforced or not – may exert independent control across some unknown number of responses and dimensions. This power of reinforcement must be considered when attempting to explain operant behavior.

Non-contingent effects. To understand when and how responses vary, we must consider both non-contingent and contingent effects of reinforcers. This distinction is important. For example, presenting a treat to a puppy may elicit tail wags, jumping, and generally high activity, this despite the fact that the treat was presented independently of the puppy's behavior or even contingent upon the puppy momentarily sitting quietly. I will discuss three important sources of non-contingent influences on variability: adventitious events, or "accidents"; decreases in reinforcement; and reinforcement uncertainty (see also Lee, Sturmey & Fields, 2007; Neuringer, 2002).

Environmental Accidents. Bandura (1982) described cases of accidental experiences leading to major changes in life's paths, such as happening to sit next to a particular person on an airplane or coming across a particular passage in a book (see also Taleb, 2007). Accidents are as important in science as in everyday life, leading to new experiments and discoveries (Beveridge, 1957). The rich variety of commonly experienced environments provides a continual source of behavior change that can't be predicted prior to the experience itself. That organisms are highly sensitive to such accidents is indicated by the importance of unexpected stimuli—CSs, discriminative stimuli, unconditioned stimuli, and primary reinforcers—in conditioning both Pavlovian and operant responses (Mirenowicz & Schultz, 1994; Rescorla & Wagner, 1972). Depending upon the needs of the moment, one can modify the influence of such environmental accidents. For example, working in a constant environment might help to put good work habits under stimulus control, as was done by B. F. Skinner when he established a particular location and stimulus light as a discriminative stimulus for his writing behavior. Opening a book of famous aphorisms or walking through a novel area of a city works in the opposite direction, increasing the likelihood of new ideas or behaviors.

Decreased Reinforcement. Withholding needed nutrients often causes variability to increase. This appears to be true across many species. For example, e-coli bacteria demonstrate random tumbling when the nutrient medium becomes less rich, and predictable, repetitive, straight-ahead movement when nutrients increase to more favorable levels. The increased bacterial variability due to the adversity of low nutrients parallels, in some ways, what happens in the genome of bacteria, where adverse environmental conditions result in a high mutation rate. Increased phenotypic variation is also induced by environmental stressors (for a discussion related to operant variability, see Roberts & Gharib, 2006)

Operant-response variability also increases when reinforcers are withheld, for example, under conditions of extinction. In a notable experiment by Antonitis (1951), rats could easily produce food reinforcers by poking their noses anywhere along a horizontal opening. When food was later withheld, variability of nose-poke locations increased appreciably, that is, during the period of extinction. Increased variability due to extinction was later confirmed in many dimensions of response: location (Eckerman & Lanson, 1969); force (Notterman & Mintz, 1965); topography (Stokes, 1995); number (Mechner, 1958). A point worth reiterating is that while extinction increases variability, the set of variations is organized around the originally learned response, another illustration of bounded stochasticity. For example, if lever-pressing is the operant that produces food pellets, a rat may vary the ways in which it presses when food is withheld, but much of the behavior will be organized around pressing (e.g., Stokes, 1995).

Neuringer, Kornell and Olufs (2001) quantified the bounded nature of extinction-induced variability. In that experiment, rats were reinforced for repeating a single sequence across two levers and

a key: Left lever, Right lever, Key, in that order. To be reinforced, the rat had to press the left lever once, then the right lever, then push the key. Figure 3 shows the distribution of the relative frequencies of each of the possible sequences (proportions of occurrences) during the conditioning phases (filled circles) and during extinction of responding (open circles). LRK was, of course, most frequent during the reinforcement phase, with other somewhat similar sequences falling off in terms of their frequencies. LRK was also most frequent throughout the extinction phase, and the two curves were quite similar. Of
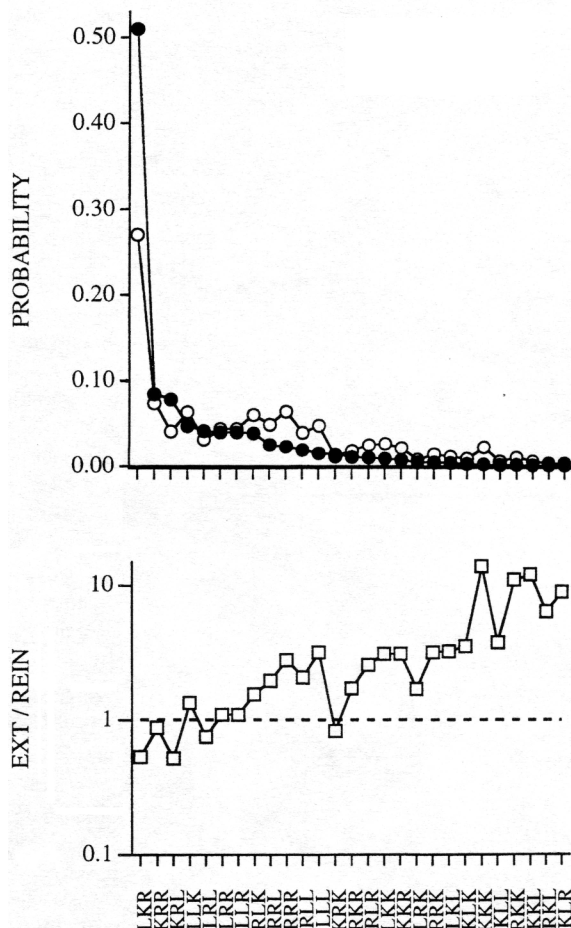

Figure 3. The top graph shows the proportion (or probability) of occurrences of the three-response patterns shown along the x-axis during a period when a single sequence (LKR) was being reinforced (filled circles) and during a period of extinction, when reinforcers were withheld completely (open circles). The bottom graph shows the ratio of responding during extinction to responding during reinforcement. Together the graphs show that patterns of responding during extinction were similar to those during reinforcement, but high-frequency sequences decreased and low-frequency sequences increased during the extinction phase.

course, because reinforcers were withheld during the extinction phase, response rates fell to very low levels. Also shown at the bottom of the figure are the ratios of response proportions during reinforcement and extinction phases (that is, the ratio of the curves in the upper graph). The take-home message here is that the basic form of the behavior was maintained during extinction while variability increased due to the generation of unusual or highly unlikely sequences. Extinction was therefore characterized as resulting

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, Kornell & Olufs, 2001, p. 79).

In daily life, the variability induced when reinforcement is withheld can provide the opportunity and behavioral substrate for learning new responses. That lesson can be applied to one's own behavior as well as when attempting to teach or influence others. When in a "rut," perhaps being unproductive or dissatisfied, it may help to remove and avoid commonly consumed reinforcers that had been produced by habitual behaviors.

Reinforcement Uncertainty. As an anticipated reinforcer is approached, in time or space, responding tends to become increasingly repetitive and predictable. This was shown for ratio schedule of reinforcement (Cherot, Jones, & Neuringer, 1996) and has been documented as well in Pavlovian schedules where activity becomes highly localized and stereotyped as a potential sexual reinforcer is approached (Atkins, Domjan & Gutierrez, 1994) as well as in ethological studies of animal behavior (Craig, 1918; Timberlake & Lucas, 1985). Roberts and co-workers relate effects such as these to "expectancy": As reinforcer expectancy increases, variability (in this case for lever-press hold-down durations) decreases (Gharib, Gade & Roberts, 2004). One interpretation of these effects is that response variability is generated by reinforcement uncertainty.

A commonly studied example is found in situations where organisms can choose among options, but where there is uncertainty as to which particular choice will return a reinforcer or when that will happen, for example, under concurrent schedules of reinforcement. Under such schedules, the moment of availability of reinforcers cannot be predicted but a scheduled reinforcer remains available until it is collected. The schedules may differ in terms of their richness, for example, the schedule on the left might provide reinforcement three times more frequently than the schedule on the right, but there is no way to predict exactly where the next reinforcer will be located. The common finding is "generalized matching," shown by a power-function relationship between relative frequencies of responses to the two alternatives and relative frequencies of obtained reinforcements (Baum, 1979). Also found is that choices are emitted stochastically rather than in any predictable pattern (Jensen & Neuringer, in press; Nevin, 1969). The reasons for why, in the presence of uncertain outcomes, choices should also be uncertain (stochastic and unpredictable in terms of the instance) are not clearly understood, but may be related to explorations of uncertain environments without excessive demands on memory capacities (Viswanathan et al., 1996).

Interactions. Accidents, decreased reinforcement, and reinforcer uncertainty are discussed as "non-contingent" effects because variability does not control these events. Non-contingent and contingent effects – the latter due to direct effects of reinforcing variability – often occur jointly, or interactively. Additional phases in the Grunow and Neuringer experiment described above provide a clear example. To recall, in the first phase of that experiment, four groups were reinforced for high, medium-high, medium-low, and low levels of response-sequence variability. In each of the groups, reinforcement was provided whenever a rat met its particular contingency requirement. There followed two additional phases in which the overall frequencies of reinforcement were systematically lowered by providing reinforcement only intermittently for satisfying the variability contingency. In the case in point, a Variable Interval (VI) schedule of reinforcement was superimposed on the variability contingency: first a VI 1 min (such that food pellets were limited to an average of once per min) and then VI 5 min (limiting food pellets to no more than once every 5 min). The way the schedules worked was that once a variable interval timed out, the first trial that met the threshold contingency (highly demanding of variability in one case, less demanding in another, even less in a third, and very little demanding of variability in the 'low' group) ended with a reinforcer. All other trials ended with a brief timeout.

The right side of Figure 4 shows the effects of these changes in reinforcement frequency on *rates of responding* in each of the four groups. As reinforcement frequencies were lowered, responding became less frequent, that is, response rates fell, and this was the case equally for all four groups. The left portion shows the effects on levels of response variability, indicated by U-value, a measure of sequence uncertainty. The higher the U, the closer was responding to a random model. The individual curves represent the four variability requirements and the x-axis represents frequencies of reinforcement. Of some interest, an interaction was observed: when low levels of variability were reinforced (the 0.37 group), decreases in reinforcement frequencies yielded *higher* levels of variability. When high levels of variability were reinforced (the 0.037 group) the opposite occurred, that is, decreasing reinforcements led to *lower* variability. The intermediate groups showed intermediate effects. A similar interaction was reported with respect to delayed reinforcement (Wagner & Neuringer, 2006).
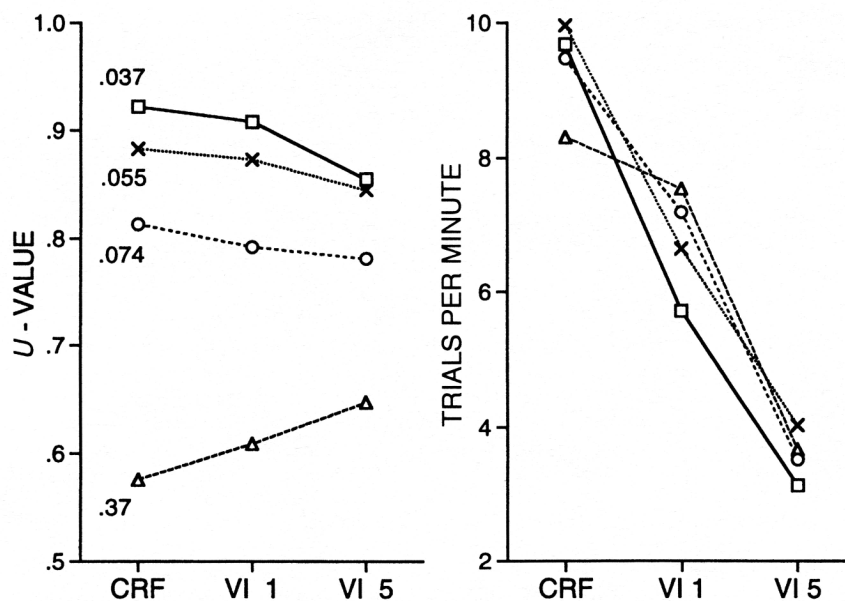

Figure 4. The left graph shows U-values, an index of behavioral variability, for each of four groups of rats: Open squares (.037) represent a group reinforced for highest levels of response-sequence variability; x's (.055) and open circles (.074) experienced variability contingencies that were increasingly lenient; and open triangles (.37) were reinforced even for relatively low levels of variability. During one phase (CRF) reinforcement was provided each time that an animal met its required contingencies. VI 1 and VI 5 indicate phases in which reinforcement frequencies were appreciably lowered – to no more than once per min (VI 1) and once per 5 min (VI 5). The right graph shows trials per minute, an index of speed of responding, under the same conditions as on the left.

Another illuminating example of an interaction occurred when groups of rats and pigeons, in separate experiments, were reinforced either for repeating a particular sequence (some groups) or for varying their sequences (other groups). For example, in the Cherot et al. (1996) experiment mentioned above, two groups of rats were reinforced, one for repeating 4-response sequences across two levers and the other for varying. However, not every sequence that met the VAR or REP contingency was reinforced, but rather a superordinate Fixed Ratio 4 was in place, that is, the REP group had to successfully repeat sequences four times to get a single reinforcer and the VAR group to vary successfully the same number of times to be reinforced. Levels of variability were much higher in the VAR group than in the REP, as was expected from the contingent effects of reinforcers on variability, shown in the

bottom graph of Figure 5. Furthermore, as reinforcement was approached (i.e., as the last of the four successful sequences was neared), levels of variability decreased, this being shown by the decreasing functions for both VAR and REP groups. The top graph in Figure 5 shows how these influenced performance accuracy. The elicitation of lowered variability facilitated correct responding in the Repeat group but interfered with it in the VAR.
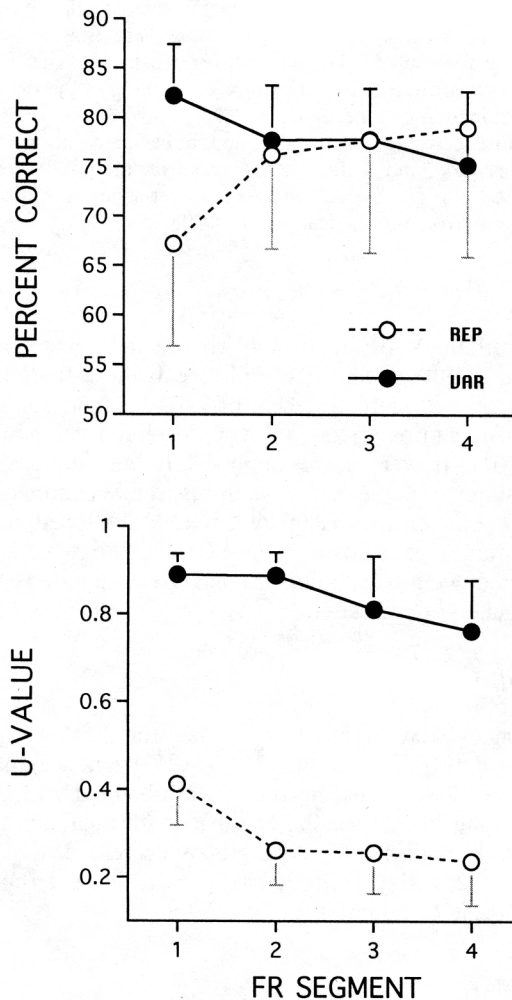

Figure 5. The top graph shows percentages of sequences that met variability (VAR) or repetition (REP) contingencies as a function of location within a fixed-ratio 4 (FR 4) schedule. The lines connect means for groups of rats and the error bars indicate standard deviations. The lower graph shows U-values, an index of sequence variability, for the two groups across the FR schedule.

The reason these findings may be important is that in many environments, repetitive behaviors are required, for example, factory work; but in other environments, highly variable (and unpredictable) behaviors are functional, for example, artistic endeavors. Differences in reinforcement frequencies, contingencies, proximities to outcomes, and the like may differentially influence how variability and repetition requirements affect responding. For example, decreasing reinforcement may affect behaviors

differently in the inventor, fashion designer, or artist than in the mail carrier, fare collector, or widget maker. Thus, in order to predict how environmental influences will affect behavioral variability, one must know initial levels of variability and the contingencies working to maintain those levels.

Reinforced variability and conditioning of new responses. Skinner suggested that operant behaviors are selected by reinforcers from a substrate of varying behaviors with the process paralleling the evolutionary process of variation and selection (Skinner, 1981). Others have supported the parallel (Baum, 1994; Hull, Langman, & Glenn, 2001; Staddon & Simmelhag, 1971). As described above, it is clear that variable behaviors can be generated by reinforcers directly contingent upon that variability, something not anticipated by Skinner or other writers on this topic, including Skinner's detractors. A question of importance is whether reinforced variability facilitates acquisition of new responses, especially difficult-to-learn ones.

Neuringer, Deiss, and Olson (2000, Exp 2) focused on that question. Thirty male Long-Evans rats were trained to press left (L) and right (R) levers in an operant chamber to obtain food-pellet rewards. During the main part of the experiment, the rats were divided into three groups, in each of which the rats were rewarded whenever they emitted a "target" sequence, namely RLLRL. The groups differed in one respect: the CONTROL group was reinforced *only* for the target sequence, with all other five-response sequences leading to a brief timeout. The VAR group was reinforced occasionally for varying, this in addition to reinforcement for every successful emission of the target sequence. The details are to be found in the original paper, but in brief, if a particular pattern was emitted *infrequently*, that is, less than about 3% of the time, it was occasionally reinforced. Thus, the VAR group was reinforced on a VI 1 min schedule for sequence variations while, at the same time, the difficult target sequence was immediately reinforced whenever it occurred. A third group served as a control for the additional food pellets, and this was an ANY group that was provided the same frequency of additional pellets, but independently of levels of variability. Therefore ANY animals *could* vary, but were not required to do so.

There were two main effects. First, the additional reinforcers caused both VAR and ANY groups to respond at high overall rates throughout the experiment whereas the absence of these reinforcers resulted in extinction of most responding in the CONTROL group. Many of the rats who were reinforced only for the difficult target sequence ended up sleeping through the training sessions. Thus, additional reinforcement motivated continued responding. The second result, shown in the top portion of Figure 6, was that only the VAR group came to acquire the difficult target sequence. Given variability of sequence generation, there was some probability that RLLRL would occur, and when that happened it was reinforced. The experiment was replicated with a different target sequence, LLRRL, and again, only the VAR group learned, this shown in the bottom of Figure 6 (see also Neuringer, 1993). As part of the Grunow and Neuringer (2002) experiment, described above, *levels* of reinforced variability were shown to contribute to the facilitative effects, with the higher the variability, the more likely that a difficult-to-learn sequence would be acquired. To recall, rats were in a chamber that contained three operanda, two levers (L and R) and a key (K), and three responses constituted one trial. One sequence was designated as the target, LKK, with other sequences occasionally reinforced when they met a variability contingency. These details are of interest because Maes and van der Goot (2006) used a procedure similar to that of Grunow and Neuringer but with human participants and obtained different results. Possible reasons for the difference will be discussed.
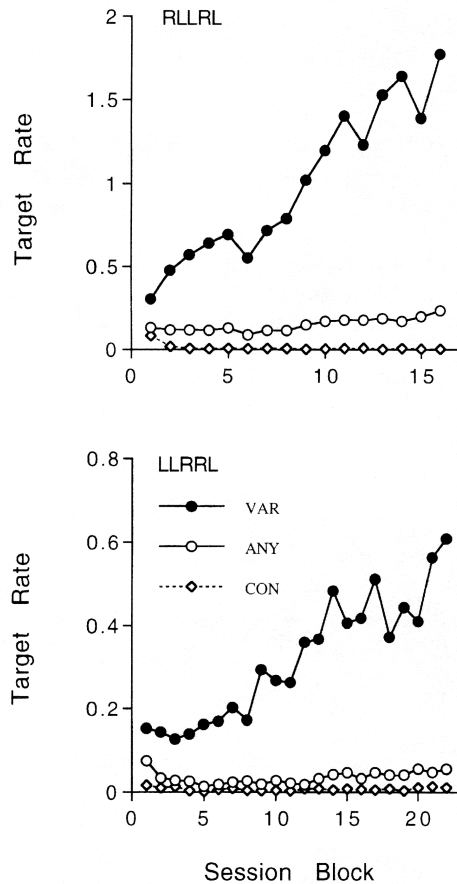
Figure 6. Rates of emission of a difficult-to-learn target sequence (RLLRL on top and LLRRL on bottom) for three groups of rats as a function of blocks of sessions (each session block shows the average of 5 sessions). One group was concurrently reinforced occasionally for varying sequences (VAR), another was reinforced at the same rate but independently of variability (ANY), and a third (CON) was not reinforced for any sequence other than the target sequences.

In the Maes and van der Goot experiment, 30 university students were occasionally reinforced for pressing three keys (1, 2, and 3) on a computer keyboard. Each trial consisted of three presses and ended either with "correct" appearing on the screen or simply being asked to continue to the next trial without reinforcement. The students were divided into three groups with all reinforced for a particular target sequence, 313, whenever that happened to be emitted. For the CONTROL group, the only way to receive "correct" was by entering the target sequence. For the VAR group, other sequences that met a variability contingency were occasionally reinforced. A second control group, YOKE, was occasionally told "correct" in addition to the target 313 but independently of variability. As in the studies with rats and pigeons, the participants in the VAR group responded more variably than those in the other two conditions. However, the VAR participants learned the target sequence *least well*, emitting fewer targets than either of the other two, with the Controls generating the most target sequences. The procedure was repeated with a more complex task involving six-response trials across two keys and a less-likely-to-occur-by-chance target, namely 211212. The results were essentially the same. Bizo and Doolan (2008) reported similar results, also with human participants: the VAR group learned least well. The Maes and

van der Goot study, and the Bizo and Doolan study, appear to have been performed carefully. The question, therefore, is why did three animal studies show that learning was facilitated when variability was concurrently reinforced and two human studies show deleterious effects?

Maes and van der Goot considered a number of procedural differences as possible explanations, and rejected many, rejections with which I agree. The most obvious difference is that of species, but in the area of reinforced variability, animal studies have been successful models of human performance. Indeed, Maes (2003) reported reinforcement-of-variability effects in human participants that paralleled in most respects those from animals. Other explanations considered differences in quality of reinforcer (food vs. "correct"), differences in baseline levels of variability (high in human, lower in rats and pigeons), differences in the "difficulty" or probability of the target sequence (very low probability in rats and pigeons, higher in humans). Among those contrasting procedural aspects that the authors think might have contributed, the frequency of reinforcement for variability was lower in the successful animal studies than in the human. Relative frequencies of reinforcement for varying versus repeating have, in fact, been shown to affect their respective probabilities of occurrence (Neuringer, 1992) and so when applying the Variation-and-Selection procedure, these frequencies should be carefully considered. Simply put, if the organism, animal or person, is highly likely to vary, then the frequency of reinforcement for variations should be lowered—otherwise motivation to learn the target may be low.

Another contributing factor may have involved motivation to respond. In the animal studies, responding essentially extinguished in the target-only control condition where no "extra" reinforcers were provided. Responding continued at high levels in all of the human experiments, including the control conditions. Thus, an interaction between motivation-to-respond and variability resulting from VAR reinforcement may have led to the positive animal findings.

Another aspect of the Maes and van der Goot (2006) procedure may have been important. The authors wrote that the "… search (by their human participants) for … explicit 'rules' might have interfered with finding the consistently reinforced target sequence. This search, and resulting interference, might be especially encouraged by the reinforcement of non-target sequences … in the VAR and YOKE conditions" (p. 91). In support, I note that the instructions provided to the participants were ambiguous. "Your task is to enter a 'correct' sequence of three keyboard keys as frequently as possible." And, "It is your task to find out what a 'correct' or an 'incorrect' sequence is." To the extent that the participants' behaviors were influenced by the second part of the instructions, those in the VAR group might have been motivated to continue to vary – as indeed was observed – in an attempt to "find out what a 'correct'…sequence is." This is especially the case since the main reward was money or course credit provided simply for participating (and independent of performance). The Bizo and Doolan (2008) instructions contained a similar ambiguity: *"Your task is to earn as many points as possible by figuring out what a 'correct' and 'incorrect' sequence is."* Again, the goal of "figuring out" may have maintained high variability in the VAR group, especially given that variable sequences were sometimes reinforced. I therefore wonder what would result from providing a large prize to the one participant who gained the most points and omitting instructions about "figuring or finding out..." The issue is of considerable importance since explicit reinforcement of variable behaviors may – or may not—serve as a tool for teaching, especially individuals who otherwise might have difficulty learning. Additional research is needed.

Autism and ADHD. Two conditions associated with learning difficulty are autism and ADHD. In both cases, there is failure of normal control over levels of variability, but in different ways, as will be discussed. For the individual with autism, levels of variability are abnormally low: repetitive, stereotypic, constant activities are defining characteristics of this disorder. However, the evidence

suggests that these levels may be influenced by reinforcing consequences.   One theory of autism is that behavioral repetitions are part of a coping mechanism to reduce high levels of arousal and uncontrollable external stimulation (Turner, 1997). The opposite hypothesis is sometimes offered as an alternative to explain the same stereotypic behavior: The individual with autism engages in repeated actions in order to *increase* stimulation (Turner). In either case, highly stereotyped behaviors are hypothesized to occur because of their consequences, indicating that the stereotypic activity is at least partly under operant reinforcement control. Turner writes, "On (these) view(s), repetitive behaviours are adaptive behaviours employed by the autistic individual as a homeostatic mechanism that regulates arousal levels" (p. 64).

Experimental studies indicate that the stereotyped behaviors characteristic of autism can indeed be influenced by direct reinforcement procedures. In one study, Miller and Neuringer (2000) reinforced 5 individuals diagnosed with autism. These 5, plus 9 control participants, were first reinforced during a baseline phase, independently of variability for playing a game on a computer. Following the baseline, reinforcement was made contingent on sequence variations. Reinforcers were points traded for things like toys, money, or edibles, depending upon the participant. There were two main results. First, the participants with autism behaved significantly less variably than the controls in both parts of the experiment. Second, both groups of participants increased significantly their variability when it was directly reinforced. The important point here is that individuals with autism, although relatively repetitive in their responding, as is characteristic of that disorder, were successfully reinforced for varying, and this was a contingent-reinforcement effect, since the control phase contained similar frequencies of reinforcements. Appropriate variations were directly reinforced.

Two experiments by Ronald Lee and co-workers extended this work. In both, individuals with autism were reinforced under a Lag schedule for varying their verbal responses to questions and in each experiment two out of three individuals with autism were successfully reinforced for appropriate verbal variations (Lee, McComas, & Jawor, 2002; Lee & Sturmey, 2006). The efficacy of direct reinforcement was also shown by Newman, Reinecke, and Meinberg (2000), but here in a situation where 2 of 3 young children diagnosed with autism learned to self-administer reinforcers contingent upon their own increasingly varied responses. Of course, at least during the initial phase, the researcher assured that the children were administering their reinforcers appropriately. Thus the experimental evidence, although not extensive, indicates that the behavior of individuals with autism can be influenced beneficially by reinforcers contingent upon variability of behaviors. Stated differently, the abnormally low levels of variability characteristic of individuals with autism may at least in part be an operant effect, influenced by contingencies of reinforcement. Because operant behaviors generally, and shaping of new operants specifically, manifest consequence-controlled variability, an important step in helping to change autistic behaviors in the direction of normalcy may be explicitly to reinforce those individuals for varying levels of variability, levels that range from unpredictable to repetitive.

ADHD is also associated with abnormal levels of variability, but in a direction opposite to that of autism. "Intra-individual variability in behavior and functioning is ubiquitous among children with attention deficit/hyperactivity disorder…" (Castellanos et al., 2005). Extremely *high* levels of variability are a defining characteristic of ADHD (Rubia, Smith, Brammer & Taylor, 2007). Such variability is seen in general behavior patterns, in responses during experimental tests, and often in reaction times. A second common identifier is lack of inhibitory control: "…children with ADHD consistently underperform controls on tasks of deliberately suppressing an activated motor response" (Nigg, 2001, p. 589).

Deficiencies in inhibitory control and abnormally high variability may be related. Assuming an operant class composed of many possible responses (Skinner, 1959), some of which are appropriate, but

many not, inadequate inhibitory control will result in emissions of inappropriate alternatives. Implied here are two things. First, operant response classes for individuals with ADHD may be abnormally large and not shaped or constrained by the demands of the situation or contingencies of reinforcement. Second, given large sets of possible responses, inhibition of inappropriate instances is absent in those with ADHD.

A question of some concern is the extent to which behavioral variability in ADHD is controlled by operant reinforcement contingencies, and the evidence indicates that, unlike the case for autism, variability may be the result mainly of non-contingent or eliciting influences. A common treatment regimen for ADHD is administration of drugs such as Ritalin, and such drugs appear to have a general, non-contingent influence on behaviors. In an animal model of ADHD, namely the Spontaneous Hypertensive Rat, or SHR, reinforcement of repetitions was relatively ineffective unless amphetamine was first administered (Mook & Neuringer, 1994). Non-contingent influences of the drug facilitated normal selective influences of reinforcers.

However, other evidence appears to contradict the "non-contingent influences theory," this evidence suggesting that behavioral variability in ADHD *is* highly sensitive to direct reinforcement. I will provide an alternative interpretation. Variability in individuals with ADHD is higher than in controls when reinforcement is infrequent but not when it is frequent (Aase & Sagvolden, 2006). This would appear to indicate that individuals with ADHD may require more frequent reinforcement than normal controls, but that the former are in fact sensitive to direct reinforcement procedures. However, recall that low frequencies of reinforcement and high uncertainty *elicit* high variability and the research on ADHD has generally not provided adequate controls for such non-contingent effects, for example, effects that are independent of the reinforcement *contingency* and due instead to frequency of non-contingent positive events.

A second case is that when reinforcement is delayed, there is a greater weakening of response in ADHD than in control participants. It has been suggested, therefore, that ADHD individuals are more sensitive to reinforcement immediacy than others. However, here, too, an alternative interpretation involves the fact that delaying reinforcement results in short-term decreases in reinforcement frequency and increased uncertainty both resulting in *elicitation* of increased variability. The eliciting effects of delay were hypothesized by Wagner and Neuringer (2006) to explain why delay increases variability generally. That the effect may be potentiated in ADHD suggests, indeed, a heightened sensitivity to attributes of reinforcement, but it may be primarily sensitivity to non-contingent or eliciting influences of reinforcing events. In the individual with ADHD, the selective effects of reinforcers may depend importantly on interactions with non-contingent or eliciting effects of the same reinforcers.

Autism and ADHD are both associated with abnormality in control of behavioral variability. In both cases, variability levels tend to be constrained, but to opposite ends of the variability continuum. Thus, it may be helpful to teach individuals with both disorders how to vary levels of variability. The evidence suggests, however, that contingent-reinforcement effects – directed to changing levels of variability—might be appropriate in autism but less so for ADHD. In the latter case, non-contingent influences, including those of drugs and overall densities of reinforcement, may be necessary before contingent-reinforcement effects come into play.

Normal operant behavior. Implied throughout the above discussion is that normal operant behavior manifests variations in levels of variability. Some variations are due to eliciting (or non-contingent) influences, some to selective or reinforcing influences, and some to endogenous effects. I would go further, however, to hypothesize that the characterization of operant behavior as "voluntary" depends upon such flexibility in variability levels. So that I am clear, I refer here to something richer and

more complex then the often noted decreases in variability observed when a newly acquired repetitive operant is reinforced. For example, when lever pressing is shaped, variability may be high at first and then systematically decreases with continued training. But many, perhaps most, voluntary operant responses show more dynamic and complex changes in variability, and sometimes moment-to-moment changes, that are adaptive within an environmental context. As an example, imagine a father attempting to entertain an infant with touches or tickles. At first, a touch to the child's stomach might be repeated, engendering smiles and laughs, but as the response decreases, or even turns to discomfort, the touches might be varied and, in some cases, done in a way that the child cannot predict. This, too, might continue if it engenders giggles and laughs, until it too leads to some manifestation of discomfort in the child and the father's behavior might then change.

The different levels of variations in "normal" operant behavior is contrasted with "abnormal" behaviors in which variability levels are restricted – to the "high" end for ADHD and the "low" end for individuals with autism. I next present empirical evidence that functional variations in variability are in fact a defining attribute of normal voluntary behavior.

Psychophysical estimates of voluntary operant response. B.F. Skinner wrote, "The standard distinction between operant and reflex behavior is that one is voluntary and the other involuntary" (Skinner, 1974, p. 44). Indeed, if one were to ask lay people to characterize operant responses, they would not have much of an idea, but most would offer an opinion about how voluntary behaviors differ from involuntary. We used this presumed equivalence between voluntary and operant behavior to assess, via a series of psychophysical experiments, the relationship of variations in levels of operant variability to appearance of voluntary behavior (Neuringer, Jensen, & Piff, 2007). Participants discriminated voluntary from non-voluntary behaviors, and that permitted us to explore the controlling dimensions. In these experiments, the participants judged whether or not a computer-simulating icon represented a normal, voluntarily performing human. There were a number of different icons, each programmed according to a different response strategy in a game-type environment where the icons repeatedly chose among three response alternatives. One icon stochastically matched its choices to relative frequencies of reinforcement --reinforcers were programmed by concurrently operating schedules -- as indicated by Herrnstein's (1970) matching law. Another icon stochastically "undermatched," that is, allocated its choices relatively equally across the three alternatives despite differences in received rewards (Baum, 1979). Yet another stochastically "overmatched," that is, allocated its choices preponderantly to the highest-reward alternative (Baum, 1979).

The rationale for this procedure was the following. Under concurrent schedules of reinforcement, a "matching" organism will respond equally across the available alternatives when reinforcers are delivered equally. As indicated above, allocation of choices is generally found to be stochastic-like in quality (without obvious patterns) and therefore difficult or impossible to predict. Thus, given equally allocated reinforcers, responding is highly unpredictable. However, when reinforcers are asymmetrically provided, such that, for example, 10 times as many reinforcers are received from one alternative than another, responding now becomes much more predictable. Although there continues to be stochastic emission, since it is now at a ratio of 10:1, predictions are much more likely to be correct. The point, here, is that a stochastic matcher under concurrent reinforcement conditions sometimes demonstrates difficult-to-predict choices and other times easy-to-predict choices. This point becomes all-the-more relevant when many more than two choice alternatives are available where differences in ability to predict (or respond unpredictability) are much greater.

Now let's consider the under- and over-matchers' performances. The stochastic undermatcher tended to respond equally across the available alternatives, no matter the distribution of reinforcements.

Stated differently, the undermatcher's behavior was relatively unpredictable throughout. Whether equal reinforcers were received, or highly asymmetric reinforcers, the undermatcher's performance could not readily be predicted. In some ways, therefore, the undermatcher was representative of the individual with ADHD, always behaving in a highly variable, noisy, and unpredictable manner. The overmatcher, on the other hand, did the opposite, tending to respond predominantly to the alternative from which most reinforcers were received, even if the difference in alternatives was small. The overmatcher's performance was highly repetitive, stereotyped, one might say autistic-like, and predictable. I have described here the extremes, but in fact we utilized the entire continuum, from undermatching through matching to overmatching, a continuum represented by the *sensitivity parameters* in Baum's generalized matching function (Baum, 1979).
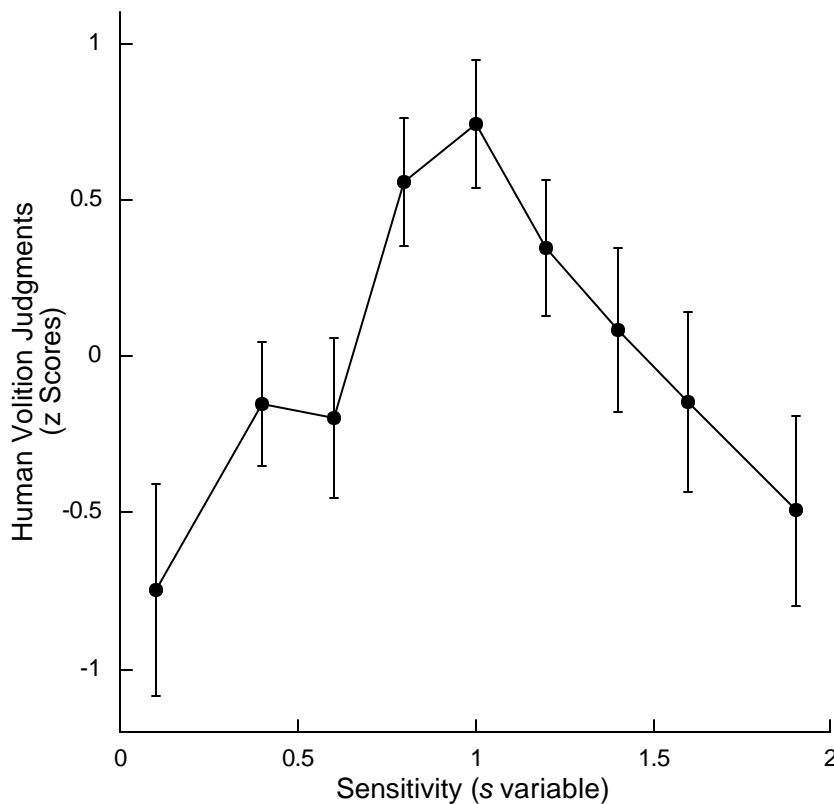

Figure 7. Probabilities (averaged across 13 human participants) of identifying an icon as a "voluntarily choosing icon," this indicated by the z-score transformations of volition judgments, as a function of the *s* value governing the choices made by the icon. An *s*=1.0 indicates exact matching of proportions of choices to proportions of received reinforcers; values less than 1 indicate "undermatching"; and values greater than 1.0 indicate "overmatching."

Here are some particulars of the experiment. Human participants were asked to compare the choices emitted by the various icons and judge whether they represented normal humans who were

voluntarily choosing in a gambling-game type of environment. The matching icon was judged to most likely represent a normal responder (Figure 7). Of course there are many alternative interpretations of this finding, but a series of control procedures pointed to the importance of variations in levels of variability – or predictability – as a key factor. One control showed that the judgments were not based primarily on the obtained frequencies of reinforcement. That is, the icon best representing the voluntary chooser was not simply the one who gained the most points. Another showed that it was not the fact that choices *matched* relative frequencies of reinforcement that was solely important. In this experiment, one icon matched choices to reinforcement but chose in a highly predictable manner (e.g., 333221333221333221…) in a case where it was reinforced approximately 3:2:1. The other icon also matched its choices to reinforcements, but did so in a stochastic and unpredictable manner (e.g., 331212233332323132…). In both cases, when proportions of reinforcements changed, choices changed, predictably in one case, stochastically in the other. An important point, therefore, is that the stochastic matcher *sometimes* behaved relatively predictably and other times not, depending upon reinforcement distributions. The stochastic matcher was judged better to represent voluntary human choice than the always predictable matcher. The take-home message here is that we perceive voluntary behavior to be related to the ability of an actor to change levels of variability – to behave in a way that is highly predictable under some circumstances but that is unpredictable in other circumstances, that is, that cannot be predicted at levels greater than chance.

Why might this research on volition be important? For thousands of years, volition and free will have been discussed and debated: These are highly valued human competencies. I suggest that Skinner was right: These notions in fact refer to an extraordinarily important behavioral competency, that of the voluntary operant. Although explanations of volition may have been in error – much as explanations of the apparent movement of the sun around the earth was in error – the phenomena are real – astral bodies move, and voluntary acts can be distinguished. A behavioral analysis of volition, as I've just offered, enables us better to understand and influence it; indeed, it provides the possibility of training individuals to better engage in volitional acts, or to behave more adaptively when "free choice" is possible. It also permits us to view abnormal behaviors within the context of volitional effects, including inabilities to vary levels of variability.

Conclusion.

During much of the history of the study of operant responses, emphasis has been on the conditioning and maintenance of repeated responses. Analyses of behavior, whether at the molecular or molar level, show much repetition. Some researchers and writers have concluded that operant conditioning and reinforcement processes apply *only* to repetitions (Schwartz, Schuldenfrei, & Lacey, 1978) and, because of that, much of the richness of normal human (and animal) behavior could not be explained by Skinnerian operant theory. A notable example was Noam Chomsky's (1959) critique of Skinner's *Verbal Behavior*. One of the criticisms was that reinforcement principles could not account for the generation of novel linguistic output. In fact, a "reinforcement affects only repetitions" position is inadequate to explain almost any complex operant. Needed is an understanding of the rich variations that underlie successful selections. Contributing importantly to that richness is the variability of the operant and its selection and control by powerful contingencies of reinforcement.

References

Aase, H., & Sagvolden, T. (2006). Infrequent, but not frequent, reinforcers produce more variable responding and deficient sustained attention in young children with attention-deficit/hyperactivity disorder (ADHD). *Journal of Child Psychology and Psychiatry, 47,* 457-471.

Antonitis, J. J. (1951). Response variability in the white rat during conditioning, extinction, and reconditioning. *Journal of Experimental Psychology, 42,* 273-281.

Atkins, C. K., Domjan, M., & Gutierrez, 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.

Bandura, A. (1982). Self-efficacy mechanisms in human agency. *American Psychologist, 37*, 122-147.

Baum, W. M. (1979). Matching, undermatching, and overmatching in studies of choice. *Journal of the Experimental Analysis of Behavior, 32*, 269-281.

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

Beveridge, W. I. B. (1957). *The art of scientific investigation*. New York: Vintage.

Bizo, L. A., & Doolan, K. (2008). *Reinforced behavioural variability*. Paper presented at the Association for Behavior Analysis meeting, Chicago, IL.

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.

Caporale, L. H. (1999). Chance favors the prepared genome. In L. H. Caporale (Ed.) *Molecular strategies in biological evolution. Annals of the New York Academy of Science, Vol. 870* (pp. 1-21). New York: New York Academy of Sciences.

Castellanos, F. X., Sonuga-Barker, E. J. S., Schere, A., Di Martino, A., Hyde, C. & Walters, J. R. (2005). Varieties of attention-deficit/hyperactivity disorder-related intra-individual variability. *Biological Psychiatry, 57,* 1416-1423.

Cherot, C., Jones, A., & Neuringer, A. (1996). Reinforced variability decreases with approach to reinforcers. *Journal of Experimental Psychology: Animal Behavior Processes, 22*, 497-508.

Chomsky, N. (1959). Review of B. F. Skinner's *Verbal behavior. Language, 35,* 26-58.

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.

Craig, W. (1918). Appetites and aversions as constituents of instincts. *Biological Bulletin of the Marine Biological Laboratory, Woods Hole, MA, 34*, pp. 91–107.

Deich, J. D., Allan, R. W., Zeigler, H. P. (1988). Conjunctive differentiation of gape during food-reinforced key pecking in the pigeon. *Animal Learning Behavior, 16*, 268–276

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

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.

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.

Gharib , A., Gade, C., & Roberts, S. (2004). Control of variation by reward probability. *Journal of Experimental Psychology: Animal Behavior Processes, 30*, 271-282.

Goetz, E. M., & Baer, D. M. (1973). Social control of form diversity and emergence of new forms in children's blockbuilding. *Journal of Applied Behavior Analysis, 6*, 209-217.

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

Herrnstein, J. (1970). On the law of effect. *Journal of the Experimental Analysis of Behavior, 13,* 243-266.

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.

Jensen, G., & Neuringer, A. (in press). Choice as a function of reinforcer "hold": From probability learning to concurrent reinforcement. *Journal of Experimental Psychology: Animal Behavior Processes.*

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.

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.

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.

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

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

Maynard Smith, J. (1982). *Evolution and the theory of games.* Cambridge: Cambridge University Press.

Mechner, F. (1958). Sequential dependencies of the lengths of consecutive response runs. *Journal of the Experimental Analysis of Behavior, 1,* 229-233.

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

Mirenowicz, J., & Schultz, W. (1994). Importance of unpredictability for reward responses in primate dopamine neurons. *Journal of Neurophysiology, 72,* 1024-1027.

Mook, D. M., & Neuringer, A. (1994). Different effects of amphetamine on reinforced variations versus repetitions in spontaneously hypertensive rats (SHR). *Physiology & Behavior, 56,* 939-944.

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. (1992). Choosing to vary and repeat. *Psychological Science, 3,* 246-250.

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., Deiss, C., & Olson, G. (2000). Reinforced variability and operant learning. *Journal of Experimental Psychology: Animal Behavior Processes, 26.* 98-111.

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.

Neuringer, A., Kornell, N., & Olufs, M. (2001). Stability and variability in extinction. *Journal of Experimental Psychology: Animal Behavior processes, 27,* 79-94.

Nevin, J. (1969). Interval reinforcement of choice behavior in discrete trials. *Journal of the Experimental Analysis of Behavior, 12,* 875-885.

Newman, B., Reinecke, D. R., & Meinberg, D. L. (2000). Self-management of varied responding in three students with autism. *Behavioral Interventions, 15,* 145-151.

Nigg, J. T. (2001). Is ADHD a disinhibitory disorder? *Psychological Bulletin, 127,* 571-598.

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

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

Pennisi, E. (1998). How the genome readies itself for evolution. *Science, 281*, 1131-1134.

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.

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

Roberts, S., & Gharib, A. (2006). Variation of bar-press duration: Where do new responses come from? *Behavioural Processes, 72,* 215-223.

Ross, C., & Neuringer, A. (2002). Reinforcement of variations and repetitions along three independent response dimensions. *Behavioural Processes*, *57*, 199-209.

Rubia, K., Smith, A. B., Brammer, M. J., & Taylor, E. (2007). Temporal love dysfunction in medication-naïve boys with attention-deficit/hyperactivity disorder during attention allocation and its relation to response variability. *Biological Psychiatry, 62,* 999-1006.

Schwartz, B., Schuldenfrei, R., & Lacey, H. M. (1978). Operant psychology as factory psychology. *Behaviorism, 6,* 229-254.

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

Skinner, B. F. (1959). The generic nature of the concepts of stimulus and response, 1935, 12, 40-65. Reprinted in Skinner, B. F., *Cumulative record* (pp. 347-366). New York: Appleton-Century-Crofts.

Skinner, B. F. (1974). *About behaviorism.* New York: Knopf.

Skinner, B. F. (1981).   Selection by consequences. *Science, 213,* 501-504.

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.

Stokes, P. D. (1995). Learned variability. *Animal Learning & Behavior, 23*, 164-176.

Taleb, N. N. (2007). *The black swan.* New York: Random House.

Thorton, J. (2006). New genes, new functions: Gene family evolution and phylogenetics. In. C. W. Fox & J. B. Wolf (Eds.), *Evolutionary genetics: Concepts and case studies* (pp.157-172). Oxford: Oxford University Press.

Timberlake, W., & Lucas, G. A. (1985). The basis of superstitious behavior: Chance contingency, stimulus substitution, or appetitive behavior? *Journal of the Experimental Analysis of Behavior, 44,* 279-299.

Turner, M. (1997). Towards an executive dysfunction account of repetitive behaviour in autism. In J. Russell (Ed.), *Autism as an executive disorder* (pp. 58-100). Oxford: Oxford University Press.

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.

Wagner, K., & Neuringer, A. (2006). Operant variability when reinforcement is delayed. *Learning & Behavior, 34,* 111-123.

Author Contact Information

Allen Neuringer,
Department of Psychology
Reed College
Portland, OR 97202
Office phone: 503-517-7403
Fax: 503-777-7785
e-mail: allen.neuringer@reed.edu