The Behavior Analyst 2012, 35, 237–241 No. 2 (Fall)

Operant Variability: Some Random Thoughts

M. Jackson Marr
Georgia Tech

Barba's (2012) paper is a serious and thoughtful analysis of a vexing problem in behavior analysis: Just what should count as an operant class and how do we know? The slippery issue of a "generalized operant" or functional response class (e.g., imitating, "relating" as in relational frame theory, attending, rule following, grammatical frames, and many other possibilities—one wonders when it will all end) illustrates one aspect of this problem, and "variation" or "novelty" as an operant appears to fall into this category. Given our traditional and fundamental operant approach to response differentiation through shaping to establish a specifiable functional response class, how do we understand "variability" as a response class? As Barba's review attests, this issue has been discussed for more than two decades without, I believe, significant resolution. What we do know is that degrees of variations in patterns of responding can be brought about by imposing special contingencies.

As with any generalized operant, no particular response can be said to belong to such a class. In the case of "random" responding, only in relation to perhaps many previous responses as a sample drawn from a probability distribution function (typically a uniform distribution) does a given response count as an exemplar. Indeed, as with a series of coin tosses, many trials must occur to provide any suitable criterion for a random sequence. It would make no sense to say whether a given heads or tails outcome belonged to a random distribution. As with this binomial example, there are many tests for determining the degree of "behavioral disorder" in a long sequence of putatively independent trials, of which the U value is common, a measure ranging from 0 (*complete order*) to 1 (*complete disorder*). In typical studies, reinforcers are presented for varying a set-length sequence of responses, say, on two different keys such that a given sequence (comprised of left vs. right responses) differs from the n previous sequences—the lag n procedure. For example, if the trial sequence length, s, is set at $s = 4$, then there are 16 possible (2^s) different sequences of left (L) and right (R) responses: LRLR, LLLL, RLLR, and so on. If reinforcement was delivered only for an emitted sequence that differed from the last 15, then all the possible sequences must have been emitted in the total of 16 trials (by chance an extremely unlikely event), yielding a U value of 1 for that series block. If the same sequence was emitted (e.g., RLRL, also by chance an extremely unlikely event) for 16 consecutive trials, the U value would be 0 for that series block. In essence, most of these studies are variations of "don't do what you did recently." In that sense, *variation* may in no way be synonymous with *random*.[1] Assuming each sequence quartet is an

Correspondence may be addressed to M. Jackson Marr, School of Psychology, Georgia Tech, Atlanta, Georgia 30332 (e-mail: mm27@prism.gatech.edu).

[1] Although *variation*, expressed as statistical variance, has a precise definition, *random* does not. As Havil (2002) notes, " Randomness is a very, very subtle concept with its properties belonging to statisticians more than mathematicians" (p. 229). The National Institute of Standards and Technology lists 16 current tests for randomness, and there are at least a dozen more; a given sequence may pass one test and not another one.

independent event, as each is emitted, the probability of emitting a different sequence from all previous ones can rapidly get less and less. The probability model is the traditional numbered balls in an urn (e.g., Feller, 1968) where each four-response sequence is represented by a ball with a different number, say, from 1 to 16. The problem is to determine in a series of n draws ($n \leq 16$) with replacement, what is the likelihood of drawing out a series of balls with all different numbers. Alternatively, what is the likelihood of drawing out a series each containing only a single number? In general, as n approaches 2^s, the probability of either of these two cases can become vanishingly small. Think of tossing a die in many blocks of six throws each and always getting the same value, or that each of the different values (1 to 6 in any order) occurred in each block! Such an outcome would be a strange sort of "disorder." Of course, as the lag n requirement is set progressively less than 2^s, the possibility of emitting a set of unique sequences improves; nevertheless, that contingencies can control emitted variations in sequences is impressive, but they may be far from random, and maybe that's the goal. From the point of view of chance, through differential reinforcement, the behavior tends to go from one set of unlikely events (sequences tend to be repeated) to another (sequences tend not to be repeated).

Barba bases his approach to response differentiation and the acquisition of a functional response class on the treatment of Catania (1973a). But Catania (1973b) also discussed this issue in detail in a chapter of the classic *The Study of Behavior*, and most recently in *Learning* (2007), using essentially the same examples. Catania (1973b) defined a functional response class this way:

> With respect to any response class, it is necessary to ask a fundamental behavioral question: Can the likelihood of responses in this class be modified by their consequences? If so, the class is called an *operant* class; it is a class of responses that is affected by the way in which it operates on the environment. (p. 53)

Presumably, modification by consequences is determined by a demonstrable order in the pattern of changes, for example, in the distribution of responses (or some property of the responses) related to reinforcement delivery (e.g., Skinner, 1938). But Catania's later presentations differ somewhat in detail from his *Behaviorism* paper. For example, in Catania (1973b) and (2007), there is no reference to a distribution of stimuli, that is, a distribution of conditional probabilities of reinforcement, or what Barba calls "S distribution." At least as described, I find this description a bit troublesome. The term *distribution* seems unclear, at least as applied to the cases under consideration. Is it, for example, a density function of a random variable? For example, if all lever presses of force less than $\boldsymbol{F}$ are never reinforced, while all lever presses greater than $\boldsymbol{F}$ are, what makes that a "distribution" except in a trivial sense? Using this as an example, here is a possible reformulation: We determine an initial distribution of response forces, R_i. That is, we determine the relation between response forces and their frequency per unit time. This distribution may simply be an "operant level" or may have resulted from some earlier explicit contingency. We now apply a new contingency to this initial distribution that, typically, will ultimately result in shift to a final distribution, R_f. For simplification, let's assume these two distributions are continuous; we could (and normally would) divide the forces into classes, but this doesn't affect the analysis.

The applied contingency (criterion for reinforcement) can be thought of in engineering terms as a forcing function (reinforcement has been analogously described in this way

before, e.g., Marr, 1992), or, as a mathematician might describe, an "operator" or "transform," operating on R_i to yield R_f. How do we assess the effects of this operation? We want to somehow measure the difference between R_f and R_i in relation to the applied contingency. Thus, simply taking the difference in total areas under the two distributions is not sufficiently precise. Assume we set our minimal force requirement to $\boldsymbol{F}_c$: All lever presses with a force equal to or exceeding $\boldsymbol{F}_c$ will be reinforced. $\boldsymbol{F}_c$ then is a point along the x axis that represents response forces ranging from 0 to some maximal force $\boldsymbol{F}_{max}$ (R_f), the highest value ever achieved in the R_f distribution. Initially, many response forces will be less than $\boldsymbol{F}_c$, but occasionally, responses in the R_i distribution may exceed $\boldsymbol{F}_c$ without the applied contingency. That is, part of the R_i distribution may encompass a range from $\boldsymbol{F}_c$ to $\boldsymbol{F}_{max}$ (R_i), the highest force achieved in the R_i distribution. The areas of interest then are for R_f: the range from $\boldsymbol{F}_c$ to $\boldsymbol{F}_{max}$ (R_f), and for R_i: the range from $\boldsymbol{F}_c$ to $\boldsymbol{F}_{max}$ (R_i). The effect of the contingency is measured by the difference between these two regions:

$$\int_{Fc}^{Fmax\,(Rf)} Rf\,dx - \int_{Fc}^{Fmax\,(Rf)} Ri\,dx$$

As already mentioned, Barba's S distribution is very limited in scope and typically is of the form (using response force as an example and S^R representing a reinforcer delivery):

$$P\left(S^R \middle| \boldsymbol{F} > \boldsymbol{F}_c\right) = 1$$

and

$$P\left(S^R \middle| \boldsymbol{F} < \boldsymbol{F}_c\right) = 0.$$

But instead of this "jump discontinuity" distribution, one could arrange, say, a "ramp" where the probability (or magnitude) of reinforcement increases linearly with force. This represents a contingency as a *feedback function* (e.g., Baum, 1981; Marr, 2006) that can dynamically tie ongoing behavior more closely to its differential consequences. Many possibilities for exploring response differentiation reside here, and I'll return to one shortly in the context of training variation in responding.

Whether one adopts Catania's description, or the above, or something else, these approaches are limited to very special cases (and I think this applies to the typical variability study) in which a shift in distributions is along some common continuous dimension like force, or discrete as with a set of response sequences. In far more interesting cases of response differentiation, the target response class and the initial response class may be utterly different, for example, successfully training a bear to ride a motorcycle, as they are said to do in Russian circuses.

The results of many such experiments demonstrate that variability in response sequences (as commonly measured by whole-session U values) can be engendered by such contingencies; that is, behavior shifts in the direction of greater variability in conformity to the imposed requirements. But as Barba points out, this raises the question of the functional response class, that is, measured variability in sequences of responses might arise as *indirect* effects of reinforcement selection of other features of responding. Barba cites Machado (1997), for example, whose studies showed that variation in response sequences can arise by differential reinforcement of switching between keys. Thus, the nominal criterion of lag n sequence differences by itself does not provide sufficient evidence of variability per se as a functional response class, never mind the putative role of an "endogenous stochastic generator" (Neuringer, 2002, 2004). I'll say more about this concept later.

To address the question of functional control of variation, Barba proposes a contingency based on cumulative U value as defining the conditions of reinforcement. I admit I'm not entirely certain of just how this would work in practice, but assuming such a contingency could be arranged so that when the cumulative U value reached some minimal value, reinforcement would be delivered, I have two reactions: First, we would still be left with the question of *how* such control is achieved; how does the imposed contingency actually shape the final performance from some starting distribution? There is an implication here that achieving a minimal cumulative U value is somehow a more direct measure defining the performance, but how would we know that this outcome had not been achieved indirectly?

A second issue involves the typical contingency arrangements alluded to earlier in discussing possibilities for defining what Barba called "S distributions" to control response differentiation. In lag n procedures, along with others, including the proposed minimal cumulative U-value contingency, there is a threshold below which reinforcement is not delivered and above which it is, what I earlier called a jump-discontinuity contingency. As a possible alternative to Barba's cumulative U-value procedure, a more nuanced feedback contingency could be arranged wherein the probability of reinforcement depended on some monotonic function of variation in emitted sequences. To take a simple example, in a nominal lag n procedure reinforcement depends on emitting, say, a four-response sequence different from the last $n - 1$ sequences; otherwise no reinforcement is delivered. But let us set the probability of reinforcement $p = 1$ for achieving this lag n requirement and a progressively decreasing probability for N different sequences $< n - 1$ (setting $p = 0$ for a just repeated sequence); in other words, a decreasing reinforcement probability as a function of the likelihood by chance of having emitted a sequence not previously emitted in a block of n. Again, such a dynamic contingency could tie behavior more closely to its consequences and, if properly studied, may offer some possibilities of seeing how such a performance is acquired.

The question of how operant variability is acquired has been the subject of some considerable experimental and theoretical efforts, as Barba has reviewed. The role of some kind of memory process has also received considerable attention, and the results differ depending on the imposed contingencies. The hallmark of random events is independence of trials as in a binomial process (whether biased or not). Then, perhaps naively, one can hypothesize that some effective variation-training contingencies shape up a "memory-less" system (Neuringer, 1991, has explored this possibility). For example, one might come to emit a random series of digits by "forgetting" the digit just emitted.[2] We know that "directed forgetting" is trainable (e.g., Zentall, Roper, Kaiser, & Sherburne, 1997); maybe something like this emerges as a feature of some of the random-response training procedures. There are conceptual problems with the no-memory idea (see, e.g., Machado, 1992); but, more important for present purposes, this cannot apply to the typical lag n procedures. The performances attained under these contingencies may be far from what chance would predict; as already mentioned, *variation* and *random* are not synonymous.

Given that, Neuringer's (2002, 2004) proposed "endogenous stochastic generator" may be a misnomer, but in my view that is the least

[2] Of course, this assumes all possible alternatives would be emitted with equal probability, like tossing a fair coin which has no "memory" of its previous outcome. Otherwise, *having no memory* could mean simply emitting the same response repeatedly.

of its problems. This appears to be a "mechanism" as name, in the tradition of Chomsky's "language acquisition device," an invention with just the properties needed to explain the findings, but with little, if any, predictive power. At best, empirical findings are said to be "consistent" with the notion. What turns such a generator on or off? How does the operation of such a generator lead to behavior? Where is it located? Is there an "endogenous deterministic generator" as well? Surely, much data would be consistent with that device. Regarding the lag *n* procedures, Machado (1992) has suggested a different mechanism: frequency-dependent selection based on differential reinforcement of infrequent behavior patterns. It is not clear to me that this concept does not share the same conceptual problems of Neuringer's stochastic generator.

In puzzling over the many mysteries occasioned by the methods and results of research devoted to "variability as a generalized operant class," I've come to see this area as representative of how little we know or understand about the complex dynamics of behavior. Perhaps the greatest challenge in developing behavior analysis as a natural science, and in vivid contrast with other natural sciences, is the lack of any clear notion about units of analysis: What does it mean, if anything, to talk about units of behavior? If such an effort is meaningful, then how do such units come about and, most mysteriously, how do they change to become new units? Such questions appear to provide the largely wasted heat driving the silly arguments pitting "molecular" against "molar" accounts of behavior, as if we really understood what we were talking about in either case. The remarkably seductive and flexible concept of the "generalized operant" (a kind of "super-unit") seems infinitely extendable, making it difficult *not* to apply it, as needed, to account for virtually any behavior we deem fit.

REFERENCES

Barba, L. S. (2012). Operant variability: A conceptual analysis. *The Behavior Analyst*, *35*, 213–227.

Baum, W. M. (1981). Optimization and the matching law as accounts of instrumental behavior. *Journal of the Experimental Analysis of Behavior*, *36*, 387–403.

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

Catania, A. C. (1973b). The nature of learning. In J. A. Nevin & G. S. Reynolds (Eds.), *The study of behavior* (pp. 30–68). Glenview, IL: Scott Foresman.

Catania, A. C. (2007). *Learning* (4th interim ed.). Cornwall-on-Hudson, NY: Sloan.

Feller, W. (1968). *An introduction to probability theory and its applications* (Vol. 1). New York, NY: Wiley.

Havil, J. (2012). *The irrationals*. Princeton, NJ: Princeton University Press.

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 keys. *Journal of the Experimental Analysis of Behavior*, *68*, 1–25.

Marr, M. J. (1992). Behavior dynamics: One perspective. *Journal of the Experimental Analysis of Behavior*, *57*, 249–266.

Marr, M. J. (2006). Food for thought on feedback functions. *European Journal of Behavior Analysis*, *7*, 181–185.

Neuringer, A. (1991). Operant variability and repetition as functions of interresponse time. *Journal of Experimental Psychology: Animal Behavior Processes*, *17*, 3–12.

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

Neuringer, A. (2004). Reinforced variability in animals and people. *American Psychologist*, *59*, 891–906.

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

Zentall, T. R., Roper, K. L., Kaiser, D. H., & Sherburne, L. M. (1997). A critical analysis of directed-forgetting research in animals. In J. M. Golding & C. MacLeod (Eds.), *Approaches to intentional forgetting* (pp. 285–287). Hillsdale, NJ: Erlbaum.