Operant Variability: Procedures and Processes

Armando Machado and François Tonneau
Universidade do Minho

Barba's (2012) article deftly weaves three main themes in one argument about operant variability. From general theoretical considerations on operant behavior (Catania, 1973), Barba derives methodological guidelines about response differentiation and applies them to the study of operant variability. In the process, he uncovers unnoticed features of operant variability research (e.g., Neuringer, 2002) and proposes interesting modifications and extensions of current experimental practices. Barba's article calls for renewed attention to important issues, and we find merit in his proposal to evaluate operant variability by comparing response distributions along a common continuous measure. We are less convinced, however, by the conceptual underpinnings that he brings to the task.

Differentiation and operant behavior. First consider Barba's claim that "differentiation is the behavioral process that demonstrates an operant relation" (p. 000). There is a sense of *differentiation* in which this claim is trivially true. Operant behavior is demonstrated by choosing a response criterion R, delivering some consequence S for any activity that satisfies R, and noting that the prevalence of R changes as a result (Skinner, 1938). By definition of *operant reinforcement,* the prevalence of criterial activities (R) must increase during reinforcement, and the prevalence of at least another activity ($\sim$R) must

decrease because all activities compete for the same total time (Rachlin & Burkhard, 1978). If these changes in the relative prevalence of R and $\sim$R are all that we mean by *differentiation,* then it is obviously true that reinforcement entails differentiation. In this case, however, there is no need to evaluate probability distributions along a continuous measure (as in Barba's Figure 2) to demonstrate an operant relation. Any evidence that the prevalence of R increases (or, in the case of punishment, decreases) will do, together with control conditions to show that the observed changes are actually due to the correlation between R and S (Thompson & Iwata, 2005).

Alternatively, by *response differentiation* we may mean something more than the changes of response distribution that are a logical consequence of the definition of *reinforcement.* Here is how Galbicka (1988) characterizes response differentiation:

> Skinner's analysis of response differentiation in *The Behavior of Organisms* invoked three processes: reinforcement, extinction, and induction (generalization). Reinforcement increased the likelihood that a response, with its associated constellation of characteristics, would be repeated. Through induction, similar responses occurred even though never directly reinforced. … At some point, induction generated a response value outside the reinforced class. The effect of extinction was to decrease the rate of responses with these characteristics, and also to decrease (through induction) the rate of similar responses. The key to differentiation was that the direct effect of reinforcement or extinction was always considered greater than the indirect effect of induction. Hence, the two classes drew apart in frequency, as criterional responses continued directly to be reinforced and noncriterional ones continued to be extinguished. (p. 344)

In Galbicka's analysis, response differentiation involves operant reinforcement

as well as processes of response induction and extinction. Notice that to speak of extinction as a process (as opposed to extinction as a procedure), it is not enough to point out that during operant reinforcement noncriterial responses are not followed by S (which is of course true in any reinforcement situation). Rather, the prevalence of at least some noncriterial responses ($\sim$R) must actually decrease as a result of not being followed by S. It is extinction in this functional sense that produces a gradual sharpening of response distributions when the reinforcement criterion is held constant, and their gradual displacement when the criterion is increased or decreased across sessions (Skinner, 1938, Chapter 8). Ultimately it is the combined effect of reinforcement, induction, and extinction that allows ''the distribution of emitted responses … to conform *closely* [italics added] to the boundaries of the class of reinforced responses'' (Catania, 1998, p. 117).

On this alternative conception of differentiation, however, it is no longer true that ''differentiation is the behavioral process that demonstrates an operant relation.'' The behavioral process that demonstrates an operant relation is reinforcement, not differentiation, and it is certainly possible to have the former without the latter. Consider Skinner's (1938, p. 87) demonstration of a prolonged increase in response rate when lever pressing (R) was followed by a single food delivery (S) and then recorded in extinction. Here no differentiation was possible because after reinforcing only one variant of R (the other variants being never emitted, they were neither reinforced nor extinguished) all variants were extinguished nondifferentially. Yet the obvious increase in response rate was enough to demonstrate operant reinforcement, and therefore an operant relation, by comparison to a control condition in which S was

presented independently of R (Skinner, 1938, p. 81).

Even in situations that leave enough time for response differentiation, the latter will not take place if for some reason the response variants being reinforced and extinguished cannot follow separate courses. Figure 1, for example, shows the number of eye blinks per minute in one child affected by Trisomy 18 (Brownfield & Keehn, 1966). Frequency of blinking increased in sessions in which the child was spoonfed after each blink (filled symbols) and decreased in control sessions in which the child was fed independently of blinking (open symbols). These data are ample evidence for an operant relation between blinking and food. Yet we have no reason to believe that response differentiation in the sense of either Galbicka (1988) or Catania (1998) ocurred in this particular case. Although logically possible, differentiation was unlikely given the pervasiveness of the child's motor impairments. In organisms that are not motorically impaired, quantitative response differentiation can always fail if the reinforced and extinguished variants are too close to each other.

Now we do not deny that response differentiation is almost always present alongside operant reinforcement. The fact remains that reinforcement and differentiation are logically distinct and that demonstrating the latter is not necessary to document the former, hence to document an operant relation. What matters is comparing the levels of prevalence of the target property (R) under operant and nonoperant conditions (see above). Of course, regardless of the control conditions employed to study reinforcement, the interpretation of the results is an exercise in causal inference (Cartwright, 1981) and as such remains open to alternative interpretations (e.g., Gardner & Gardner, 1988).

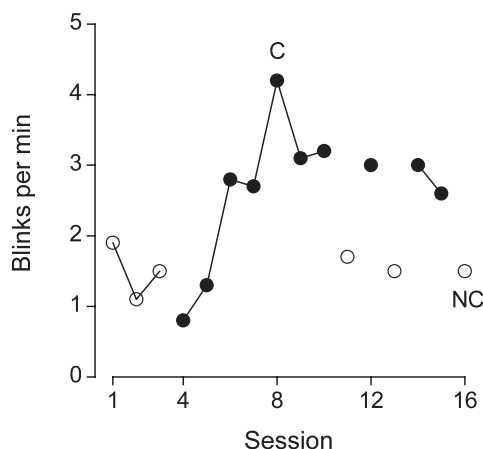*Mind the gap.* Even though we are skeptical of Barba's general conception

Figure 1. Number of eyeblinks per minute in one child affected by Trisomy 18. The filled symbols show response rate in sessions in which the child was spoonfed after each eyeblink (contingent condition, C). Empty symbols show response rate in control sessions in which food was delivered independently of blinking (noncontingent condition, NC). Data from Brownfield and Keehn (1966, Table 1, p. 414).

of operant behavior, his methodological criticisms of variability studies are well taken. It is true, for instance, that in lag *n* studies of variability the reinforcement criterion (S distribution) and the measure of its behavioral effects (the R distribution) differ from each other. Even though the S and R criteria may be correlated, the fact that they differ creates difficulties in interpreting the results. Admittedly, variability research is not unique in this respect. In Thorndike's (1898) classic studies of cats escaping a puzzle box, for example, the measure of reinforcement effects was not response rate but latency to escape. In this case, however, the conceptual gap between the hypothesized reinforcement process and its measurement can be filled by assuming that reinforced activities tend to occur earlier than the others or that shorter latencies reflect the progressive extinction of competing responses (Skinner, 1969). Whether our hypotheses are correct or not, a set of causal processes that lead from the S to the

R criterion can be formulated relatively easily.

In the case of operant variability studies, the gap from S to R distributions is not so easily filled. If the experimenter specifies Lag 25 as the reinforcement criterion (i.e., only sequences of pecks with recurrence times greater than 25 are reinforced), but measures its behavioral effects through the U statistic (itself a function of the steady-state probability distribution of all peck sequences), what processes relate the S to the R distributions (see Machado, 1992, 1993)? How does a lag-based criterion engender an observed U value? Several possibilities can be contemplated. All scientific theories take some phenomena as primitive or axiomatic, and others are derivative or emergent. In current behavior analysis, operant reinforcement is seen as a behavioral primitive, whereas contrast in multiple schedules, for example, is taken as the result of more elementary processes that act in combination (Williams, 2002). In other cases such as the matching law, whether the phenomenon should be taken as primitive or derivative remains uncertain (Williams, 1990). In the case at hand, the issue is to know whether operant variability documents an instance of operant conditioning similar to the conditioning of any other activity, or is a result reducible to more elementary processes. Neuringer and Jensen (2012) clearly favor the former hypothesis:

Variable responding is produced and maintained by reinforcers contingent upon it. Variability does not always decrease with learning, this being counter to initial theories of reinforcement. Of most importance, *particular levels of variability* are engendered by reinforcers contingent upon those levels. Variability is a dimension of behavior analogous to other operant dimensions, such as response rate, force, and topography. (p. 57)

One difficulty raised by this interpretation is the relation between stereotypy and variability. How can variability be reinforced if reinforcement

raises the probability of the responses that produce it, a process that necessarily increases stereotypy? Page and Neuringer (1985) addressed the issue with the following argument:

> Variability is as susceptible to control by reinforcement as are frequency, force, duration, location, or topography. … The following analogy may be useful: The pigeon enters the operant conditioning experiment with a class of behaviors described as pecking already intact. When the experimenter shapes key pecking, the pecking response is not being trained. Rather, the pigeon is taught where, when, and possibly how fast or hard, and so on, to peck. Analogously, there may be a dimension of all behaviors, described as variability, with which the organism enters our experiments. … Turning on or off a variability generator may be under the control of reinforcement, but the variability generator is not itself created through reinforcement. An animal may be born with the variability generator intact. (p. 450)

Clarifying the account. Two aspects of Page and Neuringer's (1985) account need to be clarified. First, Page and Neuringer do not deny that reinforcement can lead to behavioral stereotypy. On the contrary, they presuppose the stereotypy-increasing effect of reinforcement when dealing with response dimensions such as location or force. But if reinforcement can increase stereotypy as well as variability, under what circumstances does reinforcement favor one over the other? A complicating factor is that, unlike force or location, variability refers always to another dimension, as Barba points out. We are dealing not with *variability* in the abstract but with variability of another response property such response sequencing. If variability is a response dimension, then it is a second-order dimension, whereas force and duration, for example, are first-order dimensions. So the question becomes: How can the reinforcement of a response strengthen responses with equal values along its first-order dimensions (and thereby increase stereotypy) while simultaneously strengthening responses with equal values along a second-order dimension (and thereby increase variability)?

Second, if variability depends on the working of a variability generator, then what turns the generator on? Page and Neuringer (1985) answer the question by introducing an additional hypothesis. The operant conditioning of behavioral variability is said to involve a particular type of *discrimination*:

> It is advantageous for an animal to discriminate situations in which new responses must be learned from those in which previously learned behaviors must be repeated. We hypothesize that this discrimination is based on the reinforcement of diverse responses and response classes in the former case versus reinforcement of fixed, or stereotyped, responses and response classes in the latter. … When an animal is differentially rewarded for a variety of responses, it generates variable behaviors. (p. 449)

According to this account, then, reinforcement does not affect behavioral variability by making reinforced variants more probable. Instead, animals discriminate situations in which reinforcers follow repetitive behaviors from situations in which reinforcers follow novel or diverse behaviors. In the former case, stereotypy increases. In the latter case, an inborn variability generator is turned on and, as a consequence, behavioral variability increases. To explain why variability is higher under a Lag 25 schedule than under a Lag 5 schedule, the account further assumes that when the variability generator is turned on, it is attuned to the variability requirements imposed by the schedule (Neuringer, 1986, p. 74). [1]

A critique of the account. The foregoing account of operant variability faces difficulties. First, like Barba's appeal to differentiation as the basis of operant behavior, Neuringer's (1986) discrimination-based conception of reinforcement cannot deal with one-trial operant phenomena. If reinforcement depends on discrimination, then the animal needs

to experience at least two response-outcome pairs in order to discriminate whether repetitions or variations are being reinforced. To the extent that a single food delivery can result in reinforcement (Skinner, 1938, p. 87), Neuringer's account of the latter seems incorrect. His account may be salvaged by assuming two kinds of reinforcement, one that depends on discrimination and another that does not. But the issue then becomes to know how these two kinds interact.

Second, Neuringer's (1986) account of operant variability presupposes that animals can discriminate whether repetitions or variations are being reinforced. But what exactly is being discriminated in, say, a Lag 25 contingency? That the current sequence differs from all 15 (say) sequences emitted during the previous 25 trials or from only 10 of those sequences? Is the discrimination based on the entire sequence topography or only some of its parts (e.g., beginning of sequence, relative number of left and right pecks, number of switches and stays)? A discrimination-based account of operant variability will make different predictions depending on the answers it gives to these questions. But Neuringer's account gives none. A fortiori, it fails to explain how, concerning the discrimination required to turn on and tune the generator, a Lag 25 contingency differs from a Lag 5 contingency. Perhaps the reader will find it obvious that Neuringer's view predicts more variability under a lag of 25 than under a lag of 5. Our point is that this is not obvious at all.

Third, the account suggests (probably metaphorically) that animals "turn on" a random generator when they "discriminate" that reinforcement follows diverse responses. But consider a Lag 5 contingency. It seems odd that an animal capable of detecting that reinforcement follows "diverse responses" would generate as many as 30 different sequences in a 50-trial session when only six sequences, cycled through systematically, would suffice to obtain all available reinforcers (see Page & Neuringer, 1985, Figure 6). Arguing that memory limitations prevent the animal from cycling through the six sequences systematically is questionable because the initial discrimination implies no such limitations. More generally, the account grants pigeons significant memory powers to discriminate a high level of sequence variability (and tune the variability generator accordingly), but then denies them the same memory powers to vary a few sequences systematically. The account is, if not contradictory, at least implausible.

Looking for alternatives. Machado (1989, 1992, 1993, 1997) has sketched an alternative account, one that attempts to fill the causal gap between reinforcement (S) and response (R) distributions without denying that reinforcement strengthens the responses that produce it. The account starts by examining the various procedures used to condition response variability—procedures that reinforce least recent responses, sequences with higher recurrence times, novel responses, low-frequency sequences—and identifies a common thread: the presence of negative frequency-dependent selection (Blough, 1966; Shimp, 1967). More specifically, in all cases, reinforcement is more likely to follow the momentarily weakest activities and, conversely, extinction is more likely to follow the momentarily strongest activities.

When the environment implements negative frequency-dependent selection, the organism–environment interaction is strongly dynamic: As a response variant becomes weaker, it is differentially reinforced, which should strengthen it; but as a response variant becomes stronger, it is differentially extinguished, which should weaken it. At equilibrium, all variants are equally strong, and response variability is substantial. In

short, the combined effects of reinforcement and extinction, one strengthening the weakest responses and the other weakening the strongest responses, are logically sufficient to promote and maintain behavioral variability.[1] The various procedures used to condition variability differ in the proxy used to identify the momentary strength of the response variants (e.g., recency, frequency, novelty), how accurate that identification is, and how selectively they reinforce and extinguish the weak and strong response variants, respectively.

To illustrate, consider the typical lag procedure. All else being equal, a sequence's recurrence time varies inversely with its strength; the lag procedure implements negative frequency-dependent selection. But to promote variability, how high should the lag parameter be—1, 5, 25, or 50, for example? On the one hand, to extinguish the momentarily strongest sequences, the lag should be as large as possible, because the larger the lag the longer the extinction period. On the other hand, to reinforce the momentarily weakest sequences, the lag should be as small as possible, because the smaller the lag the longer the reinforcement period. It follows that the optimal lag value, that which

engenders the highest degree of variability, is unlikely to be too small (strong responses will not extinguish sufficiently) or too large (weak responses will not be reinforced, and responding itself will not be sustained for lack of reinforcement). This reasoning is consistent with Page and Neuringer's (1985) finding that sequence variability increased from Lag 5 to Lag 25 but decreased slightly for Lag 50.

With pigeons producing eight-peck sequences per trial, Page and Neuringer's (1985) Lag 25 procedure may have scored high on the two features necessary to promote and sustain behavioral variability: (a) how effectively the schedule extinguishes the momentarily stronger sequences (biases), and (b) how effectively the schedule reinforces the momentarily weaker responses. In contrast, the yoke procedure scores low on both features, because although it includes the same periods of extinction and reinforcement, those periods are not selectively coupled with the strongest and weakest sequences, respectively. More specific predictions will need to take into account how fast the sequences lose and gain strength when extinguished and reinforced, respectively.

CONCLUSION

Barba calls our attention to a major discrepancy in studies of operant variability, that between the reinforcement criterion and the measurement of its behavioral effects. But we disagree with the reason Barba invokes to lament the discrepancy. Differentiation and discrimination are not primitive processes; they derive from reinforcement, extinction, and generalization. The main reason for lamenting the discrepancy is that it reveals a major conceptual gap in our account of operant variability, namely, how high degrees of variability (e.g., measured by U scores) result from the differential

---

[1] To ask how reinforcement can increase behavioral variability is like asking how natural selection can increase genetic variability. Biologists have long recognized that when the fitness of a phenotype depends inversely on its frequency in the population (i.e., negative frequency-dependent selection), genetic variability may be promoted and maintained. The classic example is Fisher's (1930) theoretical account of the sex ratio, but Clark (1979) also stressed the importance of frequency-dependent selection to explain polymorphisms (behavioral and otherwise) in parasite–host and predator–prey systems. For experimental studies of frequency-dependent selection, see Gigord, Macnair, and Smithson (2001) as well as Hori (1993). Machado, Keen, and Macaux (2008) further explored the analogy between reinforcement and natural selection to understand the acquisition of preference in concurrent schedules.

reinforcement of sequences with high recurrence times. Neuringer's (1986) conception of operant variability suffers from various limitations and is unlikely to fill the gap. And we have sketched an alternative, behavior-analytic account that may be able to close the gap: Reinforcement may promote behavioral variability through frequency-dependent selection. Whether this account is adequate remains to be seen through further empirical and theoretical analyses.

REFERENCES

Barba, L. S. (2012). Operant variability: A conceptual analysis. *The Behavior Analyst*, *35*, 000–000.

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

Brownfield, E. D., & Keehn, J. D. (1966). Operant eyelid conditioning in trisomy-18. *Journal of Abnormal Psychology*, *71*, 413–415.

Cartwright, N. (1981). The reality of causes in a world of instrumental laws. In P. D. Asquith & R. N. Giere (Eds.), *PSA 1980* (Vol. 2, pp. 38–48). East Lansing, MI: Philosophy of Science Association.

Catania, A. C. (1973). The concept of the operant in the analysis of behavior. *Behaviorism*, *1*(2), 103–116.

Catania, A. C. (1998). *Learning* (4th ed.). Upper Saddle River, NJ: Prentice Hall.

Clark, B. C. (1979). The evolution of genetic diversity. *Proceedings of the Royal Society of London, Series B*, *205*, 453–479.

Fisher, R. A. (1930). *The genetical theory of natural selection*. Oxford, UK: Oxford University Press.

Galbicka, G. (1988). Differentiating *The Behavior of Organisms*. *Journal of the Experimental Analysis of Behavior*, *50*, 343–354.

Gardner, R. A., & Gardner, B. T. (1988). Feedforward versus feedbackward: An ethological alternative to the law of effect. *Behavioral and Brain Sciences*, *11*, 429–493. (Includes commentary)

Gigord, L. D. B., Macnair, M. R., & Smithson, A. (2001). Negative frequency-dependent selection maintains a dramatic flower color polymorphism in the reward-less orchid *Dactylorhiza sambucina* (L.) Soò. *Proceedings of the National Academy of Sciences*, *98*, 6253–6255.

Hori, M. (1993). Frequency-dependent natural selection in the handedness of scale-eating cichlid fish. *Science*, *260*, 216–219.

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. (1993). Learning variable and stereotypical sequences of response: Some data and a new model. *Behavioral Processes*, *30*, 103–130.

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.

Machado, A., Keen, R., & Macaux, E. (2008). Making analogies work: A selectionist model of choice. In N. K. Innis (Ed.), *Reflections on adaptive behavior: Essays in honor of J. E. R. Staddon* (pp. 23–50). Cambridge, MA: MIT Press.

Neuringer, A. (2002). Operant variability: Evidence, functions, and theory. *Psychonomic Bulletin & Review*, *9*, 672–705.

Neuringer, A., & Jensen, G. (2012). The predictably unpredictable operant. *Comparative Cognition & Behavior Reviews*, *7*, 55–84. doi:10.3819/ccbr.2012.7004

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

Rachlin, H., & Burkhard, B. (1978). The temporal triangle: Response substitution in instrumental conditioning. *Psychological Review*, *85*, 22–47.

Shimp, C. P. (1967). Reinforcement of least-frequent sequences of choices. *Journal of the Experimental Analysis of Behavior*, *10*, 57–65.

Skinner, B. F. (1938). *The behavior of organisms: An experimental analysis*. New York, NY: Appleton-Century.

Skinner, B. F. (1969). *Contingencies of reinforcement: A theoretical analysis*. New York, NY: Appleton-Century-Crofts.

Thompson, R. H., & Iwata, B. A. (2005). A review of reinforcement control procedures. *Journal of Applied Behavior Analysis*, *38*, 257–278.

Thorndike, E. L. (1898). Animal intelligence: An experimental study of the associative processes in animals. *Psychological Review Monograph Supplement*, *2* (4, Whole No. 8).

Williams, B. A. (1990). Enduring problems for molecular accounts of operant behavior. *Journal of Experimental Psychology: Animal Behavior Processes*, *16*, 213–216.

Williams, B. A. (2002). Behavioral contrast redux. *Animal Learning & Behavior*, *30*, 1–20.

Authors Queries

Journal: **The Behavior Analyst**
Paper: **bhan-35-02-11**
Title: **Operant Variability: Procedures and Processes**

Dear Author
During the preparation of your manuscript for publication, the questions listed below have arisen. Please attend to these matters and return this form with your proof. Many thanks for your assistance

Query Reference	Query	Remarks
1	Add Neuringer 1986 to reference list.	