Controlling and Predicting Unpredictable Behavior

Lourenço de Souza Barba^{1,2}

Published online: 18 September 2014

© Association for Behavior Analysis International 2014

Abstract Behaving predictably can be advantageous in some situations, but unpredictability can also be advantageous in some competitive situations like sports, games, and war. Can, however, unpredictable behavior be conditioned? If a contingency of reinforcement based upon the predictability of behavior generates unpredictable responding, is it possible to conclude that predictability is itself a reinforceable dimension of behavior? In this paper, I address these questions by examining the concept and measures of predictability and the procedures generally used to increase unpredictable responding. I discuss the hypothesis that contingencies based on response frequency shape the generalized operant “to vary” and an alternative hypothesis that such contingencies generate unpredictable responding by balancing the strength of each alternative response over time. I discuss the findings that support the balance hypothesis as well as its limitations. I conclude that the two alternative hypotheses may be complementary in explaining unpredictable responding.

Keywords Behavioral predictability · Operant variability · Lag n procedure · Threshold procedure · U value

Some have argued that a science of behavior should provide prediction and control over its subject matter (Skinner 1953, pp. 3–10). However, under some conditions, those whose behavior is predictable are at a disadvantage. Fighting, war, games, hunting, and competitive sports are some conditions in which unpredictable behavior may have an advantage over predictable behavior (Bryant and Church 1974; Neuringer 2002; Neuringer and Jensen 2012). But, does this imply that organisms can learn to behave unpredictably?

In what follows, I will examine contingencies of reinforcement used to control unpredictable responding and the data that led some researchers to conclude that

✉ Lourenço de Souza Barba
lourenbarba@yahoo.com.br; lsbarba@uol.com.br

¹ Centro Universitário Padre Anchieta, Jundiaí, Brazil

² R Braseliza Alves de Carvalho, 522, São Paulo, Brazil 02510-030

unpredictable responding can be reinforced. Next, I will discuss an alternative interpretation according to which the effects of such contingencies can be explained without assuming that unpredictability is itself a reinforceable dimension of behavior. First, I will briefly present the concept and measures of unpredictable responding.

Predictability as a Dependent Variable

Suppose that an urn is filled with two blue balls and two red balls (the balls are identical in all attributes except their colors). Suppose also that someone blindly selects balls one at a time and every selected ball returns to the urn after each selection. Suppose finally that one composes a binary sequence with the elements R and L. The element R is chosen to occupy the next position in the sequence only if the selected ball is red, and the element L is chosen if the selected ball is blue. The sequence generated by this process shows two remarkable properties: (1) if the sequence is long enough, it is likely to contain approximately the same number of the elements R and L and (2) there is no sequential dependence between the elements constituting the sequence. That is, the preceding events in the sequence do not affect the probability that an R or an L will occupy the subsequent positions in the sequence. Because elements R and L occur in approximately the same proportion and there is no sequential dependence, the sequence is said to be a random sequence of the events L and R (Neuringer 1986).

If the urn contains three red balls and one blue ball, the element R will correspond to approximately 75 % of the elements in a long sequence. That is, this sequence contains the element R in larger proportion than the element L, but there is still no sequential dependence between the elements in this sequence. The sequence LRLRLR (a systematic and simple alternation between the elements R and L) shows a clear sequential dependence between the responses (every R is followed by an L and vice versa). This sequence, however, contains the same number of elements R and L. Finally, the sequence LRRRLRRRLRRR contains the element R in larger proportion (it corresponds to 75 % of all sequence elements) and displays a clear sequential dependence (every L is preceded by and followed by the run RRR and vice versa). Table 1 summarizes these four different patterns that can be distinguished in the binary sequence.

If a prediction is based only on the frequency distribution of the alternative elements, the next element in the binary sequence is more predictable in pattern 1 than in pattern 2. This is because the frequency distribution of the elements is perfectly flat in pattern 2

Table 1 The four distinct possible patterns that can be distinguished in a binary sequence made up of the elements L and R

With sequential dependence	R response occurs in larger proportion than L	R and L responses occur in the same proportion
	RRRLRRRLRRRLRRRL (pattern 1)	RLRLRLRLRLRLRLRLRL (pattern 2)
No sequential dependence	Sequence blindly generated from the urn with three red balls and one blue ball (pattern 3)	Sequence blindly generated from the urn with two red balls and two blue balls (random pattern)

(each element occurs in proportion equal to 0.5), but it is peaked in pattern 1 (note that both patterns show sequential dependence). It means that the flatness of the frequency distribution of the alternative elements can be used as a measure of predictability. If the prediction is based only on the previously emitted responses, the next element in the binary sequence is more predictable in pattern 1 than in pattern 3 because only the former shows clear sequential dependence (although the frequency distribution of the alternative elements is equally peaked in both patterns). Thus, sequential dependence can also be used as a measure of predictability for the elements in a binary sequence.

Many studies have measured both the flatness of the frequency distribution of the alternative responses and sequential dependence in binary sequences of responses generated under contingencies of reinforcement (Machado 1992, 1993). Most studies that I will discuss investigated operant variability (reinforcement was contingent on variable responding). Reinforcement of variability has led to increases in the flatness of the frequency distribution of the alternative responses and/or decreases in sequential dependence between the alternative responses (Page and Neuringer 1985; Machado 1992) and hence decreased predictability. Thus, in these studies, changes in variability of responding were taken as synonymous with changes in predictability of responding.

U value is a measure usually employed in studies investigating operant variability (Neuringer 2002). U value measures the entropy or uncertainty of a system (Attneave 1959). U value reflects flatness of the frequency distribution of the units composing a response pattern (individual responses or combinations of individual responses taken as units of analysis). The general formula of the U value is

$$-\sum_{i=1}^N \frac{[pi.\log_2(pi)]}{\log_2(N)},$$

where N is the number of different alternative individual responses (or combinations of individual responses) and pi is the proportion in which each different alternative response (or each combination of individual responses) occurs in the observed pattern. U value assumes its maximum value (1.0) when the subject emits all the alternative responses with the same frequency (the frequency distribution of the alternative responses is completely flat), and it assumes its minimum value (0.0) when the subject emits a single response (the frequency distribution of the alternative responses is completely peaked).

Sequential dependence is a second type of dependent variable in studies on operant variability. Lag analysis, for example, compares the unconditional probability that an L response occurs on trial n with the probability that an L response occurs on trial n given that an L response occurred on trial $n-1$ or given that an L response occurred on trial $n-2$, and so on. Probabilities are calculated from the proportions in which each individual response, and combinations of them (pairs, triplets, quartets, and so on), occurred in trials preceding trial n . In the simple alternation sequence (LRLRLR), for example, the unconditional probability that an L response occurs equals 0.5 (the proportion of L responses), but the probability that an L response occurs on the trial n given that an L response occurred on the trial $n-1$ equals zero. In a long sequence with no sequential dependence, conditional probabilities equal the unconditional probability that a particular event occurs. Statistical tests based on Markov chains also indicate sequential dependence (Machado 1992, 1993, 1994).

Although U value essentially measures distributional flatness at a given level of analysis, differences in U values calculated at different levels can indicate sequential dependence (Attneave 1959). Conversely, sequential dependence measures can provide information about the flatness of the frequency distribution of individual responses (or combinations of individual responses) constituting the observed pattern.

Contingencies Based on Response Frequency and Predictability

Three general types of contingencies have been used to generate variable responding: frequency dependent, lag n , and threshold contingencies. Under frequency-dependent contingencies, also called contingencies based on response frequency, a response produces reinforcement only if it occurred infrequently enough in the recent past. Such contingencies are said to promote a negative frequency-dependent selection (Machado and Tonneau 2012).

Machado (1992), for example, instituted a contingency based on response frequency using pigeons and differentially reinforced pecks across two response keys. There were two individual alternative responses: peck on the left key (L) or peck on the right key (R). Each trial consisted of a single response. Machado differentially reinforced in trial n the individual response (L or R) that was the least frequent in the trials preceding trial n . The data showed that pigeons developed a stable pattern of simple alternation between the keys (something like LRLRLRLR, a pattern which contains R and L responses in the same proportion). Machado (1993) differentially reinforced triplets of individual responses that were infrequent in the past trials. The triplets were counted according to a three-response window that moved forward every time an individual response occurred. As the frequency of the triplet LLL increased, the probability of reinforcement for emitting the triplet LLL decreased (the same holds for the other triplets). Some birds developed a pattern of responding that approached a random-like pattern. Machado (1992, 1993) concluded that when infrequent individual responses were negatively selected, the subjects were likely to develop a stable and systematic response pattern. When, however, the contingency targeted higher order combinations of individual responses (like triplets of individual responses), the corresponding stereotyped pattern became too complex, and the subject then developed a random-like distribution of responses across the keys.

Lag n is another contingency used to increase variability of responding. Lag n is a contingency based on response frequency in which a response produces reinforcement only if it did not occur (i.e. occurred with frequency equal to zero) in the recent past. Using a lag n contingency, Page and Neuringer (1985) required subjects to emit a sequence of eight pecks across two response keys (a left key and a right key). A sequence of eight pecks like LLLLLRRR (four pecks on the left key followed by four pecks on the right key) constituted a trial and was reinforced only if it differed from the sequences emitted in the n previous trials. Differing meant showing a different number or pattern of L and R responses. The sequence LLLLLRRR differs from the sequence LLLLLRRRL, for example. If the six initial emitted sequences in a lag-2 contingency were LLLLLRRR, LLLLLRLL, LLLLLRRR, LLLLLRRR, LLLLLRLL, LLLLLRRR, only the first, the second, and the fifth sequences would be reinforced because only those sequences do not repeat the sequences emitted in the previous

two trials. Page and Neuringer (1985) manipulated the parameter of a lag n contingency (the n value) in a within-subject design and found that higher n values consistently produced higher U values. The lowest U values were obtained when sequence variability was permitted but was not required. In this case, responding tended to be minimally variable.

A third contingency used to increase response variability is a threshold contingency. It is a contingency based on response frequency in which a response sequence produces reinforcement only if its relative frequency in the previous trials does not exceed a threshold value chosen by the experimenter. If the threshold value is set at 0.25, for example, the sequence A produces reinforcement in the trial n only if it occurred in a proportion less than or equal to 0.25 in the trials preceding trial n . This criterion holds for any sequence. Many studies investigating operant variability employed a threshold contingency (e.g. Denney and Neuringer 1998¹; Doughty and Lattal 2001; Grunow and Neuringer 2002; Doughty et al. 2013).

Denney and Neuringer (1998) investigated whether variable responding could be brought under stimulus control by implementing a multiple reinforcement schedule. During the Vary component, the threshold contingency differentially reinforced infrequent four-response sequences, and during the Yoke component, the probability of reinforcement was yoked to the proportion of reinforced trials that the subject produced during the Vary component (i.e. during the Yoke component, reinforcement did not depend on sequence frequency). Denney and Neuringer (1998) adopted the U value as a measure of variability. In this case, the U value was calculated from the proportions in which each different sequence was emitted during an experimental session. The higher the U value, the flatter the frequency distribution of the alternative sequences. U values obtained under Vary components were higher than U values obtained under Yoke components. Similar results were obtained by other researchers (Doughty and Lattal 2001; Odum et al. 2006; Ward et al. 2008). Many studies investigating variability in sequences of two alternative responses calculated the U value from the proportions in which each different sequence was emitted (Neuringer 1991; Hunziker et al. 1996; Doughty and Lattal 2001; Grunow and Neuringer 2002; Odum et al. 2006; Ward et al. 2008; Doughty et al. 2013).

Grunow and Neuringer (2002) used a threshold contingency and manipulated the threshold value. Grunow and Neuringer required rats to emit three-response sequences on three operanda. The subjects were divided into four groups. Each group was exposed to a particular threshold value. The results revealed that lower threshold values consistently produced higher U values. Thus, when reinforcement was contingent on higher variability of responding, higher variability was produced. Findings such as these, and those described above, showing that variability of responding is sensitive to changes in contingencies and is susceptible to stimulus control indicate that variability is a reinforceable dimension of responding. According to Neuringer and Jensen (2012), “*response unpredictability can be reinforced*” (p. 513, italics added).

¹ In fact, following any reinforced sequence, the number of times each sequence occurred was multiplied by a weighting coefficient before the relative frequencies were calculated. This procedure made it possible that sequences emitted recently were weighted more than past sequences in establishing relative frequencies used as reinforcement criteria according to Denney and Neuringer (1998).

Variability as an Operant

If a contingency differentially reinforces lever press responses, it shapes and maintains the functional class *lever press* (an operant). Each instance comprising this functional class differs from other instances in its formal properties (topography, duration, force, location, inter-response time, and so on), but all instances are functionally equivalent. Neuringer (2002) claims that instances “emerge stochastically” from within the functional class (p. 697). If the contingency explicitly requires a specific topography (i.e. if only responses with that specific topography produce food), lever press topography becomes highly predictable (highly stereotyped). If, conversely, the contingency explicitly requires topographies differing from the topographies of the n previous responses (lag n contingency), topographies become more variable (or unpredictable).

According to Neuringer (2002), size and “within-classes probability distribution” (p. 697) are two attributes of functional response classes that allow one to conceptualize behavioral variability that is sensitive to differential reinforcement. Size refers to the number of members composing the class, and probability distribution refers to the probability that each member of the class is emitted. The broader the class (i.e. the higher the number of topographically distinct members composing it) and the flatter the within-class probability distribution, the more variable the responding. Contingencies based on response frequency control the size of the class and the flatness of the within-class probability distribution, according to Neuringer’s (2002) view.

Thus, under a lag n contingency (Page and Neuringer 1985) and under a threshold contingency (Grunow and Neuringer 2002), subjects learn to emit sequences of eight and three responses respectively *and* learn to emit them with configurations (number and position of L responses and number and position of R responses) that are as variable as the contingencies required (Weiss and Neuringer 2012). Therefore, not only is configuration a behavioral property sensitive to reinforcement (contingencies can select sequences containing the same number of L and R responses, for example), but the variability of configuration is also a behavioral property sensitive to operant control. Because the amount of variability is a function of the contingency requirement, it is possible to control and predict the variability levels of the formal properties characterizing the responses of an operant class.

Neuringer (2012) identifies the operant “to vary” with other generalized operants like imitation. In imitation training each reinforcer follows a particular response (a response with particular topography). The training is successful, however, not if responses with particular topography become more likely but if responses with topographies similar to the model become more likely. That is, reinforcement does not affect an intrinsic response property (like its topography), but it affects, instead, a relational response property, namely, the similarity between response topography and the model (similarity is the response property upon which reinforcement is contingent). The generalized operant “to imitate” is then shaped. Similarly, in contingencies based on response frequency, each reinforcer follows a particular response sequence (a sequence with particular configuration). The generalized operant to vary is shaped, however, not if sequences with particular configuration become more likely but if sequences with configurations infrequent in the previous trials become more likely (infrequency is the sequence property upon which reinforcement is contingent).

The Balance Hypothesis

Machado and Tonneau (2012) proposed that contingencies based on response frequency promote a dynamic interaction between behavior and environment that can explain the sensitivity of U value to the changes in the parameters of both lag n and threshold contingencies without assuming that variability is itself an operant property of behavior as claimed by Neuringer and Jensen (2012). Machado and Tonneau (2012) hypothesized that negative frequency-dependent selection continually balances the strength of each alternative response or sequence. As an alternative sequence becomes weaker (less frequent), it is more likely to produce reinforcement, and as an alternative sequence becomes stronger (more frequent), reinforcement is withheld. In fact, under both lag n and threshold contingencies, a response sequence produces reinforcement only if it is infrequent. Such contingencies strengthen currently weak responses (by reinforcing them) and weaken currently strong responses (by extinguishing them), according to Machado and Tonneau (2012). Under such contingencies, no particular sequence is consistently selected. Instead, many response sequences are emitted. The resulting behavior can be, therefore, highly unpredictable. The claim that contingencies based on response frequency produce variable responding because they continually balance the strength of each alternative response sequence is what I call the balance hypothesis.

The reasoning underlying the balance hypothesis can be elucidated by the following hypothetical case. Suppose that an experimenter deliberately aims to maintain undifferentiated responding across a number of alternative responses. A probabilistic schedule of reinforcement is not an option because it can accidentally reinforce a particular alternative response. The experimenter should arrange a contingency that both strengthens currently weak responses and weakens currently strong responses so that no particular response is selected. In order to obtain this effect, the experimenter should continuously measure the strength of each alternative response. The strength can be measured by taking the recent frequency in which each alternative response was emitted. Low frequency indicates weakness, and high frequency indicates strength. The experimenter should then arrange a contingency that continually reinforces infrequent responses and extinguishes frequent responses. This is precisely how contingencies based on response frequency work. If this experimenter instituted this contingency and produced variable responding, he could explain the obtained effect by appealing to the balance hypothesis (i.e. he could argue that the contingency continually balances the strength of each alternative response).

According to Machado (1994), a general problem in studying contingencies based on response frequency is to determine “what is strengthened by the reinforcer” delivered in such contingencies (p. 69). Unlike the hypothesis of variable behavior as a generalized operant, the balance hypothesis assumes that reinforcers delivered in such contingencies act on intrinsic sequence properties. (By “intrinsic”, I mean a sequence property that does not depend on its relation with another sequences.) Machado and Tonneau (2012) assume that when the particular sequence A produces reinforcement, the sequence A becomes more likely. Conversely, when the particular sequence A occurs and does not produce reinforcement, the sequence A becomes less likely. Thus, although lag n contingency and threshold contingency define a reinforcement criterion based on relational properties of response sequences (their absolute frequency in the n preceding trials or their relative frequency in the past trials), the differential

reinforcement that these contingencies promote acts upon intrinsic sequence properties, if Machado and Tonneau (2012) hypothesis is correct. According to the balance hypothesis, when a sequence beginning with a left response produces reinforcement under contingencies based on response frequency, sequences beginning with a left response become more likely, for example.²

If reinforcers are contingent on sequence frequency, but they act upon intrinsic sequence properties, no consistent and stable operant selection occurs because the same sequence is alternately reinforced (when its frequency decreases) or extinguished (when its frequency increases). Thus, no stable and durable functional response class is established. Although sequences beginning with a left response can become slightly more likely after the sequence LLRR produces reinforcement, for example, a stable functional class comprising sequences beginning with a left response is not engendered and maintained. This is so because sequences with an initial left response will not be subsequently reinforced if they do not meet the reinforcement criterion based on frequency. Under appropriate schedule parameters, many sequences may be emitted with roughly equal probability. Machado (1993), for example, concluded that when infrequent triplets of responses were differentially reinforced, highly variable responding emerged in some subjects because some lower order response patterns were maintained with similar strengths “due to the balance nature of schedule” (p. 124). Thus, the frequency distribution of different response patterns or response sequences can be maintained relatively flat over time by contingencies based on response frequency. But here, the distributional flatness is the effect of a contingency that continuously balances the strength of each response pattern or sequence. Unlike Neuringer’s (2002) view, the balance hypothesis assumes that contingencies based on response frequency do not engender a unitary functional class (i.e. different emitted responses or response sequences do not compose *one* operant class engendered by these contingencies). Such contingencies, instead, promote a dynamic interaction between the subject and its environment so that no unitary and stable functional class is produced.

The balance hypothesis suggests, therefore, that unpredictable behavior is promoted and maintained when continued and consistent operant selection does not take place. When nothing is consistently selected, responding is unpredictable. Unpredictable responding is a default performance if the balance hypothesis is correct. Under contingencies based on response frequency, repertoires can remain undifferentiated and undifferentiated behavior is unpredictable.

² Machado (1997) explicitly hypothesized that reinforcers delivered under a lag n contingency actually affect the intrinsic sequence property *number of switches between the operanda*. The subject switches when it shifts from one operandum to another operandum while emitting a sequence. The sequences LLRR and LLRL, for example, contain one and two switches respectively. Machado (1997) suggested that when sequence A with k switches is emitted and produces reinforcement under lag n contingency, not only sequence A but other sequences with k switches can become more likely to occur. According to Machado (1997), this induction effect could explain the high number of different sequences that lag n contingency engenders (and, consequently, the high U values obtained under such contingency). Machado (1997) tested his hypothesis by differentially reinforcing sequences with specific number of switches between the operanda. The subjects indeed emitted a high number of different sequences although reinforcers were not contingent on sequence frequencies.

Evidence Supporting the Balance Hypothesis

The balance hypothesis can explain why U value is sensitive to manipulations in the parameters of contingencies. According to Machado and Tonneau (2012), a lag n contingency imposes extinction periods on emitted sequences, and the higher the n value, the longer the extinction period for any particular sequence. In fact, a lag n contingency essentially establishes that sequence A will not produce reinforcement in the n trials subsequent to each trial in which sequence A is emitted. This implies that a lag n contingency imposes a limit on the number of reinforcers that each sequence can produce in a given number of trials and the higher the n value, the lower the number of reinforcers that each sequence can produce (in 50 trials, for example, each sequence can produce at most 3, 5 or 25 reinforcers if the prevailing contingency is lag-20, lag-10 or lag-2 respectively).

Under a threshold contingency, a sequence is reinforced only if its relative frequency in previous trials does not exceed a threshold value. Thus, the lower the threshold value, the lower the number of reinforcers that sequence A can produce before its relative frequency exceeds the threshold value (when its relative frequency is lower than that). Moreover, the lower the threshold value, the higher the number of trials that must occur without the completion of sequence A before its relative frequency equals the threshold value (when its relative frequency is higher than that). Thus, if reinforcers strengthen the particular sequence that produces them, the higher the n value or the lower the threshold value, the less each sequence can be strengthened. Hence, higher U values, obtained under more stringent contingencies, mean smaller differences between the strength levels that each different sequence reaches under contingencies based on response frequency. Therefore, the higher the n value or the lower the threshold value, the flatter is the expected frequency distribution of the different sequences emitted under lag n and threshold contingency.

This reasoning predicts that the probability of reinforcement of each sequence changes inversely with the n value and directly with the threshold value. This prediction is confirmed by data showing that higher n values in lag n contingencies and lower threshold values in threshold contingencies consistently produced higher U values and did not produce a higher percentage of reinforced trials (Page and Neuringer 1985; Grunow and Neuringer 2002) or even produced a lower percentage of reinforced trials (Morris 1989; Abreu-Rodrigues et al. 2005; Doughty et al. 2013). Because the percentage of reinforced trials does not increase as the requirement parameter values increase, the average number of reinforcers per sequence (the total number of delivered reinforcers divided by the number of different sequences emitted during the session) decreases as the number of different emitted sequences increases.³ Page and Neuringer (1985) and Morris (1989) showed that the number of different emitted sequences in fact increased as the n value increased in a lag n contingency. Grunow and Neuringer (2002) also found that the frequency distribution of the different sequences “broadened and flattened” under lower threshold values (p. 253).

³ If an increase in U value occurred, for example, when the subject came to emit 20 different sequences in a session of lag-12 contingency, instead of only 10 different sequences emitted in a previous session of lag-5 contingency, and the subject obtained 40 reinforcers per session under both conditions, the average number of reinforcers per each different sequence decreased from 4 to 2.

The balance hypothesis also can explain why responding obtained under the Yoke condition was more predictable than responding obtained under the Vary condition in Denney and Neuringer's (1998) study. During the Vary component, the contingency differentially reinforced currently infrequent sequences. Because reinforcers were delivered during the Yoke component irrespective of the sequence that had been emitted, no limit was imposed on the number of reinforcers that each different sequence could produce under this component. In the presence of the Yoke antecedent stimuli, therefore, few different sequences, or even a single sequence, could have produced all the available reinforcers whereas the available reinforcers in the Vary component were dispersed across a high number of different sequences. During the Yoke component, therefore, the contingency might have selected few different sequences. If this occurred, the onset of the Yoke antecedent stimuli was more likely to evoke the occurrence of those few sequences than did the onset of the Vary antecedent stimuli, that is, the Yoke component should produce a lower number of different sequences than Vary component. This effect could explain the higher U values obtained under Vary component.

Doughty and Lattal (2001) proposed a hypothesis consistent with the balance hypothesis to explain why variable responding was more resistant to change than was stereotyped responding. The authors required pigeons to emit four-response sequences in two terminal links of a multiple chained schedule. A threshold contingency was in effect in one terminal link, and a repeat contingency was established in another terminal link (in which only the sequence LRLR was reinforced). After performance stabilized under both contingencies, variable responding and stereotyped responding were exposed to two disruptors (prefeeding and reinforcers delivered under variable time schedules). Variable responding was more resistant to change than was stereotyped responding (resistance was measured as decreases in response rate relative to baseline conditions). The authors hypothesized that variable responding was more resistant because under the threshold contingency, all 16 different sequences were reinforced and "each sequence might have reached its own level of strength" (p. 212). This hypothesis assumes that under a threshold contingency, each particular sequence reaches a certain level of strength because each one of them is reinforced. Thus, the balance hypothesis can provide a plausible explanation as to why variable responding was more resistant to change.

Limitations of the Balance Hypothesis

Under a lag n contingency, a subject that cycles through $n+1$ different sequences always meets the reinforcement criterion. This response pattern exhibits maximal sequential dependence. Systematic cycling through all sequences can also produce all available reinforcers under threshold contingency. Contingencies based on response frequency indeed produce systematic alternation of different alternative responses when the schedule is permissive. Machado (1992) showed that subjects systematically alternated between two alternative responses when the contingency selected the currently infrequent individual response alternative. Holth (2012) reported an experiment in which four different responses were recorded. A lag-3 contingency was in effect, and the rats developed a pattern of cycling through the four alternative responses. In a broad sense, it can be said that these contingencies balance the proportion of each alternative

response. Each alternative response is equally often emitted in the pattern LRLRLR, for example. This balance of alternative responses presumably derives from a response chain in which the stimulation provided by an L response consistently controls the emission of a subsequent R response and vice versa. In this case, balance is not an outcome of an unstable responding like in the balance hypothesis. Hence, the balance hypothesis cannot explain the systematic response patterns generated by such lenient contingencies based on response frequency.⁴

If one had to predict the responding of a subject under Machado's (1992) contingency, which reinforced infrequent individual responses, he or she would propose that the subject would generate a pattern like LLLLLRRRRRLLLLLRRRRR, which alternates runs of R responses and runs of L responses with nearly the same number of R and L responses in each run. The balance hypothesis assumes that as a response (L or R) becomes more frequent, it is extinguished (losing its strength) and the alternative one becomes reinforceable. When the alternative weak response is finally emitted, it is reinforced and becomes more frequent and a new cycle begins. Cycles of successive extinctions and reinforcements balance the frequency of each alternative response by maintaining the strength of each one similarly over time. As we saw, the selection of infrequent individual response in Machado's (1992) experiment did not produce such pattern but generated, instead, a pattern of systematic alternation between the responses L and R. The effects of contingencies based on response frequency can be summarized as follows: (1) if a contingency permits but does not require variability of response, repetitive responding is generated (Page and Neuringer 1985); (2) if a contingency is lenient (lag n with low n value or differential reinforcement of infrequent individual alternative responses), systematic alternation between alternative responses is produced (Machado 1992; Holth 2012); and (3) if a contingency is stringent (lag n with high n value, differential reinforcement of infrequent triplets of responses or a threshold contingency with a low threshold value), variable responding are produced (Page and Neuringer 1985; Machado 1993; Grunow and Neuringer 2002).

Generalization tests can also reveal limitations to the balance hypothesis. If contingencies based on response frequency produce the generalized operant to vary, variable responding may be observed in novel contexts (with responses topographically distinct, for example). If variable behavior, however, emerges from unstable responding in which nothing was consistently selected by contingencies, there is nothing that will transfer to a novel situation. Weiss and Neuringer (2012), for example, differentially reinforced variable interactions with objects in one group of rats (reinforced group). During variability training, an experimenter dispensed food pellets contingent on responses directed to the objects. At most, three repetitions of the same response were

⁴ According to this reasoning, lenient contingencies based on response frequency are contingencies in which stimuli generated by previous responses acquire discriminative functions and come to control subsequent responses, the outcome being the systematic alternation between alternative responses. When contingencies are stringent, this control by antecedent stimuli is precluded. When infrequent individual responses were reinforced (Machado 1992), for example, the stable pattern that met the contingency was *RLRLRL*, a pattern that requires discriminative control by stimuli generated by the last emitted response only. When infrequent triplets of individual responses were reinforced (Machado 1993), however, the stable pattern that met the contingency was *RRRLRLLL*, a pattern that requires discriminative control by the last three emitted responses. Neuringer and Jensen (2012) themselves suggested that contingencies based on response frequency generate a random-like performance "when memory requirements exceed the organism's capacity" (p. 527). Memory can be a metaphoric way to refer to control exerted by previous responses over subsequent ones.

reinforced across a 2-min moving window. As the variability training advanced, the number of different topographies of interacting with the objects increased. Rats in another group were exposed to the same objects and consumed the same number of food pellets, but variable responses were not differentially reinforced (exposed group). During the training phase, rats in the reinforced group interacted more variably with the objects than did the rats in the exposed group. In a subsequent transfer-of-training test, food pellets were hidden within or under objects never presented before and the animals of both groups were permitted to explore the environment. Rats in the reinforced group consumed more pellets than did rats in the exposed group. This more efficient foraging behavior indicates that something learned under the variability contingency transferred to the novel context. The balance hypothesis cannot satisfactorily explain such a transfer effect.

Conclusions

Contingencies based on response frequency generate variable responding under some conditions, but it is not clear which behavioral processes underlie such effects. According to Neuringer (2012), variable responding engendered by such contingencies constitutes a generalized operant (a unitary functional response class). According to Machado and Tonneau (2012), however, contingencies based on response frequency continually balance the strength of each alternative response and variable responding emerges as a derivative effect of such contingencies. The difference between these approaches is not merely that each one of them analyzes the same behavioral process at different levels. They propose qualitatively distinct behavioral processes. Precisely, because of that circumstance, different predictions can be made about responding in a generalization test if the balance or operant variability hypothesis is correct, for example.

The balance hypothesis assumes that local processes explain the variable responding produced by frequency-dependent selection. Traditional measures of predictability, however, consider long series of response sequences. The most common U value is calculated from proportions obtained for entire experimental sessions. Molecular measures that allow one to identify changes in the strength of each alternative response can provide better evidence for the balance hypothesis (something like a cumulative record of each different sequence emitted during a session, for example). Such recording can show whether sequences sharing some properties (like beginning with a left key response) become more likely when sequences beginning with a left key response are reinforced, for example.

The literature on operant variability offers few data relating structural sequence properties and the history of reinforcement of particular sequences or group of sequences. Neuringer (1991), for example, required rats to emit four-response sequences under a lag-5 contingency. Each four-response sequences constituted a trial. He calculated the percentage at which the first response of the sequence emitted in trial n repeated the last response of the sequence emitted in trial $n-1$. Neuringer found that subjects repeated the last response of the sequence emitted in trial $n-1$ more often when trial $n-1$ terminated with reinforcement than when trial $n-1$ terminated with time-out. Doughty and Lattal (2001) obtained variable responding under a threshold contingency,

but they found some stereotyped patterns in responding by some subjects. For one subject, most of response sequences that were emitted at or above the proportion 1/16 contained a left key response as a third response in the sequence (there were 16 possible different sequences). For another subject, most of the sequences emitted at or above the proportion 1/16 contained a repetition (two right key responses or two left key responses) as the two final responses in the sequence. Machado (1997) examined the internal structure of eight-response sequences emitted at stable state under a lag n contingency. Most subjects showed preference for one key (L or R) in the first response of the sequence. As the trial advanced, they tended to switch to other key. Cohen et al. (1990) showed the four most frequent sequences emitted under lag-5 contingency. Doughty et al. (2013) showed that under a lenient threshold contingency (threshold value=0.30), the most frequent sequences terminated with three responses on the same key.

These studies, however, did not show the history of reinforcement of each different sequence or groups of sequences (data showing whether sequences terminating with a L response were more frequently reinforced than were sequences terminating with a R response or what proportion of sequences terminating with repetition was reinforced or what proportion of certain sequence as whole was reinforced, for example). Such data can provide evidence allowing one to confirm or to reject the balance hypothesis and can also indicate whether reinforcement acted upon sequences as a whole or upon specific sequence properties.

Holth (2012) suggested that a mix of basic processes could explain variable responding obtained under contingencies based on response frequency. Because response sequences are complex units, reinforcement and extinction can affect individual responses or switches between the manipulanda that compose the sequences. Along with the lack of stimulus control by previous responses, this mix of cyclically reinforced and extinguished sequence components can result in the variable response pattern generated by contingencies based on response frequency. Variability induced by extinction—e.g. increases in variability of responding that is observed when operants are submitted to extinction (Neuringer 2002)—is probably part of the mixed basic processes that can explain the effects of contingencies based on response frequency. As a strong (frequent) response sequence is extinguished, other weak (infrequent) sequences are more likely to occur and be reinforced.

Even if contingencies based on response frequency shape the generalized operant to vary under some conditions, the possibility that such contingencies increase variability of responding as a derivative effect under other conditions should be considered. That is, the balance hypothesis and the operant variability hypothesis may not be mutually exclusive. It is possible that each one of them applies to different cases. They can represent complementary approaches to explain increase in variability of responding. Because the basic processes underlying the effects of contingencies based on response frequency are not clear, it is difficult to account for even more complex effects of such contingencies. For example, human subjects were able to generate binary sequences that could not be statistically distinguished from random sequences (Neuringer 1986). However, without a clearer understanding of the basic processes, it is hard to access the role of verbal instructions in the production of random binary sequences, for example.

Discovering variables controlling behavioral predictability may be useful not only because unpredictable behavior can be adaptive in sports or games; the question has

theoretical relevance since the very conceptual framework of behavior analysis assumes that behavior is, to some extent, a predictable phenomenon. Moreover, the conceptual and empirical investigation of operant variability may also elucidate generalized operants (a question worth studying in its own right). Finally, techniques derived from such findings can be applied to treat human behavior characterized by highly stereotyped (highly predictable) responding, like those observed in some children diagnosed with autism and other developmental disabilities (Neuringer 2002).

Acknowledgments I would like to thank Patrick Diamond for reviewing the style of this paper. I would also like to thank Maria de Lourdes Passos for her helpful comments on the content and style of this article. Finally, I would like to thank the reviewer, Allen Neuringer, for his helpful comments and suggestions on the content of this article.

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. doi:10.1901/jeab.2005.33-03.
- Attnave, F. (1959). *Applications of information theory to psychology: a summary of basic concepts, methods and results*. New York: Holt-Dryden Book: Henry Holt.
- Bryant, D., & Church, R. M. G. (1974). The determinants of random choice. *Animal Learning & Behavior*, 2, 245–248.
- 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.
- Denney, J., & Neuringer, A. (1998). Behavioral variability is controlled by discriminative stimuli. *Animal Learning & Behavior*, 26, 154–162. doi:10.3758/BF03199208.
- 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. doi:10.1901/jeab.2001.76-195.
- Doughty, A. H., Giorno, K. G., & Miller, H. L. (2013). Effects of reinforcer magnitude on reinforced behavioral variability. *Journal of the Experimental Analysis of Behavior*, 100, 355–369.
- Grunow, A., & Neuringer, A. (2002). Learning to vary and varying to learn. *Psychonomic Bulletin & Review*, 9, 250–258.
- Holth, P. (2012). Variability as an operant? *The Behavior Analyst*, 35, 243–248.
- Hunziker, M. H. L., Saldana, 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. doi:10.1901/jeab.1996.65-129.
- 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 responses: some data and a new model. *Behavioural Processes*, 30, 103–130.
- Machado, A. (1994). Polymorphic response patterns under frequency-dependent selection. *Animal Learning & Behavior*, 22, 53–71.
- 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. doi:10.1901/jeab.1997.68-1.
- Machado, A., & Tonneau, F. (2012). Operant variability: procedures and processes. *The Behavior Analyst*, 35, 249–255.
- Morris, C. J. (1989). The effects of lag value on the operant control of response variability under free-operant and discrete-response procedure. *The Psychological Record*, 39, 263–270.
- 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. doi:10.1037/0097-7403.17.1.3.

- Neuringer, A. (2002). Operant variability: evidence, functions, and theory. *Psychonomic Bulletin & Review*, 9, 672–705. doi:[10.3758/BF03196324](https://doi.org/10.3758/BF03196324).
- Neuringer, A. (2012). Reinforcement and induction of operant variability. *The Behavior Analyst*, 35, 229–235.
- Neuringer, A., & Jensen, G. (2012). Operant variability. In G. J. Madden, W. V. Dube, T. D. Hackenberg, G. P. Hanley, & K. A. Lattal (Eds.), *APA handbook of behavior analysis: methods and principles* (Vol. 1, pp. 513–546). Washington, DC: American Psychological Association.
- Odum, A. L., Ward, R. D., Barnes, C. A., & Burke, K. A. (2006). The effects of delayed reinforcement on variability and repetition of response sequences. *Journal of the Experimental Analysis of Behavior*, 86, 159–179. doi:[10.1901/jeab.2006.58-05](https://doi.org/10.1901/jeab.2006.58-05).
- Page, S., & Neuringer, A. (1985). Variability is an operant. *Journal of Experimental Psychology: Animal Behavior Processes*, 11, 429–452. doi:[10.1037/0097-7403.11.3.429](https://doi.org/10.1037/0097-7403.11.3.429).
- Skinner, B. F. (1953). *Science and human behavior*. New York: The Free Press.
- Ward, R. D., Kynaston, A. D., Bailey, E. M., & Odum, A. L. (2008). Discriminative control of variability: effects of successive stimulus reversal. *Behavioural Processes*, 78, 17–24. doi:[10.1016/j.beproc.2007.11.007](https://doi.org/10.1016/j.beproc.2007.11.007).
- Weiss, A., & Neuringer, A. (2012). Reinforced variability enhances object exploration in shy and bold rats. *Physiology & Behavior*, 107, 451–457.