Reinforcement and Induction of Operant Variability

Allen Neuringer
Reed College

The target paper by Barba (2012) raises issues that were the focus of my first two publications on operant variability. I will describe the main findings in those papers and then discuss Barba's specific arguments.

Can Variability Be Reinforced?

Neuringer (1986) trained students to respond in random-like fashion by providing feedback based on 10 statistical tests of randomness. The research challenged conclusions from more than 50 prior publications that people did not have the ability to respond randomly (Brugger, 1997). However, reinforcement procedures had never before been utilized. Rather, human participants had simply been asked to respond randomly (e.g., by calling out heads and tails as would be expected from a coin flip). My procedure differed in that the students entered 100 responses per trial on the 0 and 1 keys of a computer, with feedback following each trial. At the beginning of training, responses differed significantly from a random model, but with training, the same statistical tests indicated approximations to randomness. Figure 1 shows the diverse nature of the feedback provided at the end of training. No single measure suffices as a test of randomness because randomness is a complex, multifaceted phenomenon. Indeed, statisticians and mathematicians have long debated the very definition of randomness.

Initially the participants saw only the first line, representing the percentages of 0s emitted in each trial, with the goal being to generate the

same distribution of 0s and 1s as expected from a random generator. It is important to note that the students' task was not simply to produce the same average number of 0s and 1s as the random generator, or 50% on every trial, but to distribute the percentages across trials, as would be expected for a random output, as indicated by the columns. The article can be consulted for details, but the students' goal was to equalize the number of trials across the columns. When a student generated equal percentages of 0s and 1s in a way that was statistically indistinguishable from the random model, a second statistic on a second line was added as feedback, which represented the number of changes from 0 to 1 or 1 to 0 in each trial. Note that *changes* is exactly the same as the *switches* discussed by Barba and below. When both 0s and changes met the random criterion, a third, then a fourth, and so on, statistic was added. Four of the statistics, namely RNG1, RNG2, C1, and C2, were information- or entropy-based statistics related to U value, a measure discussed in detail by Barba. At the end of training, all participants met the criterion of no statistical difference between their responses and the random-number model on all 10 statistics. The feedback at the end of each trial was presumed to serve as a conditioned reinforcer, with a large reinforcer (paid time off) contingent on success in meeting the randomness criterion. The results therefore support the claim that random-like behavior can be reinforced.

Is Variability an Operant Dimension?

To experimentally analyze operant variability, Page and Neuringer

Address correspondence to the author at allen.neuringer@reed.edu.

Descriptive statistic	CL1	CL2	CL3	CL4	CL5
%0 (low)	11	14	8	12	13
CHANGE (low)	9	11	10	18	8
RUN1 (low)	13	13	9	12	12
RUN2 (low)	10	10	18	8	11
RUN3 (low)	12	12	12	12	12
RUN4-10 (low)	15	9	11	8	15
RNG1 (var)	13	13	6	14	12
RNG2 (var)	11	11	10	12	13
C1 (low)	7	12	18	9	11
C2 (low)	6	17	11	11	10

Figure 1. Feedback table from Neuringer (1986).

(1985) simplified the procedures and measures described above. They reinforced sequences of responses that had not been emitted by pigeons across a number of previous trials, specified by a lag contingency. Trials consisted of eight responses on left (L) and right (R) keys, with a trial ending in food reinforcement if the lag contingency had been met and a time-out if not. In one experiment, for example, reinforcement was based on the current trial's sequence differing from each of the last 50 trials. This Lag 50 contingency resulted in highly variable responding, and most trials ended with reinforcement.

A critical test of the operant nature of variability was whether the reinforcement contingency was responsible. As test, a yoked control was applied in which frequencies and distributions of reinforcers were identical to that in the Lag 50 Vary condition, but variability was not required. As a result, response variability decreased to much lower levels. An important second indication that variability is an operant dimension was discriminative control: The pigeons learned to respond repetitively when the keys were one color and variably when the keys were a different color. But Page and Neuringer (1985) provided additional evidence. Levels of variability were shown to be controlled by the contingency: As the requirement became more demanding (higher lag) variability generally increased, a finding replicated in many experiments

(e.g., Grunow & Neuringer, 2002; Machado, 1989). Later publications showed that choices to vary or repeat were controlled by relative frequencies of reinforcement, much as is the case for choices between two operanda (Abreu-Rodrigues, Lattal, dos Santos, & Matos, 2005; Neuringer, 1992). Despite this wealth of evidence, Barba has argued against the operant nature of variability.

Barba Argument 1

Because reinforcement contingencies differ from behavior-change measures, operant variability has not been demonstrated. According to Barba, demonstration of an operant depends on the following: The "response" that changes during its conditioning must be identical to the "response" on which reinforcement is contingent during shaping. In the present case, "variability" is hypothesized to be a conditioned "response," but the variability on which reinforcement is contingent differs from the variability used to demonstrate operant-type changes. Specifically, in many cases, reinforcement is contingent on meeting a lag contingency (one indicator of variability), but successful conditioning is shown by changes in U value (a second indicator of variability). Because of this disconnection, Barba argues that the evidence is not sufficient to show that variability can be reinforced.

In many publications, however, there is no discrepancy. Although most studies of operant variability provide U-value data, there are also many cases in which the reinforcement contingency and behavior-change measures are identical, a point noted by Barba. For example, Morgan and Neuringer (1990) compared rats' response-sequence variability under a Lag 5 Vary schedule with that under a yoked condition (Yoke). The effects of three different types of operanda were studied. The

results showed that reinforced variability was highest when lever presses were required, next highest with key pushes, and lowest when overhead trapezes had to be pulled; these effects were also seen in the Yoke conditions. U-value statistics were provided to compare across operanda conditions and across Vary versus Yoke, but so also was percentage of variation, defined as the percentage of trials in which the Lag 5 criterion was satisfied. Percentage of variation was computed under both Vary and Yoke conditions (but only in Vary did it relate to presentation of reinforcers). In other words, meeting the Lag 5 contingency was required for reinforcement, and the frequency of meeting that contingency was one of the measures of performance change. Furthermore, changes measured by percentage of variation were quite similar to the U-value changes. Many other publications on operant variability provide both U values and measures of meeting the reinforcement criteria (e.g., Cherot, Jones, & Neuringer, 1996; Denney & Neuringer, 1998). The results are consistent: Reinforcement of variability (via lag contingencies, threshold contingencies, or others) leads to high levels of variability, measured in a variety of ways.

Let us also consider Barba's general point, namely that for an aspect of behavior to be described as operant, measures of behavior change and reinforcement contingency must be identical. That is not the case for "generalized operants," or whenever the operant is a general class of responses. Imitation is one example. Reinforcers may be based, in one case, on a child waving her hand when you do, and jumping up and down in another, and so on. But ultimately what is learned is to imitate. Imitation is the conditioned operant. Behaving unpredictably is a similar kind of phenomenon in which a particular instance is reinforced but measurement is of the more general

"respond unpredictably." In each of these cases, a response class is conditioned, whereas individual members of the class are reinforced. (According to Skinner, 1935/1961, the same is true even for the simplest operants, e.g., lever press. I add that during the successive approximation shaping process, microswitch closures are not reinforced, but they serve as an index of the conditioned response class.) No one instance represents the entire class. Specification and measurement of the class may require different (or additional) measures than those employed in the individual reinforcement procedures.

Barba Argument 2

U value is not a general measure of variability. There may be confusion as to what is meant by *general*. U value has been typically used to assess the overall entropy, or information value, in a set of responses. As such, comparisons can be made, in terms of U value, across conditions within an experiment and across publications. But U value is not "general" in the sense of replacing or representing all other measures of variability or randomness, nor does a single, overall measure exist. U value can be "gamed" by responding in a strategic or predictable manner; for example, if there are four possible sequences (two operanda, trial lengths of two responses), then responding LL, LR, RL, RR, LL, LR, RL, RR, ... will generate the maximum U value, thereby indicating "randomness." Lag contingencies can similarly be met via a strategy of repetitive sequencing. For example, assuming the same parameters as just described, then responding in the same way under a Lag 3 contingency would result in *every* trial being reinforced (i.e., 100% reinforcement). But, in both cases, responses were repetitive and predictable.

Why, then, use U value as a measure of variability and lag (or

another common procedure, threshold) as a contingency? One reason is that it follows Skinner's second unformalized principle of scientific practice: "Some ways of doing research are easier than others" (Skinner, 1956, p. 224). Lag schedules are readily programmed, and U values easily computed. And U value shares a close relation with other studies and fields in which entropy is the measure of variability. Strategic responding has rarely been observed in nonhuman animals except occasionally when lag values are low (e.g., in Lag 1 or Lag 2 types of schedules; Machado, 1992; Manabe, Staddon, & Cleaveland, 1997). Such strategizing has sometimes been observed in human participants under lag contingencies, and should be looked for. Autocorrelations and Markov analyses will identify such strategizing. Absent such strategic responding, U value has documented strong effects of lag-based or threshold-based reinforcement.

Barba Argument 3

Lag schedules reinforce switching rather than variability. Much of the target article is based on research by Machado (1997) in which reinforcement of switches was hypothesized to account for the variability generated by lag schedules. In one of Machado's experiments, reinforcement was contingent on one or more switches across L and R operanda in an eight-response trial. Thus, for example, an LRRRRRRR sequence would be reinforced, and it didn't matter whether this sequence was repeated or not; reinforcement depended only on the presence of one or more switches. A surprising result was that this simple contingency generated very high variability; indeed, variability was essentially equal to that under a Lag 25 schedule, also shown by Machado. It seemed reasonable, therefore, to claim that reinforcement of switching explained the variability. An important implication was that

variability was not itself reinforced and, consequently, that variability is not an operant. Stated differently, so-called operant variability can be explained by reinforcement of switching.

Page and Neuringer (1985) considered two different explanations. One was that under lag contingencies, previous sequences served as discriminative cues for what not to do (i.e., the pigeons remembered the previous sequences and did something different); in other words a *memory* hypothesis. The second hypothesis was that pigeons (and many other species) have the potential to respond in random-like or stochastic fashion, referred to as the *random* hypothesis. The evidence from Page and Neuringer favored the random hypothesis. First, it was deemed unlikely that pigeons could remember 50 eight-response sequences, as would be required under Lag 50. Second, when lag values were increased from 5 to 50, percentage of trials that met the lag contingency changed in a way similar to that of a random-generator model. Third, when number of responses per trial were systematically manipulated, probability of meeting the lag contingency changed as predicted by the random hypothesis and opposite to that predicted by the memory hypothesis (see also Jensen, Miller, & Neuringer, 2006). In other words, Page and Neuringer concluded that reinforcement of variability explained the lag results.

To evaluate Barba's (and Machado's) alternative explanation, let us consider results from two additional experiments by Machado (1997). In one, two or more switches per trial (rather than one or more) were required for reinforcement; in a second, the frequencies of switches expected from a random source were reinforced differentially. These two procedures generated levels of variability that were quite similar to one another and similar as well to variability under the one-or-more switch condition. If reinforcement of switching

were responsible for variability, then parametric changes in the switch contingency should affect the variability, but that was not the case. Stated differently, variability was not correlated with reinforcement parameters. Might something other than reinforcement of switching account for all three of Machado's findings?

One consistent aspect of the three procedures was that sequences that contained no switches were never reinforced. That is, eight Ls in a row or eight Rs were never reinforced. (In Machado's third experiment all-L and all-R sequences were each reinforced with a .03 probability, which effectively meant extremely rarely or never.) These two sequences are the most likely to be emitted when variability is not required (e.g., Hunziker, Saldana, & Neuringer, 1996; Machado, 1992), and therefore extinction of the two most probable sequences may have contributed, perhaps importantly, to the observed high levels of variability. I hypothesize that extinction of all-L and all-R sequences induced variability.

Induction of variability does not depend on direct reinforcement. For example, variability is induced when previously experienced reinforcers are withheld (i.e., during extinction, or when reinforcement frequencies are lowered; Lee, Sturmey, & Fields, 2007). Concurrent schedules of reinforcement induce highly variable allocations of responses at the micro level while molar distributions can be described by a power-function relation between overall choices and obtained reinforcers (Jensen & Neuringer, 2008). High levels of choice variability are also seen in studies of "spontaneous alternation" and in radial-arm mazes, again these being induced effects because reinforcers do not depend on the variability. And other variables, such as distance between operanda and the type of operanda, influence induction effects (e.g., Morgan & Neuringer, 1990, as described above).

Therefore an alternative to Machado's and Barba's switch hypothesis is that

(a) extinction of preferred response sequences (LLLLLLLL and RRR-RRRRR) induces relatively high sequence variability, and (b) adventitious reinforcement of induced variability maintains the high levels. Induction of variability no doubt plays a role in other operant-variability studies, but must be considered in conjunction with reinforcement-of-variability effects, both contingent and adventitious. Thus, for example, under Yoke contingencies, variability, although low, is maintained at levels generally greater than zero. Induction (due to intermittency of reinforcement and other influences) plus adventitious reinforcement may account for variability under Yoke. Yoke conditions and Machado's more-than-one-switch contingency are similar except for the withholding of reinforcement for preferred sequences in the latter case. This difference, I suggest, may explain the relatively high levels of variability under Machado's switch procedures.

An important take-home message is that both induction and direct reinforcement (or selection) must be considered when accounting for the variability of operant responses. Both often contribute to outcome, with induced effects sometimes facilitating reinforced variability and at other times counteracting it. Cherot et al. (1996) provide an illustrative example. Pigeons were required to meet a Lag 4 contingency (L and R operanda, four-response trials), but to do so four times (FR 4) for each food reinforcer. With an asterisk indicating satisfying the Lag 4 Vary contingency and dash indicating not, an example is as follows: LLRR*, RRRL*, LLRR-, LLRR-, RLLR*, RRRL-, LRRR* + food reinforcer. Thus, a pigeon had to satisfy the Vary requirement four times to be fed. A second group of pigeons had exactly the same conditions with one exception: They were required to satisfy a Repeat contingency, namely the sequence had to be one that had been emitted at least once during the

preceding four trials (with LLLL and RRRR being excluded). Figure 2 shows the results, both in terms of percentage correct and U values for each group. Consider first the U values in the bottom graph. The Vary pigeons responded much more variably than did the Repeat pigeons. Thus, the Vary versus Repeat reinforcement contingencies were highly effective in controlling variability. But note also that U values decreased as the food reinforcer was approached (i.e., as the fourth segment neared in the FR 4), and this was true for both Vary and Repeat. Other publications have documented that variability decreases as a reinforcer is approached in time or space; this is sometimes referred to as effects of anticipation, expectancy, or proximity to reinforcement (Gharib, Gade, & Roberts, 2004; Stahlman, Roberts, & Blaisdell, 2010). The results of this induced variability are shown in the top graph: Repeat performances improved with proximity to the reinforcer, and Vary performances were degraded. Thus both induction effects (due to proximity to reinforcers) and contingency effects (due to direct reinforcement of variability) combined to generate variability and therefore success rates. Induction facilitated performance in one case (Repeat) and interfered with it in the other (Vary). A final observation: In this as well as most other studies, the variability contingencies exerted a greater effect on levels of variability than did induced effects (e.g., Grunow & Neuringer, 2002). Variability contingencies are powerful.

Application: Extinction Plus Reinforced Variations

Withholding reinforcement from, or extinction of, undesirable repeated responses (as modeled by withholding reinforcement for all-L and all-R responses in Machado's experiment) can be combined with direct reinforcement of variability among desired

Figure 2. The top panel shows the percentage of trials in which two groups of pigeons satisfied their respective contingencies and were reinforced with food. The Repeat (Rep) birds were required to repeat response sequences, and the Vary (Var) birds were required to vary the sequences. The x axis represents location in a fixed-ratio (FR) 4 requirement indicating that each group was required to satisfy the contingencies four times to obtain food. The bottom panel represents the U values generated under Repeat and Vary conditions, again as a function of location within the FR 4 requirement.

responses. This “extinction plus reinforced variations” may be helpful as a therapeutic procedure (see Neuringer, 1993, for a related procedure with pigeons). As in my interpretation of Machado's findings, withholding of such reinforcement is likely to induce variability that can then be strengthened and maintained by reinforcers that are contingent on variability. A second phase could be to select particularly desirable responses from the

variations. Selection from reinforced variations has been shown in animal models (Neuringer, Deiss, & Olson, 2000) but so far not with human participants (Maes & van der Goot, 2006). However, successful application of reinforced variability in individuals with autism (Lee, McComas, & Jawor, 2002) suggests that an extension to "extinction plus reinforced variations" is worth future study.

REFERENCES

- Abreu-Rodrigues, J., Lattal, K. A., dos Santos, C. V., & Matos, R. A. (2005). Variation, repetition, and choice. *Journal of the Experimental Analysis of Behavior*, 83, 147–168.
- Barba, L. S. (2012). Operant variability: A conceptual analysis. *The Behavior Analyst*, 35, 213–227.
- Brugger, P. (1997). Variables that influence the generation of random sequences: An update. *Perceptual & Motor Skills*, 84, 627–661.
- Cherot, C., Jones, A., & Neuringer, A. (1996). Reinforced variability decreases with approach to reinforcers. *Journal of Experimental Psychology: Animal Behavior Processes*, 22, 497–508.
- Denney, J., & Neuringer, A. (1998). Behavioral variability is controlled by discriminative stimuli. *Animal Learning & Behavior*, 26, 154–162.
- Gharib, A., Gade, C., & Roberts, S. (2004). Control of variation by reward probability. *Journal of Experimental Psychology: Animal Behavior Processes*, 30, 271–282.
- Grunow, A., & Neuringer, A. (2002). Learning to vary and varying to learn. *Psychonomic Bulletin & Review*, 9, 250–258.
- Hunziker, M. H. L., Saldana, R. L., & Neuringer, A. (1996). Behavioral variability in SHR and WKY rats as a function of rearing environment and reinforcement contingency. *Journal of the Experimental Analysis of Behavior*, 65, 129–144.
- Jensen, G., Miller, C., & Neuringer, A. (2006). Truly random operant responding: Results and reasons. In E. A. Wasserman & T. R. Zentall (Eds.), *Comparative cognition: Experimental explorations of animal intelligence* (pp. 459–480). Oxford, UK: Oxford University Press.
- Jensen, G., & Neuringer, A. (2008). Choice as a function of reinforcer "hold": From probability learning to concurrent reinforcement. *Journal of Experimental Psychology: Animal Behavior Processes*, 34, 437–460.
- 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., & 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.
- 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. (1992). Behavioral variability and frequency-dependent selection. *Journal of the Experimental Analysis of Behavior*, 58, 241–263.
- 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.
- Maes, J. H. R., & van der Goot, M. (2006). Human operant learning under concurrent reinforcement of response variability. *Learning and Motivation*, 37, 79–92.
- 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.
- Morgan, L., & Neuringer, A. (1990). Behavioral variability as a function of response topography and reinforcement contingency. *Animal Learning & Behavior*, 18, 257–263.
- Neuringer, A. (1986). Can people behave "randomly"? The role of feedback. *Journal of Experimental Psychology: General*, 115, 62–75.
- 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., Deiss, C., & Olson, G. (2000). Reinforced variability and operant learning. *Journal of Experimental Psychology: Animal Behavior Processes*, 26, 98–111.
- Page, S., & Neuringer, A. (1985). Variability is an operant. *Journal of Experimental Psychology: Animal Behavior Processes*, 11, 429–452.
- Skinner, B. F. (1956). A case history in scientific method. *American Psychologist*, 11, 221–233.
- Skinner, B. F. (1961). The generic nature of the concepts of stimulus and response. In *Cumulative record* (enlarged ed., pp. 347–366). New York, NY: Appleton-Century-Crofts. (Original work published 1935)
- Stahlman, W. D., Roberts, S., & Blaisdell, A. P. (2010). Effect of reward probability on spatial and temporal variability. *Journal of Experimental Psychology: Animal Behavior Processes*, 36, 77–91.