Tweedledum and Tweedledee: Symmetry in Behavior Analysis

M. Jackson Marr
Georgia Tech

Abstract

Symmetry is revealed when some non-trivial transformation of a system leaves the system unchanged or invariant. Symmetry is a pervasive feature in many sciences from physics to embryology. In physics, for example, symmetry is reflected in fundamental laws. Can this be said of behavioral findings and principles? I explore this question by discussing several examples from behavior analysis including the operational and functional aspects of reinforcement, stimulus and schedule control, the three-term contingency, and putative scale-invariance of behavioral principles—a feature conferring unity to the field of behavior analysis.

Key words: symmetry, three-term contingency, reinforcement, punishment, stimulus control

Resumen

La simetría se muestra cuando alguna trasformación no trivial de un sistema deja este sin cambios o invariante. La simetría es una característica generalizada en muchas ciencias desde la física a la embriología. En física por ejemplo la simetría se refleja en leyes fundamentales. ¿Se puede decir lo mismo de los resultados y principios conductuales? Exploro esta cuestión al discutir varios ejemplos desde el análisis de conducta incluyendo los aspectos operacionales y funcionales de reforzamiento, estímulo, y programas de reforzamiento, la contingencia de tres términos, y la supuesta escala de invarianza de los principios de la conducta—una característica que confiere unidad al campo del análisis de la conducta.

Palabras clave: simetría, contingencia de tres términos, reforzamiento, castigo, control de estímulos.

Most, if not all behavior analysts would identify their field as a natural science, taking a rightful place with the physical and biological sciences. I have addressed this issue and some of its implications in a previous paper (Marr, 2009) with the aim of identifying communalities as well as differences between behavior analysis and other unquestioned natural sciences. I primarily emphasized ontological, empirical and explanatory schemes in our behavioral science and, more particularly, behavior analysis as a branch of the biological sciences. I listed a number of other issues relating behavior analysis to concerns in other natural sciences, one of which will be my focus in this paper—the role of symmetry. In a brief commentary in The Behavior Analyst some years ago (Marr, 2006a), I introduced the concept of symmetry as it might apply to some theories and findings in behavior analysis. In this essay, I will expand on some aspects of my earlier treatment.

Fundamentally, a natural science reflects special discernment in observations, selections, and distinctions with respect to the phenomena of interest as well as with theories of those phenomena. With respect to phenomena, some distinctions are based on fairly direct observations while others may emerge from more detailed study, including indirect observations and theoretical formulations. Some phenomena

---

1 Address: School of Psychology, Georgia Tech, Atlanta, GA 30332-0170. email: mm27@prism.gatech.edu
may initially appear quite distinctive, but may turn out not to be. For example, rain and snow certainly look very different, but they are simply different phases of the same substance, H\textsubscript{2}O. The motion of the moon and an apple falling to earth are both actions of a single mechanism—gravity—the moon is also falling toward the earth. And, as Faraday and Maxwell taught us, electricity and magnetism, seemingly quite different phenomena, turn out to be dynamically interdependent.

When a science has not yet reached a proper threshold of understanding its phenomena of interest, distinctions may be vague and provisional, if made at all. With new discoveries and accompanying formulations, many distinctions may emerge while others vanish. In the most sophisticated sciences, a minimum of distinctions yield a maximum of explanatory power—think of Newton's three laws of motion that underlie all of classical mechanics—a science of remarkable integrity and scope. Behavior analysis is impossibly far away from that achievement; but, as I will try to argue, it does have a kind of unity that can be revealed by considerations of symmetry.

**What is symmetry and why should we care?**

At its most basic, a system shows symmetry when there exists at least one non-trivial transformation that leaves the system unchanged. Such is familiar with certain geometric figures and shapes. For example, an equilateral triangle rotated 60 degrees by a perpendicular axis through its center is unchanged, and three such rotations return it to its original position. Other figures or shapes may have more or less degrees of symmetry—a circle (or sphere) can demonstrate an infinite number, while most arbitrary forms have none (except, say, a rotation that returns it to its original position—an example of what I call a trivial transformation). While the feature of symmetry with respect to certain figures and shapes is well-known, what is less well-known, but far more important is that the concept applies to principles, concepts, and experimental findings as well. The laws of physics, for example, do not depend on position, time, or, more generally, states of motion—we say they are invariant under these transformations. Because the application of the term ‘symmetry’ in these ways may be unfamiliar, I’ll present a bit more detail on these points before pursuing some examples in behavior analysis.

Lisa Randall (2005) remarks in her popular book on contemporary physics that: “When a physical system has symmetry, you can describe the system on the basis of fewer observations than if the system has no symmetry” (p. 193). But this is not merely a matter of convenience. The very structure of physics is founded on symmetries and the breaking of them—from mechanics to electromagnetics to relativity to particle physics (see, e.g., Park, 1988). But sciences as seemingly far apart as embryology, crystallography, botany, and organic chemistry also embody principles of symmetry.

Einstein’s special theory of relativity was actually occasioned by symmetries in electromagnetism and the theory’s astonishing accomplishments emerge from just two postulates. In terms of symmetry: (1) special relativity asserts that the laws of physics are invariant under translations of position, uniform motion, or time. And (2), the speed of light will be measured the same regardless of states of motion. One implication (of very many) of these two postulates is the equivalence of matter and energy—another symmetry. Einstein’s general relativity, the extension of special relativity to accelerated systems, is an account of gravity involving invariant relations called tensors—mathematical formulations underlying all fundamental laws of physics.

One very deep manifestation of symmetry was shown early in the 20\textsuperscript{th} century by the mathematician Emmy Noether (e.g., Neueschwander, 2010). Her theorem connected the sorts of invariances I’ve mentioned to fundamental laws of conservation—of mass, momentum, charge, and
energy; these laws emerge from such invariances. Virtually all of modern physics is founded on symmetries, or symmetry-breaking. An everyday example of symmetry-breaking is found at a circular dinner table at a banquet. Bread plates are placed to the left of each place setting and there is a circular symmetry around the table. This symmetry commonly leads some diners to be unsure just which bread plate is theirs. However, once one person chooses a plate, the symmetry is broken and all the rest of the diners fall in line, perhaps with some relief. Broken symmetries in nature are not uncommon, though they can be mysterious. Somehow after the Big Bang, what we call matter became overwhelmingly predominant over anti-matter. In biochemistry all amino acids, which could occur in both L and D enantiomers (molecular mirror images), only occur in the L-form in living tissue. I’ll say more about symmetry-breaking later.

How might the various aspects of symmetry be revealed in behavior analysis? There are many if the field and its results are looked at with the sort of perspectives I’ve just discussed. The remainder of this essay is devoted to several examples, some perhaps more convincing than others. There is, of course, no claim to have exhausted the possibilities; no doubt, readers will think of other examples.

One way to provide some organization to symmetries in behavior analysis is to start with the fundamental scheme of the three-term contingency as shown in Figure 1. We may then inquire how symmetry principles apply to each of the components of this scheme—antecedent, response class, and maintaining consequence. I’ll discuss each of these components separately, and then propose why the three-term contingency as a whole illustrates a symmetry.

Figure 1

Three-term Contingency

The functioning three-term contingency is a fundamental concept in behavior analysis and can only be understood as an interactive feedback system—the function of each “element” depends on (i.e., is defined by) the others—an example of circular symmetry. For example, the $S_D$ is defined in relation to the operant class, $R_o$, it occasions, as well as its control being maintained via the reinforcer, $S_R$, which, in turn, acts via feedback to control the occurrence of the operant class. The effect of the $S_R$, in turn, depends on the occurrence of the selected operant class, $R_o$, under control of the $S_D$. The horizontal arrow conceals the contingent relation supporting the interactive system. Newton’s second law as an abstraction shows a similar interdependency. A net force is revealed by the acceleration of a mass. But then how is mass to be defined—by that which is accelerated by a net force? Acceleration is a vector having no meaning independent of net force and mass. Of course, in practice, we can send a spacecraft to Neptune applying this law—it works.
Behavior-stimulus relations--stimulus control

Perhaps the simplest expression of operant stimulus control is embodied in the expression:

\[ B = f(S) \]

That is, to the extent to which the probability of a given class of behavior can be demonstrated to depend on values of an antecedent stimulus, we define the behavior to be under stimulus control. Obviously, there is much left to unfold here, but for present purposes I’ll stick with this beginning.

The function above is best illustrated by a generalization gradient. For example, after suitable training with a single value of a stimulus, one can show that, in the absence of further training, the organism will tend to respond to other values of the original training stimulus dimension (see, e.g., Dinsmoor, 1995). In the absence of imposing further constraints and contingencies, the typical shape of these gradients shows symmetry around the training stimulus; the gradient slopes on either side tend to be equal. Much early work on generalization by Guttman and Kalish (1956) and many others (e.g., Dinsmoor, 1995) showed a remarkable symmetry in gradients around the training stimulus. Later, Honig, Boneau, Burstein, and Pennypacker (1963) showed this effect with inhibitory gradients as well. Similar effects were also demonstrated with punishment (Honig and Slivka, 1964). But, as is well known, generalization gradient symmetry can be broken by prior discrimination procedures along the stimulus dimension of interest to produce an asymmetric peak-shift (e.g., Hanson, 1959).

But the slope of generalization gradient points to another, deeper symmetry—that between generalization and discrimination—steeper slopes emphasizing discrimination; shallower slopes, generalization—each characterized by the rate of change in behavior as stimulus values are changed. Discrimination and generalization are thus two sides of the same coin we call “stimulus control.” Referring to one or the other depends not simply on selected procedures but the degree to which we wish to emphasize differences or similarities in a behavior class as we vary a stimulus.

More complex stimulus control reveals further examples of symmetry, but I will discuss some of these later when I treat the three-term contingency itself as an example of symmetry.

Consequences

In standard behavior-analytic textbooks, events serving as effective consequences are typically organized in terms of (1) reinforcers, i.e., those events said to shape and maintain behavior; and (2) punishers, i.e., those events said to attenuate or eliminate behavior. Within each of these two categories, distinctions are made with respect to procedures, namely whether the reinforcing or punishing effects of events result from their onset or offset contingent on the occurrence of a selected operant class. These distinctions are labeled “positive” (onset) and “negative” (offset). Thus we distinguish positive or negative reinforcers and positive or negative punishers. Clearly, given this scheme, there are several possibilities for demonstrating symmetric (or non-symmetric) effects. Is it the case, for example, that positive and negative reinforcers show equivalent effects (and the same for positive and negative punishers)? Moreover, given that reinforcers can increase the probability of behavior and punishers decrease the
probability of behavior, could we characterize these effects as *anti*-symmetric? In other words, are reinforcement and punishment functional inverses of each other—“mirror images,” if you like?

With respect to positive and negative reinforcers, Michael (1975) and Baron and Galizio (2005) have asserted that there are apparently no significant functional distinctions to be made between the two. If such is the case, then reinforcer effectiveness (by various measures) is *invariant* under a simple inversion of procedure; that is, onset as opposed to offset of pertinent events in relation to a selected behavior class. In my commentary (2006) on Baron and Galizio (2005), I raised the issue of symmetry considerations as well as provided a critique of what I saw as some major conceptual problems. The latter focused on how the authors had framed the putative controlling variables for the proposed equivalence of positive and negative reinforcement. Essentially, they argued that in any given situation where some behavior is said to be under the control of a reinforcing consequence, there can be ambiguity as to whether the behavior is maintained by the delivery of a consequence or the removal of an aversive event. In one of their examples a child turns on a TV and watches a cartoon. Is the child’s behavior under control of the *onset* of a cartoon or the *offset* of, say, boredom? If one believed in the identity of positive and negative reinforcement, what difference would it make? Of course, in this situation, it could be *both*, or *neither*—then what? Moreover, regrettably, Baron and Galizio framed the action of consequences in terms of a limited, non-functional “pleasure versus pain” account. Even if such an account were creditable, as I commented, “…the pursuit of pleasure does not imply the prior condition of pain, nor does the avoidance of pain imply the pursuit of pleasure.” (Marr, 2006a, p. 127)

In any case, there is a sense in which the putative symmetry of positive and negative reinforcement is a trivial assertion—both conditions reflect the action of reinforcement—as shown by numerous studies. But there is a potential symmetry-breaking issue here. “Reinforcement,” however demonstrated, reflects only one property of such events—increasing the probability or maintaining some behavior class. But, for example, reinforcers may show releasing, eliciting, and discriminative actions as well; and for a true symmetry with respect to the onset versus offset procedures, one would have to show equivalence in all the properties of the maintaining events—unlikely, in my view.

With respect to the positive versus negative dimension, to the degree to which a symmetry might apply to reinforcement, one might ask if this also applies to positive and negative punishment. Though there may be a relevant literature here, I’m not familiar with it, so will not pursue this issue further beyond suggesting from the standpoint of symmetry, there may be a functional equivalence. Readers may be able to supply some examples.

Returning to Baron and Galizio’s (2005) pleasure-versus-pain dichotomy, there is some evidence that reinforcement and punishment are, in some sense, mirror images of each other; that is, they are *anti*-symmetric, just another form of symmetry (e.g., de Villiers, 1980; Farley, 1980; Farley and Fantino, 1978). This position is not without controversy as some of the discussions of the Baron and Galizio paper attest. In fact, there is evidence to the contrary, for example, Rasmussen and Newland (2010).

The analysis of putative symmetric relations between positive and negative reinforcement, and between reinforcement and punishment clearly depend on disentangling (1) operations or procedures, (2) the effects of those operations, and (3) the reasons for the effects seen. For example, the case of positive versus negative reinforcement shows that in terms of procedure, the arrangements are reversed (e.g., onset as opposed to offset of events) while the effects are similar (otherwise, why invoke the term “reinforcement” for both?). With the relation between reinforcement and punishment, the opposite applies. Let us assume that we are only considering what are typically called positive procedures. If we adopt Azrin and Holz’s (1966) definition of a punisher, namely as a response-produced event that
decreases the subsequent probability of a given response class, then the procedures for reinforcement and punishment are the same, but with opposite outcomes. One complication here is that to study its effects, punishment must be superimposed on a reinforcement baseline. The asymmetry between punishment and reinforcement is shown principally by a subtractive effect of punishment on a reinforcement baseline.

In all these cases, putative symmetric relations depend on a careful experimental and conceptual analysis and not simply on ostensibly functional definitions. For example, Azrin and Holz (1966) defined punishment as being the opposite of reinforcement, but, by their own experimental analysis, this led to contradictions, namely because, as previously pointed out, consequent events, those we call reinforcers and punishers, typically have multiple effects.

Behavior-consequence relations

Skinner’s opening sentence in *Verbal Behavior* (1957, p.1): “Men act upon the world and change it, and are changed in turn by the consequences of their action” can be, in itself, a reflection of symmetry. Consider shaping. You may recall seeing the old cartoon in the Columbia *Jester* where one rat in a Skinner box remarks to another, “Boy, have I got this guy conditioned. Every time I press the bar down, he drops in a piece of food.” In shaping behavior we recognize the symmetry between the behaviors of the shaper and the shapee. As Skinner pointed out long ago, the behavior of each controls the other in a kind of acquisitional dance (e.g., Skinner, 1972, pp. 122-123). On a more subtle level, we see something similar in feedback dynamics selecting and controlling even complex performances in resonance with their consequences. For example, in treating molar accounts of behavior Baum (1989) discussed how response rate, controlled by a prevailing reinforcement rate (what he calls “O-rules” or functional relations), is in a dynamic dance with how reinforcement frequency is, in turn, controlled by response rate (“E-rules” or feedback functions) (see also Marr, 2006b).

Some of the most compelling reflections of symmetry are revealed in the effects of behavior-consequence relations we call *schedules of reinforcement*. Here, there are numerous examples, and, no doubt, many left to be discovered.

Just as the laws of physics do not depend on states of motion, time, or place, the patterns of behavior engendered under schedules of consequences appear to operate over an enormous range of species—from bees to babies—a biological phenomenon shared perhaps only with certain fundamental biochemical pathways operating from yeast to humans. Moreover, the patterns of responding engendered under schedules can remain invariant under transformations of consequent events, operant classes, manipulanda, as well as species (e.g., Kelleher and Morse, 1968). For example, Barrett and Katz (1981) show fixed-interval performances of squirrel monkeys maintained by food, cocaine administration, stimulus-shock termination, and response-produced shock—all the patterns are identical—without their labels one could not tell any difference.

Long ago, Cook and Catania (1964) showed equivalent fixed-interval performances maintained by food and escape; but, in addition, the administration of a variety of drug classes (e.g., chlorpromazine—“anti-psychotic,” imipramine—“antidepressant,” chlordiazepoxide—“anti-anxiety”) showed the same effects within classes, independent of whether the behaviors were maintained by food or escape—a symmetry under transformation of drug class. While drug-behavior interactions don't always show behavior-consequent-independent effects, there are many cases when drug effects across classes are similar under similar schedules, but with different maintaining events—a challenge to the still common view of interpreting behaviorally-active drug effects largely in terms of motivational variables.
Features of schedule-controlled performance can show a special form of symmetry called *scale-invariance*, that is, changes in scale leave certain properties unchanged (see, e.g., Marr, 2004 for more detailed discussion and many examples). A beautiful example is from Dew (1970) where he showed that certain quantitative features of a fixed-interval performance (the “scallop”) remained constant over three orders of magnitude in fixed-interval value. Other, more subtle, examples are found in comparing performance features of second-order schedules with first-order schedules especially with respect to possible scale-invariance of functional response units (see, e.g. Marr, 1979 for details).

We also see examples of a kind of *symmetry-breaking* with respect to schedule performance. These are reflected in sudden shifts in behavior, either from no responding to responding or the reverse. Both fixed-interval and fixed-ratio responding illustrate the former in that at some point after a reinforcer presentation, a sudden shift occurs from no responding to responding. The reverse effect is seen in performance under large ratio requirements which is characterized by “break and run” patterns of responding. After all the years of the experimental analysis of schedule responding, we still don’t have a good account of these phenomena.

As a final example of symmetry in the context of behavior-consequence relations, earlier I discussed Emily Noether’s Theorem connecting symmetries with conservation laws in physics. In a conserved system, some value remains constant even though elements of the system may take on different values. For example, with colliding billiard balls, even though each ball may change its momentum, the total momentum remains constant, in other words invariant. Physicists describe this situation as an example of the law of conservation of linear momentum. Something analogous has been proposed with behavior (e.g., Baum, 2010; Herrnstein, 1970).

Herrnstein (1970) in deriving his “hyperbola” describing a relation between response rate (or time allocated) and prevailing reinforcement rate, assumed that (1) all behavior is choice, (2), the strict matching law held, and (3) the total behavior in a given situation remained constant, in other words, the behavior is conserved. More recently, Baum has made a similar argument using, for example, a pie chart to characterize behaviors occurring during a given period of time (2010). Presumably, such a conservation of behavior principle would have to assume that all the behaviors of interest would be mutually incompatible!

### Three-term contingency and beyond

In setting up the experimental or applied conditions involving antecedents, behaviors, and consequences, we have some operational (i.e., procedural) latitude in selecting each of these elements. But in a functioning three-term contingency neither of these elements can stand alone; they only have meaning through their mutual interaction; each, in some sense, controls as well as depends on the other. This can be described as a circular symmetry. The old, but still heard, argument that reinforcement is a circular concept and is thus meaningless, can have a point, but misses a much more important point, namely that reinforcement is a relational term whose meaning is found in specific application as in a three-term contingency; in simple words, *it works*. Virtually the identical argument has been made with respect to Newton’s Second Law of Motion. One challenge is to define “force” independently of the other terms. Yet, who could deny the powerful applicability of this law? It is a pillar of classical mechanics, surely one of the greatest creations in the history of science—it, too, works.

The symmetry inherent in the three-term contingency can be extended to *n*-term contingencies, as Sidman’s (e.g., 1994) work with equivalence has taught us. Equivalence classes give us an additional and explicit example—symmetry is a definitive aspect of such relations. Each of the characteristics of an
equivalence relation reflects a kind of symmetry: “Identity” (or “reflexivity”) is what I called earlier a “trivial” transformation; “symmetry” is self-evident; and “transitivity” is an example of circular symmetry.

Sidman also went further in asserting that equivalence relations were established among all the elements participating in the contingency. Thus, many possible symmetries could characterize various n-term contingencies. Presumably, such relations could be extended to relational frames (e.g., Hayes, Barnes-Holmes, and Roche, 2001), but I'll not pursue that here as I am far from expert in this domain.

Coda

Earlier I raised the question of why should we, as behavior analysts, care about a concept as abstract as symmetry. Hermann Weyl (1952, p. 5) in his classic book on symmetry asserts:

“Symmetry, as wide or as narrow as you define its meaning is one idea by which man through the ages has tried to comprehend and create order, beauty and perfection.”

Not only may considerations of symmetry in behavior analysis confer the advantage of our needing fewer observations as Randall (2005) suggested, but they reveal an internal and consistent unity to the field, as well as affirm the enormous range of its applications. The principles of behavior analysis operate at all levels of behavioral and biological complexity—from pigeon key pecks to human cultural practices—an astonishing invariance conferring a beautiful unity to the field.

In a larger sense, I believe the most seductive quality of science is its most elusive—beauty. Without question, this is why I made science my career. As behavior analysts, basic and applied, most of us have been at least implicitly attracted to the field because we have a sense of its unity and its elegant simplicity; through these we strive for plausible accounts of the most complex phenomena we know about—behavior.

References


Marr, M. J. (2004). Dimension in Action and the Problem of Behavioral Units. In Jose´ Burgos and Emilio Ribes (Eds.), Theory, basic and applied research, and technological applications in behavioral science: Conceptual and methodological issues (pp. 151-177). Universidad de Guadalajara, Mexico.


Michael, J. (1975). Positive and negative reinforcement: A distinction that is no longer necessary; or a better way to talk about bad things. *Behaviorism, 3*, 33-44.


