



On Response Strength and the Concept of Response Classes

David C. Palmer¹ 

Accepted: 17 June 2021 / Published online: 29 June 2021
© Association for Behavior Analysis International 2021

Abstract

Simon et al. (2020) argue that the concept of response strength is unnecessary and potentially harmful in that it misdirects behavior analysts away from more fruitful molar analyses. I defend the term as a useful summary of the effects of reinforcement and point particularly to its utility as an interpretive tool in making sense of complex human behavior under multiple control. Physiological data suggest that the concept is not an explanatory fiction, but strength cannot be simply equated with neural conductivity; interaction with competing behaviors must be considered as well. Decisions about appropriate scales of analysis require a clarification of terms. I suggest defining behavior solely in terms of its sensitivity to behavioral principles, irrespective of locus, magnitude, or observability. Furthermore, I suggest that the term *response class* be restricted to units that vary together in probability in part because of overlapping topography. In contrast, *functional classes* are united by common consequences; they vary together with respect to motivational variables but need not share formal properties and need not covary with acquisition and extinction contingencies.

Keywords Definition of behavior · Functional classes · Inhibition · Response classes · Response competition · Response strength

In an earlier article, I attempted to show how the concept of response strength can be used in the interpretation of complex human behavior (Palmer, 2009). The thesis of the article was speculative, namely, that all behavior lies on a continuum of response strength, that response strength is additive, crossing at some point, a “threshold,” at which the behavior is emitted. As it approaches this threshold, I argued, incompatible responses are in competition with one another, with only the strongest response being emitted, the effect possibly filtered by inhibitory processes. I believed then, and I still believe, that the scheme I proposed is probably “correct,” in the sense that future work

✉ David C. Palmer
dcpalmer@smith.edu

¹ Department of Psychology, Smith College, Northampton, MA 01963, USA

will support it at both the physiological and behavioral levels, but at the very least, I find the proposal to be a useful heuristic when attempting to interpret behavior under multiple control. In such cases, the behavior of interest is often a unique event that cannot be understood in terms of data aggregated over time. I concluded my article by saying, “I am uncertain whether the concepts of response strength and response probability are necessary, but I think they do more good than harm” (p. 59).

Simon et al. (2020) recently reviewed the concept of response strength in this journal, with a special focus on my article, and concluded that, to the contrary, the concept does more harm than good. They argued that the notion of response strength is at best superfluous; worse, it encourages a misconception about the appropriate temporal window for viewing behavior. The editors have kindly invited me to respond.

At the outset I want to make clear that their article does a good job of identifying the pitfalls of hypothesizing beyond one’s data, mingling levels of analysis, and introducing terms of uncertain ontological status. Furthermore, they did not distort my article in order to score debating points. It is true that I found myself objecting to this or that point, but no doubt the fault lies as much in my exposition as in their interpretation. Moreover, when I grant them their premises, namely, that behavior is best understood as patterns extended in time considerably beyond the level of molecular “responses,” I agree with them. I don’t think the concept of response strength is usefully applied to extended patterns such as taking the train to work as opposed to driving a car, or going out to eat as opposed to making dinner at home. These are coherent patterns of behavior, but they are not usefully considered “responses.” I have no quarrel with someone who prefers to analyze behavior at a different degree of magnification. As Hineline (2001) observed, order can be found at multiple scales of measurement. (See Shimp, 2020, for a discussion of the numerous interpretations of appropriate levels of analysis.) Nevertheless, I find it useful to speak of response strength, in particular when a particular target response is the object of inquiry, and I see no good reason to avoid doing so, except in the rare cases in which objective measures make interpretive exercises unnecessary.

At the behavioral level, we say that reinforcement “strengthens” behavior in the context in which it occurs. The term is used as a summary term for a family of behavioral effects, such as an increase in rate, an increase in running speed, a decrease in latency, prepotency over competing behavior, an increase in response amplitude, or per Catania (2017), increased resistance to change. Both Skinner (1938, 1957) and Keller and Shoenfeld (1950) explicitly invoked “strength” as a summary term for these various effects and used the term and its cognates several hundred times each in that sense. The term *response probability*, stripped of its precise mathematical sense, is often used as a synonym of strength, with usage depending on context: we do not speak of the probability of a response that has just been emitted, but in consideration of its latency and amplitude, we might speak of its strength. When predicting a response, we are more likely to speak of its probability rather than its strength, but it comes down to mere speculation about where response strength lies with respect to a threshold. According to my understanding of such things, response strength, as commonly used, is a hypothetical construct, at least as the term is used by MacCorquodale and Meehl (1948). Skinner (1957) noted that the “notion of strength” of a verbal operant was “based on” various empirical measures. Herrnstein (1961) suggested that multiple-key concurrent operant procedures might provide “significant empirical support” for the

concept of response strength, better at least than data from a single-key procedure. These locutions imply that strength is “real” but cannot be measured directly; rather, it is merely implied or glimpsed by empirical measures. This implies, in turn, that as science advances, the concept of response strength may someday be fully understood and no longer hypothetical.

The uncertain status of response strength in an objective science has not gone unnoticed. Moore (2008) suggested that the concept of response strength is “perhaps metaphorical” (p. 109). Michael (1993) proposed replacing it with frequency and relative frequency: “This usage makes it possible to avoid such controversial terms as *response strength* and *response probability*” (p. 192; emphasis in original). In this respect, at least, Simon et al. (2020) are firmly on Michael’s side: “Understanding response strength as an intervening variable, synonymous with response probability, would avoid the risk of reification. Then, a response’s rate does not indicate its strength but instead determines it” (p. 691).

But this will not do. How do we determine the probability of a response? What is the denominator? Probability has a clear meaning in mathematical models, where we can specify a sample space and a sampling procedure, but not when the sample space is unknown, or continually changing in unknown ways, as in complex environments. In behavior analysis, we can use the term loosely, as I do, under control of many of the same variables that control appeals to response strength, or we can use it as some sort of derivative of rate, as suggested above, in which case it can be given a precise meaning but with limited application. Rate is an appropriate measure of behavior only in certain contexts. In discrete trial procedures, like the semantic priming procedure, rate is fixed by the intertrial interval and trial length, whereas strength is indexed by latency. Rate is notoriously inappropriate as a measure of the strength of verbal behavior: a comedian whose joke is received with riotous laughter does not tell the joke again until the audience has changed, possibly weeks or months later.

Why Speculate about Response Strength and “Latent” Behavior?

As I have said on many other occasions, we should distinguish the several purposes of science:

Science serves us in two ways. First, it underpins our mastery of the physical and biological world: We should not like to do without our vaccines, antibiotics, semiconductors, or internal combustion engines. But perhaps an even more important service is to resolve mysteries about nature. Science offers beautiful, elegant, and, often, deeply satisfying explanations for complexity and order in nature, and if forced to choose, we might prefer to live in a cave, with our understanding intact, than in a wonderland of gadgets, benighted by superstition. (Palmer, 1998, p. 3)

To master nature requires complete control of the subject matter, but to understand it requires only that we can see at least one way an outcome might have occurred given established scientific principles. We might not be able to create a platypus in the laboratory, but we can understand how such a beast might have evolved without

benefit of a designer, and this dispels the temptation to cling to superstitious alternatives. In short, we do not speculate about unobserved events in our experimental analyses. Such speculations are reserved for interpretive exercises in which we try to make sense of phenomena that lie outside the reach of experimental control by appealing only to established principles and hence tentatively claim it as within the embrace of our science. The first interpretive exercise to which I devoted any considerable effort was how to understand recall solely from a behavior analytic perspective (Palmer, 1991). It is one of the most ordinary of human behaviors, and if we are to claim that our science is comprehensive, we must provide such an interpretation. Consider an everyday question such as “What did you have for breakfast yesterday?” The contingency calls for a particular response, but the current constellation of discriminative stimuli is of little help; the question could be asked in almost any context. We can’t appeal to previous experience with the question: it may occur only once; if not, any reinforcement for having answered “eggs and bacon” in the past will interfere with, rather than facilitate, correctly saying “French toast” today.

The answer cannot be explained by appealing to rates, averages, probabilities or temporally extended patterns of behavior. It is a unique event. The target response is “in our repertoire,” in the sense that there is a set of conditions under which it will be emitted, but it is not currently being emitted. Simply acknowledging these two states of the response establishes, at the very least, a dichotomous scale of response strength. However, the task requires that we go to work on increasing the strength of the target response by the mnemonic strategy of providing ourselves with successive supplementary probes to the point that the response is actually emitted. If we can plausibly assume that our successive self-probes were relevant, that is, that stimulus control was additive, then the scale is better seen as continuous, or at least ordinal. The threshold of responding is usually of great practical importance, but there is no reason to assume that additional probes would not have led to a response of shorter latency and unusual vigor. That is, the response threshold need not represent the upper end of the scale. I find nothing in this account that should bother a behavior analyst, and I believe it provides a wholly behavioral explanation of an everyday phenomenon without resorting to the troublesome metaphor of memory storage. (See Palmer, 1991, for details.) The concept of strength, as used here, is an interpretive tool, not an experimental measure; it is admittedly an imprecise term, but it captures the presumed effect of successive probes on behavior and thereby resolves a behavioral puzzle.

A principal virtue of behavior analysis is that it enables us to make sense of the complex, fleeting, and unique experiences of daily life. In the laboratory, we necessarily mete out stimulus changes sparingly in service of experimental control, but in everyday affairs we are simultaneously bombarded by myriad discriminative stimuli. The reader need merely glance around to acknowledge that, in the visual modality alone, discriminative stimuli are everywhere, each putatively calling for one or more discriminative responses, typically mutually incompatible. When we see no evidence of the corresponding discriminative responses, are we to assume that these stimuli are behaviorally inert?

Our experience prompting behavior in the classroom suggests otherwise. If a child fails to answer a question, merely providing an initial speech sound may be sufficient to evoke the response, possibly with notable emphasis and short latency. We cannot take refuge in the assumption that the two sources of control are a compound stimulus—

they might never have occurred together before—nor can we simply count discriminative stimuli; six prompts are not necessarily more effective than one. The concept of a gradient of response strength and a threshold of emission accommodates the fact that some discriminative stimuli make greater contributions than others, even when a target response has not been emitted. Thus, I infer that the myriad discriminative stimuli in your ken at this moment are all exerting a behavioral effect. Such effects may be negligible, but they reveal themselves when supplementary stimuli, perhaps equally insignificant in themselves, appear.

Verbal behavior is almost always under multiple control, and when a verbal response is the object of study, invoking the concept of a gradient of response strength is particularly tempting. In *The William James Lectures*, Skinner (1948) found it helpful to appeal to below-threshold behavior:

Our basic datum, we may recall, is not a verbal response as such but the probability that a response will be emitted. . . . One can be said to possess a number of different verbal responses in the sense that they are observed from time to time. *But they are not entirely quiescent or inanimate when they are not appearing in one's own behavior.* . . . A latent response with a certain probability of emission is not directly observed. It is a scientific construct. But it can be given a respectable status, and it enormously increases our analytical power. (p. 25; emphasis added)

For example, Skinner's article on the verbal summator was devoted to the effect of indistinct nonverbal stimuli on the evocation of echoic behavior (Skinner, 1936). Because the form of the stimuli bore only prosodic similarity to the responses and varied from person to person, he inferred the contribution of other variables and called their influence "latent."

Putting Empirical Meat on the Bones of the Concept of Response Strength

The different empirical measures of response strength mentioned above are not equivalent to one another, cannot be easily translated into one another, nor are they equally useful. That rate and latency are taken as measures of the same thing implies that there is a thing to be measured that manifests itself in different ways. What might that "thing" be? Skinner (1938) assumed that "the basic concept of synaptic conductivity or its congeners is for most purposes identical with reflex strength" (p. 421). This was an uncontroversial statement even at the time. Sherrington (1906), who is credited with introducing the term "synapse" (Foster & Sherrington, 1897) was aware that synapses differed in their "facility of conduction" (p. 155) and assigned them hypothetical integer values to illustrate how a sensory stimulus might be differentially allocated to various effectors.

Simon et al. (2020) prefer that the line between behavior and neurophysiology be kept distinct, but they embrace functional explanations at every level and suggest that we supplement our understanding of behavior by reference to evolutionary contingencies. I heartily endorse the latter proposal but not the first. I see no reason to avoid

consulting relevant neurophysiological findings, and together with John Donahoe have been attempting to integrate such findings with both behavioral and evolutionary concepts and neural network modeling for over 40 years in what we call a *biobehavioral interpretation* of behavior (e.g., Donahoe & Palmer, 1989; Donahoe & Palmer, 2019; Palmer & Donahoe, 1992). Skinner (1938) and others have argued that a science of behavior can proceed independently of the science of neurophysiology, which is demonstrably true, but that fact should not be taken as a prescription that peering within the nervous system is forbidden. Even Skinner (1974) acknowledged the advantages of an integrated analysis:

The physiologist of the future will tell us all that can be known about what is happening inside the behaving organism. His account will be an important advance over a behavioral analysis, because the latter is necessarily “historical”—that is to say, it is confined to functional relations showing temporal gaps. . . . He will be able to show how an organism is changed when exposed to contingencies of reinforcement and why the changed organism then behaves in a different way, possibly at a much later date. What he discovers cannot invalidate the laws of a science of behavior, but it will make the picture of human action more nearly complete." (p. 215)

The purpose of my earlier article was to show how response strength, as a construct, together with a gradient and a threshold, helps us interpret, at the behavioral level, complex tasks like problem solving and recall, and I made only oblique references to physiology. The concept of response strength is widely used in our field, and my purpose was to exploit it, not to justify it. I was content to leave it as a hypothetical construct, accepting that the term was pejorative to behaviorists’ ears. But the relevance of physiological facts was obvious to Skinner and Sherrington and is far more conspicuous today. The effects of reinforcement on synaptic changes are now known in some detail, and the concept of “strength” at this level of analysis is by no means hypothetical. Neuromodulators associated with reinforcing stimuli can strengthen interneural conductivity by increasing the number of synapses at an axon terminal, by increasing the release of neurotransmitter by the upstream neuron, and by increasing the number of neurotransmitter receptors in the downstream neuron. (See Kandel, 2006, for a summary of his contributions to this topic.) The activity of individual neurons reveals both gradients and thresholds. The depolarization of the membrane at the dendrites increases incrementally until a threshold is reached at which point the neuron “fires” an action potential, which in turn affects the polarization of membranes in downstream neurons. The effect is determined by frequency of input, number of inputs, type of inputs, location of inputs, and synaptic efficacies. The strength of the input can be quantified in terms of potential difference across the membrane at the axon hillock. A near-threshold neuron will fire given only a slight change in stimulation. It would not be fanciful to describe the change in depolarization as a change in “strength” of an action potential at the level of the individual neuron. The term might indeed be superfluous but is not misleading.

It requires only a modest adjustment of our analytical microscope to capture behavioral events. Consider the flexion reflex, as revealed by the work of Sherrington (1906) and others. The reflex is only observed when the stimulation reaches a threshold

intensity. At first glance, there appears to be no sense in which the reflex has “strength” below the threshold stimulus intensity. However, Sherrington found that two subthreshold stimuli presented in sequence, or presented simultaneously but close to one another on the limb, could be sufficient to elicit the reflex. Such observations were among the first examples of the additivity of stimulus control, a concept frequently invoked by Skinner and subsequently empirically supported in operant behavior by Wolf (1963) and in a variety of preparations by Weiss (e.g., Weiss, 1964, 1967, 1977). A subthreshold stimulus might appear to have no effect on its target response, but the phenomenon of additivity of stimulus control invites speculation that a reflexive response is “not entirely quiescent” in the presence of a subthreshold stimulus.

Two decades after Sherrington’s published *The Integrative Action of the Nervous System*, Edgar Adrian, who shared the 1932 Nobel Prize with him, showed that neurons coded stimulus intensity, not by the magnitude of action potentials, but by their rate of firing (e.g., Adrian & Zotterman, 1926). In particular, he showed that the firing rate of sensory neurons increased monotonically with increases in stimulus intensity, before leveling off at some maximum. Taken together, Sherrington’s behavioral data and Adrian’s physiological data show how two subthreshold reflexes can differ quantitatively, even though no relevant movement has occurred in either case.

By surgically severing the spinal cords of his experimental subjects, and by controlling the locus of stimulation, Sherrington isolated “reflex arcs” to single pathways consisting of sensory neurons, interneurons, motor neurons, and those muscle fibers actuated by the motor neurons. At the cellular level, thresholds must be exceeded at each junction before a movement occurs, but each element contributes quantitatively to the whole. Figure 1 is a schematic diagram that integrates the data of Sherrington and Adrian. It illustrates that a reflex can differ quantitatively both above and below its threshold. One might object that a reflex is not a reflex until the threshold is met, but the stimulus and its downstream effects are elements of the reflex, and an analysis of such events extends, and does not interfere with, our understanding of behavior. When we notice that a subsequent weak stimulus elicits the response, we do not assume that the first stimulus was irrelevant. Let us take the provisional position of saying that a subthreshold stimulus along with any downstream effects on interneurons and motor neurons changes the reflex “strength” and that that change is revealed in the effect of other variables. The example illustrates both the concept of a gradient of strength and a threshold of emission.

But what of more complex behavior? Most behavior is far too complex to be understood by reference to a wiring diagram, but we can work backward from motor behavior to the motor cortex for additional insight. An experiment by Georgopoulos et al. (1986) is particularly illuminating. They trained monkeys to touch a light that could appear in any of eight positions, in a circular array, while they recorded the activity of several hundred relevant neurons in the motor cortex. By correlating the direction of hand movement with the activity of each neuron, they were able to derive a model that enabled them to determine, with very small error, the direction of hand movement by looking only at the pattern of neural activity and not at the hand itself. The principal point of the article was to show that coordinated directional movement is the result of the activity of a population of neurons with quite different “preferences.” To put it loosely, a coordinated movement of the hand to a 90° position is mediated, in part, by neurons that fire most rapidly to stimuli at a 45° position and by others that fire

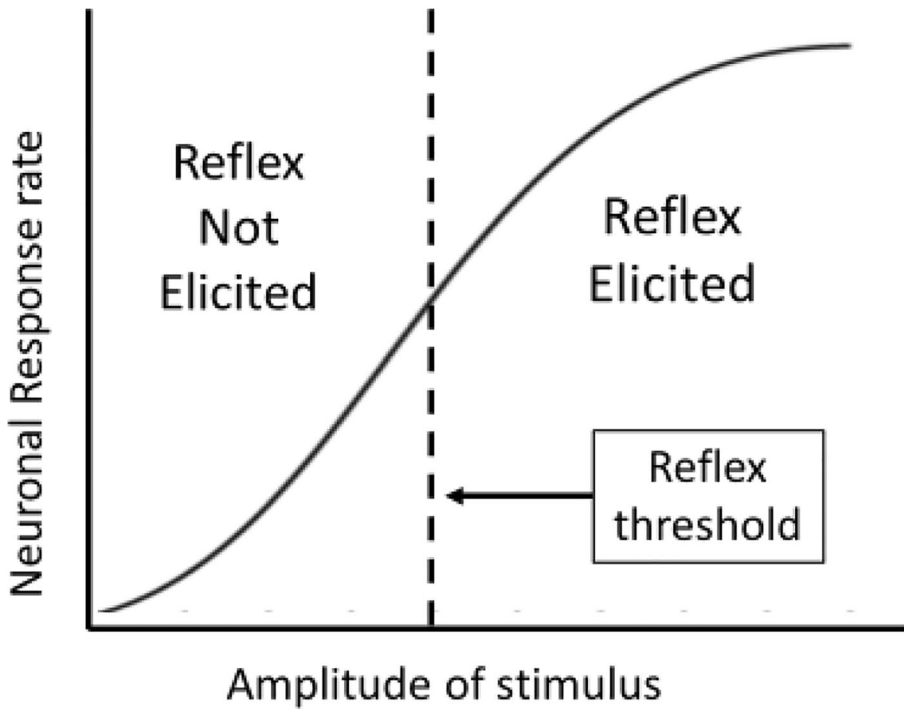


Fig. 1 Schematic Diagram Integrating Findings of Sherrington (1906) and Adrian and Zotterman (1926). *Note.* At the behavioral level, the reflex is all-or-none, but a consideration of the underlying neural activity shows how two reflexes can differ from one another both above and below the threshold. For example, a reflex close to the threshold can become actuated by the addition of a weak second stimulus, while another might require the addition of a relatively strong supplementary stimulus. The neural activity in this figure can be taken to represent that of the population of sensory neurons mediating the reflex, with interneurons and motor neurons following analogous trajectories at unknown points as stimulus intensity increases

most rapidly to stimuli at a 135° position, as well as many others with various orientation preferences. This finding by itself is of considerable interest to behavior analysts, but it is a secondary result that is particularly relevant for the present discussion. In one condition, they trained the monkeys to press the light, not when it came on, but a few seconds later when it dimmed. They found that, during the waiting period, the population of motor neurons began to fire sufficiently to predict directional movement (Georgopoulos et al., 1989). That is, at the level of the activity of motor neurons, the response was relatively strong, but at the level of actual movement, it was latent.

So can we agree with Skinner and define response strength in terms of neural networks and synaptic efficacy, in effect as an intervening variable? I think not. It is not my purpose to reduce behavioral concepts to physiological ones: relevant facts from physiology are sufficient to suggest that the concept of response strength as a behavioral term is more than an explanatory fiction, but we almost never have such facts available to us when we invoke the term. Moreover, our understanding of the physiology of behavior is in its infancy. The emission of behavior requires a consideration, not just of the target response, but all competing behavior as well, and we know little about such behavioral interactions at either the behavioral or physiological levels.

Inhibition and the “Bottleneck”

Adrian’s work illuminated the role of stimulus intensity with respect to reflex thresholds, but the effects of intensity are modulated by other variables. The intensity of a discriminative stimulus is often relevant, but a shouted question to a student is usually no more likely to evoke a correct answer than a soft one, and organisms can respond to stimulus diminution or offset as well as to stimulus onset. Furthermore, all behavior must be understood in the context of other behavior that may facilitate or interfere with the target behavior. Sherrington found that stimulus control can be not only additive but subtractive. In a state of rest, an animal’s limb is held in a stable position by the tonus, or slight activation, of opponent extensor and flexor muscles. He observed by palpation that when a flexion reflex is elicited, the extensor muscles simultaneously go limp, and vice versa, a phenomenon he called *reciprocal inhibition*. He inferred the role of inhibitory interneurons and located them in the spinal cord, inferences now known to be correct in both cases. The adaptive significance of the phenomenon is obvious: If natural selection has decreed that flexion would be useful in response to an insult to the leg, any activity of opposing muscles would be detrimental. By simultaneously stimulating opponent reflexes and interacting reflexes at various intensities, Sherrington was able to demonstrate abrupt moment-to-moment shifts in the corresponding reflexive responses. That is, the effect of a strong stimulus must be understood only in context of competing reflexes.

In earlier articles (Palmer, 2009, 2013) I suggested that an analogous inhibitory process might occur with all behavior. (See also Catania, 1969; Catania & Gill, 1964.) I offered the heuristic of a bottleneck at the response threshold to account for the fact that behavior is typically smooth and coherent with no sign of competing responses. I hope that I am not alone among behavior analysts in thinking that this is a puzzle that needs to be solved. We occasionally see response blends, but only rarely, given the number of opportunities, and even blends tend to alternate sequentially. If we are walking toward a tree, we can walk to the left of it or to the right of it. Let us stipulate that in our past we have done both with equal frequency and to equal effect. We will either walk to the left or to the right, but we will neither walk into the tree nor will we oscillate back and forth awkwardly. (We do so with pedestrians only because they move left just as we are moving right.) That is, although both responses are putatively equally strong, only one is emitted and that one is emitted smoothly, with no hint of interference from an incompatible response. In a simple world in which stimuli came at us one at a time, this would occasion no surprise, but in the world as it is, we are assaulted by a blizzard of stimuli. Some sort of filtering is necessary, and I showed how a hypothesis of blind mutual inhibition within each response system, analogous to reciprocal or lateral inhibition, could accomplish such filtering. Let us suppose that each of five incompatible movements of an effector is perfectly equally potentiated. By hypothesis, each will try to shut the others down, but the balance is perfect, and no preference emerges. But let a butterfly flit by, so to speak, so that one response becomes stronger by the smallest possible increment. The lucky response would emit a slightly stronger inhibitory signal than the others, which would cause the returning inhibition to be slightly weaker at the next time step (that is, at the next action potential), and the following time step would see a somewhat greater discrepancy. The system would rapidly go out of balance, at an

accelerating rate, and one response would be emitted without interference from competing responses.

This scheme is consistent with what is known about motor coordination, in which muscle groups implementing a response are potentiated, while neighboring muscles are inhibited, a phenomenon called *surround inhibition* (e.g., Beck & Hallett, 2011; Sohn & Hallett, 2004). Furthermore, seizures are understood, in part, as a temporary failure of inhibitory processes, leading to a behavioral free-for-all—just the kind of chaos we might expect, in the absence of filtering, from simultaneous exposure to myriad discriminative stimuli. As a result, at the level of motor behavior, where one movement is, or is not, compatible with another, the proposal is plausible and has empirical support. Whether it can be extended to account for the orderliness of all behavior is speculation, but I cautiously recommend the hypothesis to any others, if any there be, who are perplexed by the smoothness of behavior in a chaotic world.

What is Behavior?

Simon et al. (2020) make a vigorous case for taking a broad and temporally extended view of behavior as the activity of the whole organism, and they object to an analysis that suggests that behavior can be broken up into discrete bits. They point out that a still photograph of an organism can seldom tell us what the animal is doing. As I mentioned above, I have no doubt that order can be found at different levels of magnification, and I cheer them on in their enterprise. As a pragmatic matter, our level of analysis will depend upon the tools at our disposal. A glance at the weather report will tell us whether the beach will be crowded tomorrow, whereas an inventory of relevant moment-to-moment contingencies is pragmatically out of the question. But I would remind them that the broader one's unit of analysis, the more within-unit variability one will find, with corresponding ambiguity in what to measure.

For anyone who tries to interpret complex human behavior, such as recall, problem solving, visualizing, and subvocal speech, the question of how to define behavior is vexing. A variety of definitions have been suggested (see Todorov, 2017, for a sample), but it is not surprising that I prefer my own: *behavior is any change in an organism that is sensitive to independent variables in accordance with some or all of those behavioral principles that have emerged from an experimental analysis*. Is “imagining a potato,” or a burst of neural activity in the hippocampus, or covertly moving a chess piece behavior? The question reduces to whether that activity, in an orderly way, can come under stimulus control, or can extinguish, or can habituate, or can be reinforced, etc. In other words, it is how the phenomenon behaves, so to speak, that determines whether we call it behavior. Unlike most alternatives, this is not a prescriptive definition but an empirical one, and like other empirical definitions (What is a coydog? A radish? Verbal behavior?) it is likely to yield fuzzy answers continually open to refinement. The further we stray from controlled observations, the more we depend on plausible inferences. In practice, we assume the behavioral status of any variable that is similar or analogous to one that has been well studied: moving a switch on a novel bit of apparatus can be assumed to sufficiently like key-pressing to require no special evaluation. Likewise, to the extent that covert speech is merely low-amplitude speech beyond the reach of our current instruments, we can provisionally consider it behavior;

no special ontological status is assumed. Whether we can call perception, neural activity, or subjective experience behavior awaits empirical and technological advances, but tentative answers can be inferred from the extent to which doing so helps us bring otherwise mysterious phenomena under the umbrella of science. For example, Mechner (2010) has offered a behavioral interpretation of exhibitions of “blindfold chess,” that is, those in which grandmasters play multiple games of chess simultaneously without sight of the boards. Analogous interpretive exercises allow us to understand the phenomena of evolutionary biology, astronomy, and historical geology. Science routinely tentatively extends its reach beyond phenomena that can be controlled experimentally.

Behavior is necessarily extended in time, as are relevant contingencies. A contingency analysis of innate behavior would be scaled in millennia, but I find the contingency of reinforcement to be a particularly useful level of analysis for interpreting human behavior, with the operant as the appropriate unit. The terms *operant* and *response class* are commonly used interchangeably, but there is a subtle distinction. An operant is a unit of analysis that embraces the antecedents, response, and consequences of those contingencies in an organism’s environment that are stable enough over time to select reliable units. The verbal response *bank*, for example, can be a member of several operants depending on whether it applies to the side of a river, a financial institution, or a type of billiard shot, among other things. A response class is the behavioral element of a particular operant, and once acquired, it can easily come under control of other variables to form new operants. For example, when learning an unfamiliar name, such as Siobhan or Czeslaw, it may take many trials before the response class emerges as a fluent, coherent unit, but once acquired, it will transfer immediately to another person with the same name. That is, the same response class can be a member of different operants. It is the coherence of response classes that is of primary interest in what follows.

I find that the standard view of the response classes requires refinement. For Skinner (1935), the response class was an empirically defined unit of analysis, and he offered a procedure for deciding one. In particular, he proposed systematically narrowing and expanding one’s defining criteria of stimulus and response elements until a point of maximum orderliness was achieved. His proposal is analogous to adjusting the barrel of a microscope back and forth until the image snaps into focus. When he applied this approach to the behavior of a rat lever-pressing in an operant chamber, he discovered that the most useful definitions of stimulus and response units were not crystalline, but fuzzy. One lever-press was likely to differ from another in force, duration, location, and topography, and the corresponding stimulus conditions presumably varied somewhat from one response to the next according to the posture of the animal. In short, the reinforcement of a response of a particular topography led to the emergence of responses of somewhat different topography under somewhat different stimulus conditions. He found that defining his response according to the criterion of “behavior that operates the lever” gave him data that was orderly, in the sense that it gave reliable patterns of data in response to changes in independent variables. (Of course, these empirical definitions are not immutable; stimulus and response classes can be carved up by differential contingencies.)

According to Catania’s (1973) analysis, “behavior that operates the lever” is an example of a descriptive response class, that is, those responses scheduled for

reinforcement by the experimenter. In contrast, the functional response class is the set of responses that vary systematically with exposure to such a contingency. They need not be the same. For example, the minimum force required to close the switch of a manipulandum might sufficiently determine the descriptive response class, but the reinforcement of responses, so defined, typically generates a distribution that includes a considerable number of subcriterion responses that ordinarily go unmeasured (e.g., Notterman & Mintz, 1965; Pinkston & Libman, 2017). However, the fact that descriptive response classes are commonly sufficient to yield useful data, without the trouble of surveying functional classes, has led to the neglect of the latter in prevailing definitions of response classes. A standard textbook definition of a response class is “a group of responses of varying topography, all of which produce the same effect on the environment” (Cooper et al., 2007, p. 703). Although the phrase “of varying topography” is appropriate so far as it goes, it leads to the inference that topography is irrelevant to the definition of a response class, an inference that in my experience is widespread in our field. That is, it suggests that doing a sum by hand, using a calculator to compute it, and asking one’s friend to do the sum are all members of the same class. This inference has undesirable implications.

The concepts of response classes and stimulus classes are important because they capture empirical data in a way that reinforcement of an “essential response” to an “essential stimulus” does not. If reinforcement strengthens all members of a class of variable exemplars of the prior response, *then we have a useful predictive principle*, viz.: In the future, we expect to see responses similar to, but not necessarily identical to, the reinforced response, in contexts similar to, but not necessarily identical to, the context at the moment of reinforcement. If perfect identity were required, we wouldn’t get orderly data, because perfect identity would almost never occur.

Skinner (1953) understood stimulus and response classes in terms of shared elements: “The control acquired by a stimulus is shared by other stimuli with common properties or, to put it another way, that control is shared by all the properties of the stimulus taken separately” (p. 134). Of course, this position is not unique to Skinner (e.g., Blough, 1975; Estes, 1950; Guthrie, 1935), but he took an equally molecular view of response classes:

A more useful way of putting it is to say that the [response] *elements* are strengthened wherever they occur. This leads us to identify the element rather than the response as the unit of behavior. It is a sort of behavioral atom, which may never appear by itself upon any single occasion but is the essential ingredient or component of all observed instances. The reinforcement of a response increases the probability of all responses containing the same elements. (Skinner, 1953, p. 94; emphasis in original)

This elemental view of response classes is implicit in the phenomenon of response induction, and it gives us both a predictive principle *and an explanatory principle*, in the sense of a mechanism by which such classes hang together but vary somewhat from one instance to the next: novel elements would be carried along by previously reinforced elements and would be reinforced in turn. Over the course of repeated exposure to a contingency, all relevant elements would be selected, and the stimulus and response classes would become distinct (cf. Iversen, 2002). From this perspective,

a response class is a set of response elements that tend to “hang together” in contingencies of reinforcement. They could hang together structurally, as in muscles of the forearm, or they could hang together temporally because they act in coordination, as in the muscles of the diaphragm, larynx, tongue, and throat when we speak. This perspective restores a topographical criterion to the concept of response classes whose members covary in strength, and it accommodates Catania’s (1973) observation that reinforcement creates coherent classes some of whose members may fail to satisfy the relevant contingency.

If we are going to use the concept of a topographical response class as justification for the appearance of novelty—that is for the appearance of any one of a set of somewhat different responses—on the second occasion, it follows that reinforcement must affect all members simultaneously on the first occasion and all subsequent occasions. Likewise, if our concept of response classes were to include all responses that are functionally equivalent—that is, that lead to a common consequence irrespective of form—then all such topographies would have to be reinforced simultaneously. The former implication is plausible: the reinforcement of one response will affect other responses in the class to varying degrees because these other responses will overlap topographically with the reinforced response. But the latter implication is untenable: if the occasional “rogue” topography that presses the lever in a Skinner box—e.g., tail-lashing while nibbling a screw head on the side wall—were reinforced along with every lever press, the two topographies would soon reach asymptote and become equally frequent. A second problem with identifying response classes solely by common consequences is that it suggests that responses on extinction, or on variable schedules, are members of a different class from those responses that are reinforced. In such cases, a punctilious adherence to a functional definition would violate the generality that the concept was invoked to explain.

What, then, is a “response element?” If two responses can objectively be said to differ, the difference between them reveals one or more response elements. If they cannot be said to differ, then there is no variability that needs to be explained. Thus the grain of variability from one response to another reflects the elemental nature of response classes. At the level of motor behavior, we can speculate that the action of a motor neuron and its effector is an element, as suggested by the findings of Georgopoulos et al. (1986, 1989) cited earlier. The question of how covert behavior can have an elemental nature, or how one instance might vary from another, is not easily answered, but on the assumption that covert behavior is a physical phenomenon, the extension to that domain is not unjustified in principle. For Georgopoulos’s monkeys, neuronal activity that could plausibly be called covert reaching had the same, or similar, elemental nature as overt reaching. In the end, it is an empirical question.

Skinner’s clearest statements about the role of topography and consequences in the concept of the operant appear in *Contingencies of Reinforcement* (1969) in endnotes that he appended to the reprinting of his article, “Operant Behavior” (Skinner, 1963):

It is not enough to say that an operant is defined by its consequences . . . *the topography and the consequences define an operant*. . . . “Proximity seeking” is not an operant, or any useful subdivision of behavior, *unless all instances vary together* under the control of common variables, and this is quite unlikely. (Skinner, 1969, pp. 127–128; emphasis added)

That members of a response class must vary together as a function of common variables provides us with a useful criterion for distinguishing responses that serve the same function from those that share both topography and function. A simple test would be to determine if the extinction of one exemplar extinguishes others. Is doing a sum on a calculator in the same class as doing it by hand? Not by Skinner's criterion. If we discover that our batteries are dead, using the calculator will quickly extinguish, and we will turn to doing it by hand, or some other way. As an empirical and interpretive matter, it is useful to have a name for classes of events that vary together under control of common variables, i.e., the response class.

However, a definition of response classes in terms of both topography and function ignores an important generalization: responses with common consequences but different topographies can still covary according to function (cf. Staddon, 1967). A motivational variable will have broad effects on all behavior relevant to some consequence, or classes of consequences (Michael, 1982, 1993). In applied settings, this covariation is often revealed in response hierarchies. For example, the extinction of one response maintained by escape is likely to be followed, in order, by the emergence of second and third responses, etc., that serve the same function and will be extinguished in their turn (e.g., Lalli et al., 1995). Thus, I propose that we distinguish between *response classes*, and *functional classes*. The former rise and fall together in reinforcement and extinction contingencies because members have common elements. The latter vary together according to motivational variables because they share common consequences, but they do not covary with reinforcement and extinction, except to the extent that they share common elements.

Conclusion

I have long been dismayed by the ease with which psychologists invoke explanatory fictions in the interpretation of behavior, so I find myself sympathetic with Simon et al. (2020) in their insistence for conceptual clarity. I agree to this extent: the concept of response strength as a hypothetical construct plays no role in the experimental analysis of behavior, at least at our present level of technology. However, I do not think it can be conceived of as an intervening variable. It would indeed be both convenient and desirable if we could equate response strength with objective measures such as rate or relative frequency per opportunity. It would then be analogous to the concept of gravity, at least to my 19th-century understanding of the concept. Gravity is revealed in a number of seemingly incommensurable ways, such as the path of orbiting planets, the motion of pendula, the thud of a feather in a vacuum, one's pressure on a bathroom scale, and noting the superficially surprising fact that there are two tides a day, not one. But Newton was able to reduce the concept of gravity to an elegant formula relating mass and distance, and the concept, so defined, neatly ties together these disparate phenomena and innumerable others, allowing us to make sense of fragmentary data. It also allows us to make assumptions, untestable in practice, about everyday phenomena: we assume that two marbles on a desktop are attracted to one another, although, given the limits of our measurement tools, such attraction is "covert." By adding mass to the marbles, we could incrementally increase the "strength" of the mutual attraction,

eventually reaching a “threshold” at which the gravitational attraction was just sufficient to overcome friction, hence “emitted.” Unfortunately, at present our mastery of the relevant behavioral phenomena does not suggest an equivalent equation for response strength at either the behavioral or physiological levels. Hull (1943) made a valiant attempt to treat response strength as an intervening variable, i.e., as a function of a variety of other variables, but the mathematical operators were fanciful and the variables themselves troublesome. Our sciences of behavior and physiology are still “pre-Newtonian.”

However, I have found the concept of response strength useful as an interpretive tool when experimental analysis is either impossible or insufficient. It is the nature of interpretive exercises that one can't be sure they are right. Perhaps someone can offer alternative interpretations of recall and of other complex behavior that make no reference to response strength. If those interpretations are more parsimonious, so much the better. I admit to profound skepticism that we can offer a comprehensive explanation of human behavior without reference to private events and moment-to-moment contingencies (e.g., Palmer, 2011), but I see no reason to doubt the utility of considering relatively molar contingencies as well, and I have no interest in legislating how a science of behavior should advance. However, I am inclined to think that science will move in the other direction, that the behavioral and physiological facts of complex behavior will eventually become so thoroughly understood that the concept of response strength will be, not a hypothetical construct, but an empirical measure.

Acknowledgments I thank Jason Bourret, Jon Pinkston, Matt Normand, and members of the Bourret lab at Western New England University for helpful discussions of some of the topics covered here.

Declarations

Conflict of Interest The author has no conflicts of interest and has complied with all relevant ethical standards.

References

- Adrian, E. D., & Zotterman, Y. (1926). The impulses produced by sensory nerve endings. Part 2. The response of a single end-organ. *Journal of Physiology*, *61*(2), 151–171. <https://doi.org/10.1113/jphysiol.1926.sp002281>.
- Beck, S., & Hallett, M. (2011). Surround inhibition in the motor system. *Experimental Brain Research*, *210*, 165–172. <https://doi.org/10.1007/s00221-011-2610-6>.
- Blough, D. S. (1975). Steady state data and a quantitative model of operant generalization and discrimination. *Journal of Experimental Psychology: Animal Behavior Processes*, *1*(1), 3–21. <https://doi.org/10.1037/0097-7403.1.1.3>.
- Catania, A. C. (1969). Concurrent performances: Inhibition of one response by reinforcement of another. *Journal of the Experimental Analysis of Behavior*, *12*(5), 731–744. <https://doi.org/10.1901/jeab.1969.12-731>.
- Catania, A. C. (1973). The concept of the operant in the analysis of behavior. *Behaviorism*, *1*(2), 103–116 <https://www.jstor.org/stable/27758804>.
- Catania, A. C. (2017). *The ABCs of behavior analysis*. Sloan.
- Catania, A. C., & Gill, C. A. (1964). Inhibition and behavioral contrast. *Psychonomic Science*, *1*, 257–258. <https://doi.org/10.3758/BF03342897>.

- Cooper, J. O., Heron, T. E., & Heward, W. L. (2007). *Applied behavior analysis* (2nd ed.). Pearson Merrill Prentice Hall.
- Donahoe, J. W., & Palmer, D. C. (1989). The interpretation of complex human behavior: Some reactions to *Parallel distributed processing*, edited by J. L. McClelland, D. E. Rumelhart, and the PDP Research Group. *Journal of the Experimental Analysis of Behavior*, 51(3), 399–416. <https://doi.org/10.1901/jeab.1989.51-399>.
- Donahoe, J. W., & Palmer, D. C. (2019). *Learning and complex behavior*. Ledgeotop Publishing (Original work published 1994).
- Estes, W. K. (1950). Toward a statistical theory of Learning. *Psychology Review*, 57(2), 94–107. <https://doi.org/10.1037/h0058559>.
- Foster, M., & Sherrington, C. S. (1897). *Textbook of Physiology*. Macmillan.
- Georgopoulos, A. P., Crutcher, M. D., & Schwartz, A. B. (1989). Cognitive spatial-motor processes. 3. Motor cortical prediction of movement direction during an instructed delay period. *Experimental Brain Research*, 75(1), 183–194. <https://doi.org/10.1007/BF00248541>.
- Georgopoulos, A. P., Schwartz, A. B., & Kettner, R. E. (1986). Neuronal population coding of movement direction. *Science*, 233, 1416–1419. <https://doi.org/10.1126/science.3749885>.
- Guthrie, E. R. (1935). *The psychology of learning*. Harper.
- Herrnstein, R. J. (1961). Relative and absolute strength of response as a function of frequency of reinforcement. *Journal of the Experimental Analysis of Behavior*, 4(3), 267–272. <https://doi.org/10.1901/jeab.1961.4-267>.
- Hineline, P. N. (2001). Beyond the molar-molecular distinction: We need multiscaled analyses. *Journal of Experimental Analysis of Behavior*, 75(3), 342–347. <https://doi.org/10.1901/jeab.2001.75-342>.
- Hull, C. L. (1943). *Principles of behavior: Introduction to behavior theory*. Appleton-Century-Crofts.
- Iversen, I. H. (2002). Response-initiated imaging of operant behavior using a digital camera. *Journal of Experimental Analysis of Behavior*, 77(3), 283–300. <https://doi.org/10.1901/jeab.2002.77-283>.
- Kandel, E. (2006). *In search of memory: The emergence of a new science of mind*. W. W. Norton.
- Keller, F. S., & Shoenfeld, W. N. (1950). *Principles of psychology*. Appleton-Century-Crofts.
- Lalli, J. S., Mace, F. C., Wohn, T., & Livezey, K. (1995). Identification and modification of a response-class hierarchy. *Journal of Applied Behavior Analysis*, 28(4), 551–559. <https://doi.org/10.1901/jaba.1995.28-551>.
- MacCorquodale, K., & Meehl, P. E. (1948). On a distinction between hypothetical constructs and intervening variables. *Psychological Review*, 55(2), 95–107. <https://doi.org/10.1037/h0056029>.
- Mechner, F. (2010). Chess as a behavioral model for cognitive skill research: Review of *Blindfold chess* by Eliot Hearst & John Knott. *Journal of the Experimental Analysis of Behavior*, 94(3), 373–386. <https://doi.org/10.1901/jeab.2010.94-373>.
- Michael, J. (1982). Distinguishing between discriminative and motivational functions of stimuli. *Journal of the Experimental Analysis of Behavior*, 37(1), 149–155. <https://doi.org/10.1901/jeab.1982.37-149>.
- Michael, J. (1993). Establishing operations. *The Behavior Analyst*, 16(2), 191–206. <https://doi.org/10.1007/BF03392623>.
- Moore, J. (2008). *Conceptual foundations of radical behaviorism*. Sloan.
- Notterman, J. M., & Mintz, D. E. (1965). *Dynamics of response*. John Wiley & Sons.
- Palmer, D. C. (1991). A behavioral interpretation of memory. In L. J. Hayes & P. N. Chase (Eds.), *Dialogues on verbal behavior* (pp. 261–279). Context Press.
- Palmer, D. C. (1998). The speaker as listener: The interpretation of structural regularities in verbal behavior. *Analysis of Verbal Behavior*, 15, 3–16. <https://doi.org/10.1007/BF03392920>.
- Palmer, D. C. (2009). Response strength and the concept of the repertoire. *European Journal of Behavior Analysis*, 10(1), 49–60. <https://doi.org/10.1080/15021149.2009.11434308>.
- Palmer, D. C. (2011). Consideration of private events is required in a comprehensive science of behavior. *The Behavior Analyst*, 34(2), 201–207. <https://doi.org/10.1007/BF03392250>.
- Palmer, D. C. (2013). Do we need a concept of behavioral inhibition? *European Journal of Behavior Analysis*, 14, 97–103. <https://doi.org/10.1080/15021149.2013.11434448>.
- Palmer, D. C., & Donahoe, J. W. (1992). Essentialism and selectionism in cognitive science and behavior analysis. *American Psychologist*, 47, 1344–1358. <https://doi.org/10.1037/0003-066X.47.11.1344>.
- Pinkston, J. W., & Libman, B. M. (2017). Aversive functions of response effort: Fact or artifact? *Journal of the Experimental Analysis of Behavior*, 108(1), 73–96. <https://doi.org/10.1002/jeab.264>.
- Sherrington, C. S. (1906). *The integrative action of the nervous system*. Yale University Press.
- Shimp, C. P. (2020). Molecular (moment-to-moment) and molar (aggregate) analyses of behavior. *Journal of the Experimental Analysis of Behavior*, 114(3), 394–429. <https://doi.org/10.1002/jeab.626>.

- Simon, C., Bernardy, J. L., & Cowie, S. (2020). On the “strength” of behavior. *Perspectives on Behavior Science*, 43(4), 677–696. <https://doi.org/10.1007/s40614-020-00269-5>.
- Skinner, B. F. (1935). The generic nature of the concepts of stimulus and response. *Journal of General Psychology*, 12, 40–65. <https://doi.org/10.1080/00221309.1935.9920087>.
- Skinner, B. F. (1936). The verbal summator and a method for the study of latent speech. *Journal of Psychology*, 2, 71–107. <https://doi.org/10.1080/00223980.1936.9917445>.
- Skinner, B. F. (1938). *The behavior of organisms*. Appleton-Century-Crofts.
- Skinner, B. F. (1948). *Verbal behavior: The William James lectures*. Unpublished manuscript available at www.bfskinner.org. Accessed 11 Jan 2021.
- Skinner, B. F. (1953). *Science and human behavior*. Macmillan.
- Skinner, B. F. (1957). *Verbal behavior*. Appleton-Century-Crofts.
- Skinner, B. F. (1963). Operant behavior. *American Psychologist*, 18(8), 503–515. <https://doi.org/10.1037/h0045185>.
- Skinner, B. F. (1969). *Contingencies of reinforcement*. Prentice Hall.
- Skinner, B. F. (1974). *About behaviorism*. Alfred A. Knopf.
- Sohn, Y. H., & Hallett, M. (2004). Surround inhibition in human motor system. *Experimental Brain Research*, 158, 397–404. <https://doi.org/10.1007/s00221-004-1909-y>.
- Staddon, J. E. R. (1967). Asymptotic behavior: The concept of the operant. *Psychological Review*, 74(5), 377–391. <https://doi.org/10.1037/h0024877>.
- Todorov, J. C. (2017). Sobre uma definição de comportamento. *Perspectivas em Análise do Comportamento*, 3(1), 32–37. <https://doi.org/10.18761/perspectivas.v3i1.79>.
- Weiss, S. J. (1964). Summation of response strengths instrumentally conditioned to stimuli in different sensory modalities. *Journal of Experimental Psychology*, 68(2), 151–155. <https://doi.org/10.1037/h0049180>.
- Weiss, S. J. (1967). Free operant compounding of variable-interval and low-rate discriminative stimuli. *Journal of the Experimental Analysis of Behavior*, 10(6), 535–540. <https://doi.org/10.1901/jeab.1967.10-535>.
- Weiss, S. J. (1977). Free-operant compounding of low-rate stimuli. *Bulletin of the Psychonomic Society*, 10, 115–117. <https://doi.org/10.3758/BF03329297>.
- Wolf, M. M. (1963). Some effects of combined *S*^{ps}. *Journal of the Experimental Analysis of Behavior*, 6(3), 343–347. <https://doi.org/10.1901/jeab.1963.6-343>.