Timberlake's theories dissolve anomalies

Peter R. Killeen

Arizona State University, United States

ARTICLE INFO

ABSTRACT

Two of Timberlake's major contributions, amongst numerous other good notes, are Behavior Regulation Theory (BRT), and Behavior Systems Theory (BST). BRT was a refinement of the Premack Principle. What both got right was that reinforcers are responses, not stimuli. For BRT, they were responses that were occurring below the rate at which they otherwise would given free access to them. BST was a larger ethological framework for our science of behavior. We have always needed it, as it opens an important window on our field. With that window closed, it is easy to stumble over a half-dozen anomalies in the dark, ones that we say *humph* to, scratch our heads, and then move on. When illuminated by BST, however, such anomalies become keys to a deeper understanding of our subject. This paper reviews numerous anomalies that make sense within the joint framework of BST and BRT, and Dickinson's Dual-Process theory of learned behavior. No longer anomalous in that context, all that is now left to do is test the validity and productivity of this general framework for those many strange cases.

1. Introduction

Two of Timberlake and colleagues' major theoretical contributions are seldom considered together. The first of these is his and colleagues' theory of reinforcement, Behavior Regulation Theory (BRT: Timberlake, 1984; Hanson and Timberlake, 1983), aspects of which are known as Response Deprivation Theory (Timberlake and Allison, 1974), Molar Equilibrium Theory (Timberlake, 1980), and Disequilibrium Theory (Jacobs et al., 2019). The second contribution is his and colleagues' situating of behavior at large within an ethological framework, Behavior Systems Theory (BST: Timberlake and Lucas, 1989; Timberlake, 1994, 1993, 2000). These two theories play well together, but Timberlake did little to develop them ensemble (cf. Timberlake, 1993). When they are combined a gratifying simplification occurs, in that some of the outstanding problems which the field has struggled with, or habituated to—anomalies—no longer appear anomalous. But that shift of perspective also shifts us significantly away from traditional theories of reinforcement, a refocus that some may find disorienting.

2. Behavior regulation theory (BRT)

2.1. The problem

David Premack was a creative innovator in experimental, developmental, and comparative psychology. Early in his career he recognized that something was amiss with traditional behavioral theories. The problems were a heritage of operationalism, which blossomed briefly in the early 20th century. SS Stevens was one of the founders of operationalism in psychology, and a colleague of and influence on Skinner. Stevens held that experimental psychology should concern itself with clear (operational) definitions of stimulus and discriminated response, and nothing else. Nothing should be admitted to the canon of science that couldn't be given a definition in terms of recipes for creating and measuring stimuli and responses (Stevens, 1935). In a paper a decade later that borrowed its title from Stevens, Skinner agreed that this was good practice (although he could do better yet, and speak about public unobservables—thinking and imagining and remembering—by rebadging them as responses (Skinner, 1945/1984 Skinner, /, 1984 Skinner, 1945/1984; Killeen, 2019a, 2019b; Moore, 1975; Zuriff, 1979)). For Skinner, responses were what the animal was observed to be doing; responding was the dependent variable, and exteroceptive stimuli, serving both as signs (discriminative stimuli, S^Ds) and as reinforcers (reinforcing stimuli, S^rs) were the independent variables—what the experimenter or the environment was doing to the subject. There are other more flexible behaviorisms (e.g., Staddon, 1999; Burgos and Killeen, 2018), but this is the one that conditioned our behavioral weltanschauung in ways that predisposed some of the anomalies.

This was an ideal scientific scenario: independent variables under the control of an experimenter, with operationally clean definitions of the stimulus and response. But it had limited generality. Sometimes reinforcers don't reinforce. A behaviorist cannot ask his daughter: "What, you don't like my pumpkin pie!?!"; or at least cannot credit the girl's answer that she was just full. Dispositions and internal states such

as likes are not part of the behaviorist's vocabulary. How then does one in general identify the conditions that will make an event a reinforcer, beyond imposing a deprivation regimen and seeing if that works? This is not a promising approach in the extra-laboratory world. This was the problem that Premack tackled; and Timberlake after him.

2.2. The solution

What is the reinforcing stimulus for a game of racquet ball or golf? Is it the pleasure of hearing the opponent say, "You won"? Does winning games when that is not said put your play into extinction, and losing games punish your trip to the links? Hardly. It is the pleasure of playing. It is the pleasure of behavior. It is the pleasure of seeing your responses improving, when that happens. Reinforcers are responses; responses generating environmental feedback.

There is an apparent problem of circularity here. We like to treat reinforcers as independent variables: Surely the reinforcer is a cause, not an effect, we say. Well, a historical cause, sometimes, as generally causes do not come after their effects. Whether food will cause responding depends on the state of the animal on the occasion: Food is an effective independent variable when the animal is hungry. Experimentalists control deprivation regimens to establish such incentive motivation (Edwards et al., 2019). But it is the eating of the food, rather than its mere presentation that matters. With a bit of taste aversion, or a bit of satiation, the organism may respond for the food, but then will not eat it, and responding for it eventually stops. A fine dinner is clearly a stimulus, an independent variable. It provides affordance for a hungry organism to eat, and perhaps to eat slowly, paying attention to each bite, savoring it (Killeen and Jacobs, 2017a, 2017b). That is a different response from wolfing down gruel, or refusing a third helping of dessert, no matter how much one enjoyed the first. When experimentalists do not control deprivation regimens, assuming that a stimulus will be a reinforcer is generally dicey, but can be addressed experimentally with the help of BRT (Timberlake and Farmer-Dougan, 1991).

What kinds of behaviors, then, might serve as reinforcers? Building on prior research showing that rats will learn to press a lever to operate a running wheel (Kagan and Berkun, 1954), Premack struggled with systematic ways to conceptualize reinforcement (Premack, 1959). His eventual solution was that responses that were more probable in a situation will serve to reinforce responses that were less probable (Premack, 1965). When you are hungry, you are more likely to make an immediate date for dinner than one for golf. Dining is more probable than golfing, and you will engage in otherwise improbable behaviors to make that dinner happen. Conversely, a less probable response will punish a more probable one if made contingent upon it (Premack, 1971). Probability, however, has had a checkered career in psychology—as it has in science in general (Stigler, 1999). It is a logical system requiring maps to behavior or other events, and those maps are typically the source of many of the problems involving statistics (Killeen, 2007, 2019).

Skinner held that reinforcement changed the probability of a response, and he took response rate as a measure of probability. But probability and rate are measured in different dimensions: Probability is dimensionless along the unit interval, whereas response rates have dimensions of s^{-1} on the positive line, and can easily exceed 1. Maps between rate and probability *are* straightforward (Killeen et al., 2002), but seldom used in our practice. What Skinner could not say was that reinforcement strengthened an association [as that is an unobservable] between context and disposition to respond, as Pavlov and Thorndike did; or that it affected synaptic connectivity, as they and Hebb did; he could only say that reinforcers would often, if not always, increase probability or rate, which he took as a sign of strengthening [albeit often ephemeral, as that potency is contingent on the state of the organism]. I shall not discuss Premack's solutions to mapping probability to behavior (e.g., Terhune and Premack, 1970), as better ones await in

the next paragraph. Premack soon tired of "respectfully bemused gazes at the ends of his presentations" (personal communication, 1973), and switched his research endeavors to ones with higher probability of reinforcement (e.g., Premack, 2010; Premack and Woodruff, 1978).

2.3. The solution improved

Eisenberger and colleagues (Eisenberger et al., 1967) demonstrated that differential probabilities in baseline conditions would in some cases fail to predict the reinforcement relation: A less probable response in baseline could reinforce a more probable response, if in the testing context, the less-probable response was depressed below its baseline rate. (Premack also commented on that depression, which was one of the seeds of behavioral economics). Of course, baseline and testing are different contexts: What these investigators showed was not so much the failure of Premack's Principle, as the need to measure probabilities in a context similar to that of testing. As Timberlake and Wozny noted, "different factors determine responding in baseline and schedule conditions" (1979, p. 468). Donahoe appositely refined the Premack Principle: Preference "is defined as the proportion of time that an organism exposes itself to the stimuli that control the response when given free access to the controlling stimuli *under baseline conditions which are otherwise identical to the conditions prevailing when the contingency is present*" (Donahoe, 1977, p. 341, who developed the principle to yield several other important laws of behavior). Introduction of the testing protocol is an establishing/ motivating operation (Klatt and Morris, 2001), and thus a "reactive" measurement procedure. An extensive discussion of both the measurement problem, and of the concept of the associated set-point, may be found in Staddon (1979) and Timberlake (1984). The set-point is a hypothetical ideal level of responding; in quantitative models it is a free parameter. This solved the measurement ambiguity, but at the same time adds a degree of freedom to the model, making it more difficult to use and to disprove.

Timberlake and Allison (1974) launched a full-throated attack on the historical concept of reinforcer-as-stimulus, and introduced their own alternative, the *response deprivation* model. The name resonates with the standard motivational operation of *food deprivation,* but unlike it, it emphasizes not the operations of experimenters, but their effects on the subjects' behavior. They adopted Eisenberger and colleagues' (Eisenberger et al., 1967) nomenclature and algebraic model of the deprivation condition:

$$\frac{I}{C} > \frac{O_I}{O_C} \tag{1}$$

In Eq. (1) I is the instrumental response(s) required to make C contingent responses, and O_I and O_C are the "Operant-level" baselines rates of engaging in those responses. Contingent responses (e.g., running in a wheel, drinking water) are those whose emission is contingent upon (requires) executing the instrumental responses. The inequality of Eq. (1) is typically brought about by requiring more instrumental responses I than the animal emits in baseline to access the contingent responses, or a reduction in contingent responses (C; response deprivation). When met, this condition (Eq. (1)), will bring about an increase in instrumental responding, restoring the contingent responses toward their baseline. In older parlance, the instrumental response will be reinforced. Notice that Eq. (1) does not require significant or long-lasting deprivation of the contingent response. If I is highly elastic, so that the animal is relatively indifferent to its emission (e.g., standing in a portion of the chamber near the contingent response location) large increases in I may occur with only nominal transient decreases in C (that is, response deprivation itself can be a nominal precursor to the reinforcement process).

Although Timberlake and Allison (1974) made successful numerical predictions with variants of Eq. (1), a comparison of several quantitative models against several data sets gave the palm to Timberlake's *relative response deprivation* model, used by Timberlake and Wozny

(1979), but never itself published. Timberlake's deprivation model continued to evolve, ending with an apex, if somewhat baroque, mathematical model (Hanson and Timberlake, 1983). The best overview is found in (Timberlake, 1980). A simple exposition of a competing model (Heth and Warren, 1978) provides a tutorial and spreadsheet for the case where both contingent and instrumental responses are measured in the same (e.g., temporal) units and have similar elasticities (Jacobs et al., 2017). Heth and Warren assumed that subjects "respond to the algebraic sum of the two baseline contingency discrepancies" (p. 298) so as to minimize the sum of the absolute deviations of contingent response from its baseline level, plus instrumental response from its baseline level. Their model was generalized by Staddon (2013). Some of these regulation models (e.g., Timberlake, 1980; Staddon, 1979) assume that the total amount of all responding is constant, in some cases weighting the responses by their elasticities before summing. If there are only 2 responses an animal can make, then constraining one must force the other up (a "restriction" effect) independently of any reinforcement process. With many responses available, including doing nothing ("leisure") animals will strive toward an optimal combination of responses, and behave to minimize deviation from that *summum bonum*, called a "bliss point" (Rachlin et al., 1981; Staddon, 2013). *Set-points* can refer to that multidimensional bliss point, but more generally they refer to the ideal levels for each individual response. For a clear introduction to BRT, with examples of application, see (Jacobs et al., 2019). For a demonstration of its power in a clinical setting, see (McFall et al., 2019). Ending this section on behavior regulation theory is an observation by Gawley et al. (1987): In some cases the regulation model fails because animals persist in instrumental responses without availing themselves of the contingent responses they afforded. Such observations will be central to the anomalies discussed in Section 6.

3. Behavior systems theory

Timberlake is more a biological behavioral scientist than a laboratory learning theorist, and he brought that larger view to behavior analysts—many of whom, like adult humans in general, are myopically focused on their own field. Twentieth century behaviorists were largely concerned with levers, keys, and alleys; a few with blinks and secretions. Biological behaviorists (see Davey, 1989; Hinde, 1982; Hogan and Bolhuis, 2009; and Kamil, 1987, for examples), study the topography and coordination of higher-level units. Only a few laboratory behaviorists ventured there (e.g., Balsam and Silver, 1994; Balsam et al., 1992; Timberlake, 1983a, 1983b; Shettleworth, 2009, 1988). The icon of BST is a quasi-hierarchical architecture developed over the years by Timberlake with his students (e.g., Timberlake et al., 1989; Timberlake and Silva, 1994; reprinted as Fig. 1 in Silva et al., 2019). Timberlake (1995) later sketched a system architecture for an "urban, modern human" returning home to prepare dinner. These diagrams have become a trope. They are organizational charts, with a subsystem as its head, modes (perhaps cohering the modules as different emotional states: Gygax, 2017) branching out below, they branching to modules (tactics in acquiring another, depressed, module), finally branching off to actions that instrument those tactics. Because some actions are available from several modules, the architecture is a partial lattice. Because there are different priorities among modules nominally drawn at the same level (the gradients discussed below), these architectures are more heterarchies than hierarchies. For predecessors of them see Baerends's (1976) models of functional organization, reprinted in Shettleworth (1994a) and (Burghardt and Bowers, 2017); for a more elaborated architecture, see Timberlake (1994). For a more precise description of BST than offered here, see Bowers (2017) and Silva et al. (2019). For different behavioral systems see Hogan (1999, 1988), and Fentress (1983). For a general characterization of the hierarchy, Gallistel (1981, p. 609) gives: "Circuits at higher levels govern the operation of lower circuits by … raising the potential [for operation in
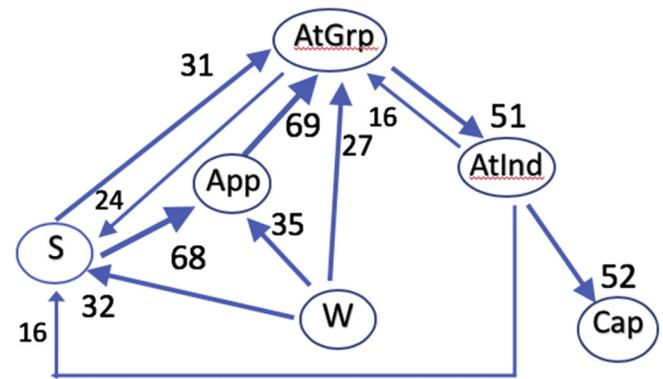

Fig. 1. Modes and modules and their transition probabilities in wolf predation when a pack discovers elk. It is based on the transition matrix in Table 3 of MacNulty et al. (2007). The numbers at the arrow tips show the probability (x100) of transition from one node to the next. *S* is a general search mode involving travelling and scanning modules. The *W* (Watch; fixating on prey while standing, sitting, or crouching) and *App* (Approach; moving toward fixated prey) are focal search modes, where the pack has caught sight of an elk herd, each governing appropriate actions. *AtGrp* is a general attack on the grouped herd, running after or lunging at it while scanning to assay its most fallible member. *AtInd* indicates a focused attack on that individual. *Cap* is capture: biting and restraining the prey, which leads to the kill 80% of the time. (Transitions do not sum to 1, as those with probabilities less than 0.15, such as the probability of 0.12 of going from *App* to *W*, are suppressed for clarity. For a dynamic version of the process, see Appendix A.

some circuits] and lowering it in others–a higher unit establishes the overall pattern to be exhibited in the combined operation of the lower units, while leaving it to the lower units to determine the details of the implementation of this pattern". A different kind of representation, a network, is shown below in Fig. 1.

Markov models can realize aspects of transitions between modules in such architectures (Sanabria et al., 2019; Berridge et al., 1987), lending quantitative dynamism to what is essentially a qualitative and static dramatis personae. Such models will need further elaboration to deal with Timberlake's characterization of behavior systems as evolving "complex causal processes" (Silva et al., 2019) that are context sensitive (e.g., the appearance of prey will cause the program to branch). Biological systems have "many kinds of potential complexities … including path dependence, interaction effects, tipping points, strategic interaction, two-directional causality or feedback loops, equifinality, and multi-finality" (Bennett and Elman, 2006, p. 264; Systems_Biology, 2019). Tools are available for analysis of causal loops (System_Dynamics, 2019), as there are for each of the other complexities, taken one at a time; but interacting, as they do, with the changing affordances of the environment makes their application notional.

A picture of causal loops in episodes of predatory sequences is drawn in Fig. 1, based on the work of MacNulty et al. (2007), who monitored wolf packs preying on elk and bison herds. The most likely tour through this network is from **W**atch to Approach to Attack the Group to Attack an Individual to Capturing, with several of the modules returning to **S**earch. Each step reinforces, and tunes, the prior. Modes and modules may be revisited, contingent on the success or failure of actions in the current module; that is, the network is recurrent. Thus, if an attack on the herd (AtGrp) is ill-fated, the wolves will return to the general search mode (S); if the individual chosen (AtInd) was elusive, the pack may return to harassing the herd (AtGrp), or may break off and return to search mode. We suppose that each of these decisions is motivated by approach to a module that, in the fray of combat, is left with a higher probability of leading to a better state of affairs. These processes are not autonomous, but highly dependent on the dynamic evolution of the environment under attack.

The following extended citation from Maier and Schneirla (1935, p. 254, quoting Whitman, 1912) further illustrates sequential organization

and releasing stimuli (e.g., the lowered head and vibration of the female's tail) that make progression from one module to the next in a pigeon's mating architecture contingent on its interaction with complementary processes in the female:

‘As the period of consummation approaches, the composition of the activities changes with the addition of new elements. Along with bowing, there is billing and fondling of each other's head, hugging or necking, jumping over the female without any attempt at mounting, opening the beak by the male, inserting of the female's beak in his, and often the shaking or rattling of the crop as if the male fed the female. The female stoops with lowered head, the male mounts with a jump, the female raises her wings and lifts her tail, while the male reaches back, moving the tail from side to side until contact (i.e., of cloacal surfaces) is effected.’ The appearance of the final stage seems to depend upon the behavior of the female, in which the vibration or spreading of wings and the raising of the tail are major components in most species.

Ethological *systems* involve all organized behavioral patterns of an organism. *Subsystems*, such as mating or defense or predation (Fig. 1), set motivational states, tune stimulus sensitivities, and prepare motor systems (Timberlake, 1994). One thing leads to another, with the sequential organization requiring environmental affordances to keep it afloat. Another system of behavioral control was noted a century ago by Thorndike in his *Law of Readiness*: "The sight of the prey makes the animal run after it, and also puts the conductions and connections involved in jumping upon it when near into a state of excitability or readiness to be made….When a child sees an attractive object at a distance, his neurons may be said to prophetically prepare for the whole series of fixating it with the eyes, running toward it, seeing it within reach, grasping, feeling it in his hand, and curiously manipulating it" (Thorndike, 1913, p. 53). It is unlikely that Thorndike's work had any influence on the development of behavior systems, yet the parallels are striking. *Modes* are intermediate-level control nodes that attune and ("prophetically") prepare a set of lower-level modules. Through diffuse hormonal regulation they prime perceptual filters and module readiness. Their effects are system-wide, and may permanently retune the system: "The hormones released at parturition change the dam's olfactory sensitivity to pup odors … and make possible the full expression of maternal behavior… If a pup is presented to a virgin female it will ignore, avoid, or possibly even attack and eat the pup" (Hogan, 2017, pp. 69, 73).

Modules "represent groupings of stimulus filters ['fixating'], mechanisms of sensory integration ['feeling it in his hand'], and motor components and programs ['curiously manipulating']" (Timberlake, 1994, p. 408; cf. Killeen and Jacobs, 2017a, 2017b; Ohman et al., 2001). They "combine to select and coordinate individual responses, termed action patterns" (p. 408). "This relatively simple system, of internal factors causing spontaneous search and increased readiness to respond to specific outside stimuli, gives rise to an amazingly varied repertoire of behavior" (Tinbergen, 1952, p. 3). Timberlake gives various examples of pre-organized modules; by varying the temporal and spatial proximity of stimuli to food consumption, he and his students are able to differentially prime different actions that are appropriate to that proximity. For overviews, see the symposium introduced by Timberlake and Fanselow (1994); (Tinsley et al., 2002; Silva and Timberlake, 2005; Bowers, 2018); and articles in this special issue.

Behavior systems theory, then, is the framing of behavior as a (lattice) hierarchy, with drives (in some nomenclature) ranging from the general (e.g., "reproductive") to the more specific (e.g., "courting") modules that prime different affordances and ready different action patterns, as appropriate to their location in the architecture. In the best of cases, the names of the modes and modules are representative of the actions that constitute them, as derived by principle-components analysis (e.g., van den Berg et al., 2003). They are possibly cohered by underlying hormonal changes that humans identify as emotions

(Burghardt, 2019), which ready action and drive attention (Ohman et al., 2001). In describing a nesting gull, Tinbergen writes: "although each of these three responses depends on a specific stimulus situation, they are all dependent on "broodiness" [a subsystem, or mode in the more inclusive subsystem of reproduction] in general. Broodiness is the state into which the bird is brought by prolactin…" (Tinbergen, 1952, p. 4). In all cases, of course, the labels are part of a theoretical structure, imposed hopefully by scientists on fluidly rearranged behavioral structures of their subjects (Killeen, 2013).

Modes provide material causes (the priming of motivationally apposite attention and disposition, often accompanied by characteristic emotions). "Each drive [subsystem] is a hierarchical system, divided in subordinated drives of a more restricted type than the general drive. For instance, the reproductive drive of most birds, if activated, controls a number of activities: courting, mating, defense of territory, nest-building, incubation, etc." (Tinbergen, 1952, p. 3; see Burghardt and Bowers, 2017, for an accompanying system architecture), with each mode's activity governed, in Tinbergen's words, by its own drive, or subdrive [*emotion* is the word that some would use today instead of drive; see e.g., Burghardt, 2019] and ranging from the general at the top of the hierarchy (e.g., reproductive) to more specific as one moves down the hierarchy (e.g., courting, or fighting with rivals), and then eventually to a consummatory action. By preparing a set of modules, and those a variety of actions, the modes are the mechanism for organized variation, providing the fodder for selection by consequences. They are "latent responses": "A historical system, an organism, has a state as well as sensitivity to stimuli and the ability to emit responses. Skinner himself acknowledged the possibility of what he called "latent" responses in humans, even though he neglected to extend this idea to rats and pigeons. Latent responses constitute a repertoire, from which operant reinforcement can select" (Staddon, 2017, p. 27).

In studying the behavior of prey confronted with a threat, Fanselow and Lester (1988, 1994) found differential readiness to behave as a function of the imminence of the threat, analogous to that which Timberlake and students have reported for appetitive responses. In the first case it is a *predator imminence gradient*, and in the latter a *consummatory imminence gradient*. Domjan (1994, Figs. 2 and 3) identified an imminence gradient for a sexual behavior sub-system. There is a substantial literature on such gradients. One study, for example, showed that the various elements of preparation for parachuting were associated with different levels of stress that increased monotonically with proximity to the jump for novices, and increased and then decreased for experts. They suggested that this emotional profile "provided a highly adapted system for the mastery of threat" (Fenz and Epstein, 1967, p. 33), presaging the work of Bolles and Fanselow. The perception of imminence is thus an occasion setter, priming different modules as a function of its strength (Killeen, 1992; Thorndike, 1913, p. 53), down to the point that different behavior patterns are elicited by forward and simultaneous pairing (Esmoris-Arranz et al., 2003).

4. Putting them together

Progress through the modules in Fig. 1, from S through the intermediary modules (W, Ap and AtGr) to the consummatory module of Cap, constitutes movement along such imminence gradients. Each transition places the animal in contexts that release behavior that is more satisfying (Thorndike), more probable (Premack) and depressed below a level at which it would have occurred in that context (Timberlake, Lorenz, et al.). A cat sighting a mouse may be eager to pounce on it; but running is the instrumental response and pouncing on it cannot occur until the running has brought it close. Similarly, it cannot run to the mouse until it has caught sight or sound of it; and it cannot clearly see the mouse until it focuses on it. Killeen (2014, 1992) suggested the obvious: *Each transition through the network is not merely instrumental, but may also become a reinforcing activity*—one that will increase the rate of behaviors that enable the animal to engage in it. In

turn, those actions and accompanying stimuli are establishing operations for a new set of actions at a successor module which, in that context, are depressed below their baseline level. The modules exist on a gradient with a preferred ordering; motion in one direction is reinforcing, in the opposite, aversive (Thomas, 1966)—until the context changes. This is most visible in scenarios such as *Umweg* problems, in which animals must move away from a goal in order to attain it; for many animals, this is difficult or impossible to do (Hebb and Williams, 1946). Just as Premack demonstrated reversibility of the drinking/running reinforcement relation, that also happens in Fig. 1. Wolves would prefer to attack the herd rather than search; but if that attack fails to separate a vulnerable individual, they prefer to return to search than continue abortive culling efforts.

Animals may dally in their actions, linger in the modules. This is either because a necessary releaser for the next step is missing; because a threshold difference is required to shift to a new module; or because the current action has become more satisfying than the next action, for which it historically had been instrumental. Progress through the modules is typically necessary to maintain the reinforcing nature of the instrumental response, but not always. Prey stalking and capture in predators such as cats may become an end in itself, a habit, engaged even when they are satiated. Pigs root when satiated (Moser et al., 2019). People sometimes chase and capture potential mates only to release them once they are emotionally committed. Hobbies provide more examples in which the links of a heterogenous chain that lead to a goal become intrinsically reinforcing: The pleasure of cabinet making is not so much to *have* that chair to sit in, but to *make it* so it can be sat in; the pleasure of catch-and-release fishing is not to eat the fish; the pleasure of sailing is not to reach yonder shore. The chair, the fish, the shore are "MacGuffins", historically necessary to develop the prior chain of behavior, even if its results are more easily and cheaply obtained in other ways. Doing it oneself adds value, both to the labor involved, and the outcome produced. It is a behavioral version of the labor (or Marxist) theory of value, in which the intrinsic value of a thing depends on the amount of socially necessary labor required to produce it; here, it is personal labor and personal value. Appetitive behavior displaces consummatory: "He keeps [his shop] like a hospital lab. He's always either cleaning something up or sharpening something. Last week I found him straightening used nails. He never gets past preparing. Preparing has been his life work. He prepares, then he cleans up" (Stegner, 2002, p. 207). Preparation is appetitive for the construction project in hand; clean-up for the next. Here they occur without the consummation afforded by the construction of that chair or desk. Selective breeding can further sever the links between modules: Pointers will hold a transition to a focal search without approaching, shepherds will approach without capturing, retrievers will capture without consuming.

From BRT we carry forward the recognition that sequences of behaviors, depressed below levels that are optimal in that mode, constitute reinforcing consequences. From BST we carry forward the observation that behavior is organized in modules that prime certain actions, and that those actions, when available, reinforce those which occasioned their availability, often releasing opportunities for actions appropriate at the next module. Various kinds of correlated actions are primed by a module, providing the variation necessary for reinforcement processes to be functional; the details of the actions released will be selected and fine-tuned by operant learning (Balsam et al., 1994, p. 335; Wasserman et al., 1975). In reviewing prey-catching by cats, Hogan (2017, p. 150) notes that "the 'correct' behavior sequences are selected on the basis of their effects ... an operant shaping process can account for all the results, with the proviso that the basic elements—locomotion, angling, pouncing and biting—are not themselves shaped." But even those elements may have been perfected in play by the kitten (Burghardt, 2005; Pellis et al., 2015). Play is a subsystem that inhibits other modes, giving actions an opportunity to be fine-tuned, and in addition making them exportable, "giving the motor mechanisms

an opportunity to become incorporated into other central mechanisms" (Hogan, 2017, p. 203; Pellis et al., 2019; Suboski, 1990). There are many modes that call on the ability to run well, so that that well-practiced module must be reachable from all of them. If, however, the engagement of that module is not path dependent, action slips may result (Norman, 1981). The activation of a mode or module changes context, so that actions such as pouncing that had low probability—were not depressed in the earlier context—are depressed in the new context, such as proximity to a prey object. Actions that get the animal close enough to pounce are selected by reinforcement.

5. Adding dual process theory

Actions that begin as operants can become automatized into more reflexive responses: "A broad spectrum of behavioral routines and rituals can become habitual and stereotyped through learning. Others have a strong innate basis" (Graybiel, 2008, p. 359). In a series of studies in which the contingent response was devalued by satiation or by conditioned aversion to it (Dickinson et al., 1995; Balleine et al., 1995; Adams and Dickinson, 1981), animals often persisted through the action chain of the modules leading to it – only to discover at the end that they had no appetite for the food. This failure of changes in incentive motivation to modulate behavior occurs when the response is over-trained, giving access to the next module with high reliability. It is a manifestation of *habits* that are no longer goal-directed *actions* (see e.g., Dickinson, 2016); they have their own behavioral momentum, insensitive to the current motivational state. The motivational organization of the modes is disengaged. Once stimuli or instrumental responses have activated one of the modules in Fig. 1, the actions run off autonomously, and only after extended experience with the devalued outcome – or perhaps never – do they come back under the control of the motivations that originally made that a relevant goal (Dezfouli et al., 2014).

New operants (still under motivational control) activate the dorsolateral striatum, associated with reinforcement; as they become habitual, control shifts to the infralimbic cortex (Smith and Graybiel, 2013). In an impressive overview of the behavioral and neurophysiological literature, Smith and Graybiel (2014) note that this distinction maps on to the reinforcement-learning/computational distinction between model-free systems (habitual; as not under the control of a model goal; open-loop) and model-based systems (goal-sensitive, instrumental, closed-loop: Dayan and Niv, 2008). They provide neurophysiological evidence that habitual patterns have become "chunked" for ready deployment as a performance unit. It is ironic that, even though Skinner has identified operant psychology as the scientific study of purposeful behavior—behavior under the control of its consequences—such a purposeful character may dwindle with extended conditioning, rendering responses more under control of stimuli that initiate the module than under the control of the consequences that shaped it. The three-term contingency of S-R-Reinf shrinks to S-R. For a recent more general overview of habits, see Wood and Rünger (2016).

In such ritualization, either variation in aspects of the reinforcing (contingent) behavior have no effect on the quantity or nature of the instrumental responding (no motivational control); or the contingent behavior may not differentially select variations in the quantity or nature of instrumental responding (no selection by consequences; Dickinson and Balleine, 1994). In either case the correlation between response and outcome goes to zero. This is the heart of *Dual-Process Theory* (DPT; Dickinson et al., 1995): Initial goal-oriented responses are, with practice and regularity, transformed into autonomous action patterns under the control of prior, but not consequential events (Watson and de Wit, 2018). Habitual behavior is most likely when the covariation between what the animal does and what it gets as a result decreases, as it may when an instrumental response pattern has been honed by its consequences to reliably generate them (Dickinson, 1985). Habit formation, much like sensory habituation, requires regularity of

conditions: Substantial variability in probability (Thrailkill et al., 2018) or timing (Colwill and Rescorla, 1988) may keep actions goal-oriented. Apparently, habitual reflex-like patterns are more computationally cost–effective than goal-sensitive patterns, and the latter morph into the former when their ability to secure the goal becomes routine. This is akin to the observation of William James that "It is a general principle in Psychology that consciousness deserts all processes where it can no longer be of use" (James, 1890/1983James, /, 1983James, 1890/1983; Compare James's "tendency of consciousness to a minimum of complication" with the minimum free-energy principle of Friston, in, e.g., Ramstead et al., 2019)

Insensitivity of instrumental responding to reinforcement was found in a study of pecking in an auto-shaping paradigm in which food was presented at the end of a CS on a percentage of the trials (in most conditions 10%), with food presented independently of pecking (Killeen et al., 2009). The linear model used to predict pecking on the ensuing trials gave greatest weight to the Pavlovian factor of presence of food on a trial. The weight for absence of food on a trial was very close to 0, it thus having little or no effect on ensuing behavior (a phenomenon we shall see recur in Section 6.6). For some conditions the constellation of parameters predicted that in continued absence of food, response probability would decrease to an asymptotic non-zero level. The data honored that prediction for at least 50 trials, as shown in Fig. 2, following the model curve.

The probability of a response on a trial had essentially zero correlation with the number of responses on the previous trial, whether those preceded an opportunity to eat ($r = .035$) or not ($r = .081$; p. 457). The failure of response rate to be affected by its outcome (e.g., a high rate on a trial followed by food did not bring about a higher than average rate on the next trial), suggests that pecking had become habitual. (A test by devaluing the reinforcer was not, unfortunately, conducted; cf. Smedley and Smith, 2018). Shull (2004) and Brackney et al. (2017) showed that responding on interval schedules after extended training is comprised of bouts, with initiation of a bout sensitive to motivational variables, but behavior within the bout insensitive; this also suggests motivational control of entry into a module, with behavior within that module becoming habitual.

Stimuli may become both incentive motivators of actions in modules, and occasion-setters for other actions occurring later in the same mode. Their role as an occasion-setter is enhanced when they are at some temporal remove from a modular transition—i.e., in the case of a long CS (Holland, 1992). Shorter CSs are more likely to support habitual behavior. (See Bonardi et al., 2017 for a recent review of the

substantial literature on the related Pavlovian-instrumental transfer.) Occasion setters often remain sensitive to outcome devaluation by satiation—they are goal driven, even when the specific modules that they modulate are not. The same is true for early links in a behavioral chain (Corbit and Balleine, 2003) as sign-tracking may be (Derman et al., 2018). Thrailkill and Bouton (2016) review these types of effects in heterogenous chains that constitute movement from one module to the next down the gradient. Because occasion setters remain under motivational control, they can protect organisms from engaging habitual modules. Once an animal becomes engaged in a habitual module, however, motivational control fails. Occasion setters are the angels saying "Don't go there"; CSs are the devil saying "In for a penny, in for a pound" (Platt, 1972). For a recent look at occasion setters, see the special issue of *Behavioral Neuroscience* (Fraser and Holland, 2019).

Repeated actions, then, are habits in the making, decoupling from the original motivation for engaging them: In the vernacular of BRT, their setpoints have changed. The well-replicated observations of dual processes by Dickinson, Balleine and colleagues add elements to both of Timberlake's theories that are necessary for them to address some of the anomalies in our field. BRT provides the causal mechanism that motivates instrumental responding; BST embeds this in a larger ethological framework, one that is sensitive to the particular mode of response and relevant eliciting stimuli, and setting the stage for an imminence gradient; DPT explains how habitual responding may become an end in itself, holding the organism in a module despite the opportunity to move on to the contingent response that originally motivated it.

6. Anomalies

Anomalies are phenomena that are abnormal, inconsistent with expectations and not obviously reconcilable with them. Dark matter and dark energy are invisible entities that have never been detected, and may never be directly observed, but constitute the majority of matter/energy in the universe. They were introduced to explain other problematic observations. Dark matter was a fix for the anomalous speed of rotation of galaxies, inconsistent with Newtonian mechanics. Dark energy was a fix for the anomalous acceleration of the expansion of the universe. "No observation in the last thirty years has been more upsetting than the discovery of the dark energy in 1998. We do not know what the dark energy is…it is just there" (Smolin, 2006, pp.149-150). Nothing changed in the ensuing decade. Wave particle duality and quantum entanglement (which Einstein called "spooky action at a distance") are also anomalous, inconsistent both with common sense and long-held world-views: "If you think you understand quantum mechanics, that means that you don't understand quantum mechanics" (attributed to Richard Feynman).

Anomalies will often drive scientific research until they are resolved, or, as in the case of quantum mechanics, until our expectations habituate. Such habituation happens with anomalies in our own field, and even with its literature: Many young members of our field may never have seen a reference to contra-freeloading or quasi-reinforcement. Great scientists notice and are bemused by what perhaps only they see to be anomalous. Einstein found the precise equality of inertial and gravitation mass anomalous—it really could not happen by chance—and making sense of it is what drove his general theory. Theoretical physicists contend with many anomalies (e.g., Fujikawa et al., 2004); Einstein died still believing a hidden variable might be found that would resolve the anomalies of quantum mechanics. Not all unpredicted events or observations are anomalous: A problematic observation becomes an anomaly when it runs counter to established theory or assumptions. Such observations are therefore often difficult for the field to accept (Garcia, 1981); one reviewer, an expert in the analysis of delayed reinforcers, called the long-delay taste aversion demonstrated by Garcia and colleagues "no more likely than birdshit in a cuckoo clock" (Freeman and Riley, 2009, p. 28). There are many anomalies in our science of animal behavior, most drawing only a blind
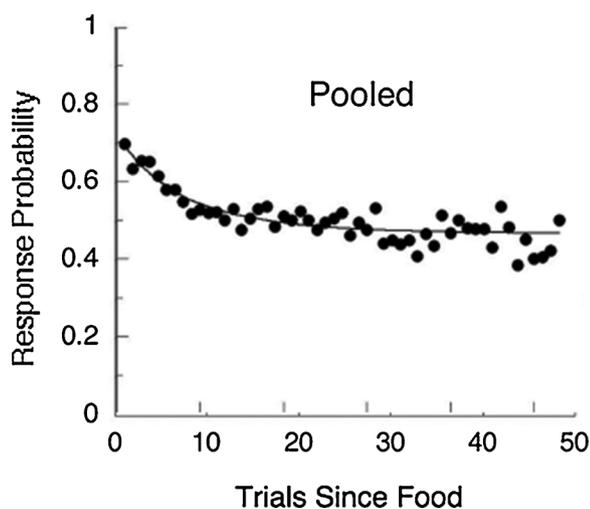

Fig. 2. The persistence of habitual responding in a percentage reinforcement sign-tracking paradigm, pooled over 5 pigeons (from Fig. 8 of Killeen et al., 2009; reproduced with permission).

eye; a few others, bird epithets.

6.1. Misbehavior

The Brelands' classic article and book (Breland and Breland, 1961, 1966) on the intrusion of action patterns into chains of behavior that they had conditioned took many students like myself by surprise; it was anomalous to those of us who had given little thought to the ecological niche of their subjects, or the possibility of "motivational infiltration of conditioned responses and conditioned stimuli": "The contamination of response classes by their reinforcers (or the drives underlying the reinforcers) was the main source of evidence against Skinnerianism used initially by the Brelands (1961)" (Herrnstein, 1977, pp. 600, 599). Sometime earlier Tinbergen had called such intrusions *derived activities* (1952). The title of the Brelands's article, *The Misbehavior of Organisms* was a riff on Skinner's *Behavior of Organisms:* Nothing in that masterpiece predicted anything like the anomalies that these expert animal trainers contended with. Misbehavior was house-broken in the laboratories of Boakes (Boakes et al., 1978) and Timberlake, who aptly defined misbehavior as "unnecessary, species-characteristic behavior that delays or prevents reward" (Timberlake et al., 1982, p. 78). It was used as an analytic tool by Timberlake and his students to dissect epochs between feedings into constituent modes.

This shift in perspective, from viewing behavior as constructed by associative processes, to the evocation of action patterns by predictive CSs, was becoming an accepted position by the end of the century, well after Lorenz observed "If J. B. Watson had only once reared a young bird in isolation, he would never have asserted that all complicated behaviour patterns were conditioned" (Lorenz, 1950, p. 233). Jenkins and colleagues (Jenkins et al., 1978) argued that the "experimental CS–US episode mimics a naturally occurring episode for which pre-organized behavior patterns exist. …the artificial signal substitutes for a natural signal, not for the object being signaled as in the Pavlovian concept of substitution" (p. 292). Such actions "are not a product of the conditioning process itself, but are imported from the species' evolutionary history and the individual's pre-experimental history" (p. 294). Thus "Pavlovian conditioning came to be seen as a process through which a whole behavior system may come under control of new causal factors" (Shettleworth, 1994b, p. 362); also see Domjan, 2005; Timberlake, 1983b; and Suboski, 1990).

Why do we not see more of these intrusions in our experimental chambers? Because Skinner and others were so skillful at "tuning" the apparatus to engage only a homogenous set of responses: "In my experience of this process, the type of intruding misbehavior reported by Breland and Breland (1961) is more the rule than the exception in seeking to reinforce new behavior or similar behavior in a new species" (Timberlake, 2001, p. 86). How often do we look, in any case? Lorenz praised H. S. Jennings and J. von Uexküll: "both of them hold that the observation *of all there is to be observed* in the behavior of a species must go *before* the quest for explanation of single items of behavior. … To Jennings we owe the conception of the *system of actions* of a species" (Lorenz, 1950, p. 233). To Timberlake we owe the introduction of field-biology into behaviorism (Timberlake, 1999).

All of the actions in Fig. 1 derive from nodes that are differentially primed by the spatio-temporal distance to consummation and signs of it; they are also differentially induced from the system's architecture when the conditions of progress through the hierarchy get stalled. Brelands's chickens innately scratch the floor rather than peck the piano keys when a peck does not immediately yield food. This derived activity can be observed in chicken-pens and Skinner-boxes alike (Marley and Morse, 1966; Su et al., 2018; Killeen, 1975, Fig. 9). When the Brelands gave their animals inedible tokens, the pigs rooted them and the racoons washed them, as these were the modules tuned by the particular spatio-temporal proximity to food that the demonstrations had arranged. They were anomalous because at the time Skinner, his students, and many others, took the position that behavior was pliant and largely controllable through contingencies of reinforcement; but the Brelands, expert trainers who had not set out to reinforce those species-typical responses, found that they displaced the operants they were trying to shape. The response to these and other anomalies (e.g., selective associability and long-delay taste aversion learning) generated the literature called "constraints on conditioning" (e.g., Domjan, 1983; Domjan and Galef, 1983; Seligman and Hager, 1972).

Adjunctive responses induced by a schedule involving consummatory actions (Segal, 1972; Baum, 2012) are often amplified by proximity to consummation. The actions are organized in time by: a) their differential memorability (intrinsic "marking": Lieberman et al., 1979, 1985; Williams, 1991); b) the differential habituation of cues that induce those responses (Balsam et al., 1994, p. 337); and by c) their competition for expression (Killeen and Pellon, 2013; Pellon and Killeen, 2015; Killeen, 2014a, 2014b). In one of the misbehaviors that I studied (1975), adjunctive schedule-induced aggression would often carry-over through the operation of the feeder, displacing the consummatory response that aroused that aggression. BST offers a systematic framework for thinking about and analyzing such actions (Burghardt and Bowers, 2017); one that offers pleasure, not surprise or suppression, when misbehaviors appear as uninvited guests to our chambers. The first step is to understand the species–typical behaviors of one's subject (Lorenz, 1950, p. 233), and observe and manipulate the affordances for those responses (Cabrera et al., 2013), as Timberlake and Lucas did so effectively in their analysis of superstitious responding (Timberlake and Lucas, 1985).

6.2. Contra-freeloading

Many psychologists, behaviorists, and ethologists implicitly assume (and often explicitly assert) that there is some economic sense to animal behavior: that animals settle into routines in which some currency is optimized. The currency could be a benefit/cost ratio of calories-in over calories-out; or calories-in over opportunity-cost; or calories-in over predator-exposure; or minimization of deviation from set-points as in BRT; etc. Misbehavior foiled that expectation but could find something of a niche in an ethological framework. Unfortunately, that framework by itself does little to help make sense of contra-freeloading. It surprised everyone (except dog owners) when animals were found to *like* to work, to make instrumental responses for food that was freely available. It amazed a generation, then was largely forgotten. As Inglis and associates noted twenty years ago: "Contra-freeloading has often been ignored because it appeared to contradict the basic tenets of prevailing theory. This is, however, a reason for investigating rather than ignoring it" (Inglis et al., 1997, p. 188). Judging from recent citations to it, the reverse happened. The first study clearly demonstrating the unsettling fact that came to be called "contra-freeloading" was Jensen's (1963), who trained rats to lever-press for pellets, giving groups 40, 80, 160, 320, 640, and 1280 training trials. He then gave them a cup of free-food at the back of the chamber. The rats lever-pressed for pellets despite the ample free food, and they did so as a monotonically increasing function of training trials. Only 1 of 200 rats chose to eat only the free food. Neuringer (1969) replicated the effect with pigeons; with 7 training trials they continued making from five-hundred to one thousand pecks every day to operate the food hopper. In an additional experiment they learned to either lever-press or key peck without prior training and continued to respond for food in the presence of free food. With an empty hopper, key pecking extinguished. Pecking for food in the presence of free food could be maintained even when it was on a VI 3 min schedule (Neuringer, 1970; VI 3 schedules deliver a reinforcer according to a random distribution of intervals averaging 3 min, and contingent on a response). An experiment by Sawisch and Denny (1973) is of special interest in the context of BRT. They replicated Neuringer's results, and then made availability of key-pecking for food (now a depressed contingent response) contingent on eating the free food—and found an increase in eating free food (the instrumental

response) when access to earning food was made contingent on it! It is as though a roadside beggar held a sign saying "Will eat for work".

A roadside zoo of animals, including humans, have been observed by many investigators to work for food when food was freely available (see Table 1 in Inglis et al., 1997). Osborne (1977) provides an authoritative review and analysis, and Inglis and associates an update and model. Osborne observed "That animals learn and maintain an operant response for food when free food is available has appeared inconsistent with general tenets of reinforcement theory" (p. 232); that is, has appeared anomalous. The premise of the present paper is that actions within a module can become intrinsically reinforcing, based on, but superseding, their ability to provide access to the consummatory response. A new setpoint for that response has been established; it has become *habitual*, as in Dual Process Theory. For this to happen, the habitual response must be well-practiced, and must belong to a different module than the ones above it (toward the top of the hierarchy) and below it (toward the bottom of the hierarchy) in the architecture. The importance of practice is illustrated by Jensen's (1963) demonstration that responding for food increased from 20% to 75% with training, and is seen in other of the studies reviewed by Osborne. The responses must be different from one another, else no transitions from instrumental to contingent can occur. There are other hallmarks of habitual responses: Animals may fail to consume a substantial portion of the food that they earn (it is the earning that matters); and considerable responding is maintained in the absence of food deprivation (Osborne, p. 225). Furthermore, when given topographically similar responses—running down a long straight alley vs. down a short one, with the former interpreted as earned, the second interpreted as closer to free—rats (rationally) preferred the short alley (Inglis et al., 1997 p. 1172). When rats had to *explore* to find food, however, they preferred to do that than go to a known location of free food (Inglis et al., 1997 p. 1176). Actions in the same *module* do not work to reinforce others in that module, because they are not a step closer to consummation, at the bottom of the architecture; they are freely available, so are not depressed below their set-point for that module. The transition matters, and stimuli that cue the transition become incentive motivators. That the action must belong in the same *subsystem*, however, is indicated by the much lower incidence of preferring earned water over free water. Manipulation of an object such as a lever is an unlikely part of the system architecture of thirst reduction.

A key way of differentiating instrumental responses from consummatory ones is by concomitant stimulus change. When entry into one free-food locale was accompanied by ambient illumination change, the pigeons preferred it over an unsigned locale (Osborne, 1977, p. 227). Rats' lever-presses for pellets were best maintained when punctuated with a concomitant stimulus change, even though such a response is already topographically different than the consumption of free pellets (which is presumptively adequate stimulus change in itself, cf. Alferink et al., 1973). In another experiment, eliminating one differentiating stimulus (the sound of the dispenser) by operating it *both* upon free food or earned consumption reduced the preference for lever-pressing: It made the responses less distinct. The preference for signaled reinforcement over un-signaled (e.g., Badia et al., 1981) could be construed as providing an opportunity for distinctive instrumental behavior during the signal (Timberlake's "extreme focal search"), unnecessary though that may be, drawn from a different module than the consummatory response. The signal encourages behaving *en route* to consumption, which, I argue, can become intrinsically reinforcing.

6.3. Activity-based anorexia

That rats take exercise in a running wheel is not strange; many animals do (Sherwin, 1998). Rats given access to a running wheel in a situation where food is given once per day, however, will typically increase running and decrease eating to the point where they do not consume enough of the available food to maintain a viable weight

(Routtenberg and Kuznesof, 1967); many will die if not removed from the experiment in time (Epling and Pierce, 1996a, 1996b). It is an anomalous effect—Boakes (2007 p. 211) calls it "paradoxical"—because no rational behavioral or economic framework would predict that rats would starve themselves to run, and fail to compensate by eating more of the available food to stay alive. Rats in the wild will run in a wheel (Meijer and Robbers, 2014), but do not starve themselves; yet standard protocols (e.g., Carrera et al., 2014) reliably produce anorexia in the laboratory. The amount of running negatively covaries with the amount of food available—when there is less food rats run more (Pierce and Epling, 1996), and when they are food-deprived they run more. Because of this relation, well-illustrated in Boakes et al. (1999); Boakes (1997) calls such running a *species-specific hunger reaction*; it is the travel module at the top of the classic Timberlakian architectures, analogous to S in Fig. 1. Dwyer and Boakes (1997) demonstrated that body weight will eventually begin to recover with pre-adaptation to the feeding schedule, the handle/consume mode will condition to time and context, attuning the rats to move into that mode under the control of circadian or other temporal cues. A recent overview and test of theories of such *Activity-Based Anorexia (ABA)* is found in (Labajos and Pellón, 2018).

Running is a general search action (Mather, 1981). It is over-practiced, a precondition for becoming a habit: Its set-point is changed. At the end of a bout of running when food becomes available, the rats have not switched out of their general search mode: They eat less food than they need, and they work less hard for that food on progressive ratio schedules than they normally would. Dwyer and Boakes (1997) got better temporal control of the once-daily feeding regimen by over-training the Handle/Consume mode; and also by temporally isolating the general search (wheel) by a 4-h delay to handling and consumption; and by getting stimulus control of the modes by limiting running/search to the bright day epoch and consumption to the dark night epoch. As is the case with general search modules such as ABA, focal modules such as hopper approach can become habitual, and in some cases block control by prior modules. Timberlake noted that "increased pretraining with food delivery prior to pairings of the bearing and food also blocked behavior directed at the bearing [i.e., misbehavior] in favor of behavior directed at the food tray" (Timberlake et al., 1982, p. 81). Pretraining can also significantly interfere with acquisition of sign-tracking in more conventional paradigms (Downing and Neuringer, 1976).

That running has become intrinsically reinforcing is suggested not only by its prevalence, and its indifference to temporal periodicities of reinforcement (Belke et al., 2018), but also by the demonstration that it induces the release of endogenous opioids: Kanarak and associates (Kanarek et al., 2009, p. 905) injected their running rats with naloxone, an opioid antagonist, and observed a "direct relationship between the intensity of running and the severity of withdrawal symptoms". If one of the general search modules in BST is amenable to such addictive habit formation, are there others that can also be so affected? (There is of course good evidence that consummatory actions may become addictions, e.g., Hebebrand et al., 2014). Is all habitual behavior addictive (i.e., self-reinforcing) behavior? Kanarek et al.'s experiments provide a paradigm for testing that hypothesis.

6.3.1. Drugging, running and pecking

Berridge and associates (e.g., Berridge, 1996; Berridge and Robinson, 2003) have made a sustained case for the distinction between incentive motivation and hedonistic attraction, while demonstrating that other paradigms such as Pavlovian and Skinnerian cannot make sense of the nature of addictions (e.g., in Robinson and Berridge, 2000). They argue that what drugs of abuse do is increase the motivational salience of cues that have in the past been associated with drug use. The cues more than release instrumental behavior; they come to elicit it. The addict will deeply want the fix, whether or not he any longer gets much pleasure from it—studies have shown that wanting increases with

use, and liking often does not, and sometimes decreases. Berridge and colleagues have shown that different brain pathways are associated with wanting and liking. Perhaps it is liking that induces the approach (Perkins, 1968), but wanting that keeps the organism doing it. Rats practiced in running do not work harder to gain access to a wheel than less practiced rats (Cordony et al., 2018); yet they run excessively when allowed. Even though running may become habitual, insensitive to outcome, approach to the wheel is not thereby more highly valued. Conversely, rats with little access to a wheel will work harder to obtain it; such instrumental responding appears to remain sensitive to the outcome.

Just as the case with drugs of addiction, there is opioid involvement with some habitual behaviors such as running; and there is persistence in others, such as sign-tracking, even when the animals are given a conditioned taste aversion to the food they are (ostensibly) responding for, and thus do not consume. Saunders and Robinson (2011, p. 1668) found that "rats prone to attribute incentive salience to a food cue [i.e., were more vigorous sign-trackers] worked harder for cocaine, and showed more robust cocaine-induced reinstatement". Even absent a taste aversion, animals in contra-freeloading paradigms will often respond for food only to then disregard it, just as cats will kill birds without consuming them, and food-cues will prompt humans to work for food when they are satiated (Watson et al., 2014). There are haunting commonalities among these anomalies—their compulsive nature, triggered by cues, free from control by central motive states. They may be related to abnormal repetitive behaviors of clinical significance (Zabegalov et al., 2019), and, most obviously to drugging addiction (Zapata et al., 2010).

6.4. Sign-tracking

Brown and Jenkins (1968) reported that pigeons would peck at a key if it was illuminated several seconds before the free delivery of food. They called this "auto-shaping", subsequently rebadging it "sign-tracking" (Hearst and Jenkins, 1974). We have lived with the phenomenon so long and intimately (Google Scholar returns 4500 articles containing one of those names, although some of them (e.g., Zimbardo, 1973) refer to a different phenomenon) that it no longer seems anomalous. But at the time it was: Pavlov's ghostly hand directing the pigeons to make a paragon ¿operant? response in Skinner's box, flipping the classical bird at selection by consequences. Like the earlier anomalies, sign-tracking provides another example of time and energy expended in activities unnecessary for reinforcement, and in some experiments interfering with reinforcement (Hearst and Jenkins's (1974) "long box" experiment; Williams and Williams, 1969; Stiers and Silberberg, 1974), satisfying Timberlake's definition of misbehavior. But just *why* do pigeons peck and rats manipulate and humans touch? Stimulus substitution? Perhaps; but that fails as a general account of Pavlovian conditioning (Timberlake and Grant, 1975; Wasserman, 1973). Jenkins and associates observed that in the case of the dogs that they were studying, "In no case could the observed action patterns be accurately described as a copy of the consummatory feeding response to the food itself" (Jenkins et al., 1978, p. 291). Viewing the *CR* (traditionally, *Conditioned Response*) as the *Conditionally Released* actions that are either conditioned or imported from the animals' evolutionary history "makes understandable the observation that induced actions are often signal-centered and are often appetitive, or reinforcer-*procuring*, rather than consummatory, or reinforcer-*consuming*" (p. 294); that is, it belonged to one of the higher modules (e.g., those entailed by general search) in the architecture, rather than being a short-circuit to the lowest consummatory module. They were instrumental hypotheses, rather than consummatory fantasies. As such, their nature depends in part on the nature of the stimuli that release them (Holland, 1977), as well as the ensuing actions that they typically empower.

What would Timberlake say? In an experiment pairing distinctive ball bearings with food or water, he found that "food-related bearings produced more complex, although not more frequent, interactions than did water-related bearings. In none of the experiments did rats lick the ball bearing related to water. The results supported a behavior-system approach, but not the stimulus-substitution or arbitrary-operant accounts of conditioned-response topography" (Timberlake, 1983a, 1983b, p. 309). Timberlake also noted that a CS duration of 2.6 s encourages "extreme focal search and handling-consuming modes" (1994, p. 214); and in another condition (the "Programmed-Exit" group[1]) a stimulus of 5.1 s elicited behaviors "analogous to goal tracking (Boakes, 1977)" (Timberlake et al., 1982, p. 66). By "extreme" Timberlake meant modules that were very close to consummation (here, eating; see Timberlake, 1994, Fig. 2c)[2]. Goal-tracking is not necessarily a different kind of behavior; it is sign-tracking to the hopper. Longer CSs promote approach to the signal, shorter ones approach to the goal, due to response competition and different delay of reinforcement gradients (see, e.g., Figs. 2 & 3 of Killeen, 2013). Would even longer (18 s) duration CSs elicit general search behaviors? Yes (Holland, 1980; Silva and Timberlake, 1998; Akins et al., 1994; Silva and Timberlake, 1997). Does the control by temporal proximity interact with spatial position of the CS? In the case of food, yes: As predicted and interpreted, using levers as CSs with different spatio-temporal relations to the delivery of a pellet, results are consistent with behavior systems theory (Silva, Timberlake, et al., 1998). A map of sexual behaviors elicited over time is drawn by Akins et al. (1994) who note some differences from behavior induced by food, in particular the fact that spatially distal stimuli are as likely to induce focal/consummatory behavior as when they are proximal.

Timberlake's (1983b) demonstration of rats' sign-tracking to rolling ball-bearings was an early prelude to his analysis of spatial control. A recent sequel used pigeons and lights on a touch-screen with lights moving toward the hopper, away from the hopper, or stationary along the path (Cabrera et al., 2009). Pecking was greatest for lights moving toward the hopper, intermediate for stationary lights, and lowest for lights moving away from the hopper. Rate of pecking increased as exponential functions of proximity of light to hopper in all cases. The standard Pavlovian hypothesis of stimulus substitution does not explain the nature and richness of these effects as well as do the gradients along the modules in BST.

Sign-tracking shows orderly change as a function of trial and intertrial durations (e.g., Gibbon and Balsam, 1981; Gibbon et al., 1977). Jenkins et al. (1981) summarized the implications of his sign-tracking experiments: "a promising candidate [explanation] is reinforcer waiting times in the CS as compared with the context" (p. 252). Jenkins's hypothesis, and a model of competition among actions instigated by

---

[1] There were 4 groups in Expt. 1. The "programmed exit" group received food at the end of a 5.1 s interval, about when an unimpeded ball would leave the chamber. The "actual exit" group received food only after the ball bearing had left the chamber. The "random group" received ball-bearings and food independently of each other. Group "CS-only" just got the ball. The actual exit group generated much more interesting data: When the food did not come until the ball left the chamber, there was vastly more "misbehavior" directed at the ball. Apparently, temporal proximity of ball manipulation and food delivery reinforced the ball manipulation, a point made tediously clear by Killeen and Pellon (2013). Absent such strong reinforcement of misbehavior, when the correlation between ball manipulation and termination of the CS is more variable, the attraction to goal at short CS durations can be more powerful than the attractions to the signs of them. Reinforced misbehavior will often outcompete attraction to the goal, but when reinforcement contingencies are not ideal for misbehavior, goal attraction may prevail (Pellon and Killen, 2015).

[2] Bill's descriptions of the behavior were more colorful than his graphs of it. With a 2.6 s interval between presentation of a ball-bearing and food, "many rats figure out ways to combine holding the bearing [sign tracking] and going to the feeder [goal-tracking]. … several rats placed the ball bearing in the feeder, picked up the pellet, and then picked up the bearing again and held it and gnawed it further"; at 7.6 s or more "The rat treats the ball bearing as a prey item, chasing, seizing, carrying it and gnawing it" (Timberlake, 1994, p. 412).

stimuli and strengthened by reinforcement (Pellon and Killen, 2015; Killeen, 2011) makes it possible to map the major quantitative features of sign-tracking (see, e.g., Eq. 3 of Killeen, 2013). This model is a quantitative realization of the observation by Timberlake and colleagues: "The variation in expression of appetitive behaviors depended on the nature of the contingency and the resultant competition for expression with other behaviors" (1982, p. 84).

Motivation plays an important and subtle role in sign-tracking (Anselme, 2016). Well-practiced sign tracking may become habitual, independent of the hedonic value of the outcome (but see Derman et al., 2018). It persists despite non-reinforcement (Williams and Williams, 1969; cf. Sanabria et al., 2006). Sign-tracking, but not goal-tracking, has been shown to be resistant to outcome devaluation with LiCl: rats would lever press, but not consume the ensuing devalued pellets (Smedley and Smith, 2018). A similar experiment using sucrose (Morrison et al., 2015) achieved a decreased consumption of sucrose, and an *increase* in lever-pressing relative to goal-tracking, as though the rats were loath to approach the sucrose. Tomie and associates (Tomie et al., 2008; Tomie and Morrow, 2018) have noted the behavioral and physiological similarities of sign tracking and drug abuse. Like many other ano malous actions, sign-tracking can become autonomous, the modules taking on a life of their own.

The most fundamental law of behavior, *approach,* seldom appears in our theories (cf. Alcaro and Panksepp, 2011; Buzsáki, 1982; Perkins, 1968; Wasserman et al., 1974), as it is typically overshadowed by the law of effect and other such formulations (Teigen, 2002). Animals approach opportunities to descend a gradient (Burghardt and Burkhardt, 2018; Craig, 1918); moving toward both stimuli and responses that enable those transitions (Deutsch, 1960; Baum, 1974). "All of the diverse positively motivated behaviors exhibited by animals [involve] the primal tendency for an animal to move from where it is to where it must be to acquire materials needed for survival" (Panksepp, 1989, pp. 12-13). Signs for "where they must be" are incentive motivators; it is the "acquisition of materials", depressed in that mode, that reinforces. Such signs of modules lower on the imminence gradient are approached because they lead to responses depressed below their setpoint in that mode. Access to those contingent responses reinforces the instrumental responses that lead to them; and so on down the chain until the act of consummation, which may then reset the motivational modes. Behavior is a trajectory through fields of such attractors (Killeen, 1989). Actions that are over–practiced in those fields gain momentum, and may become autonomous.

6.5. Quasi-reinforcement

Animals will learn to make responses for stimuli that have been paired with reinforcement (Sosa et al., 2011), even in contexts or at locations where primary reinforcement is never obtained. For nice demonstrations of such conditioned reinforcement, see Zimmerman (1963, 1969) and Horney and Fantino (1984). Working for conditioned reinforcers in the signaled absence of primary reinforcement is strange, but not quite anomalous. After all, which stimulus should an animal believe—focal or contextual—especially when the animal is needy? It makes sense to take a chance, just as it does in the case of superstitious behavior (Killeen, 1978). Animals will also, however, respond for stimuli paired with the *absence* of reinforcement; *those* should be conditioned punishers, and suppress responding. But instead they can they enhance responding, as first studied in detail by Neuringer and Chung (1967). *That* rises them to the level of anomalous, out of joint with the ubiquitous assumption that signals of the absence of a goal should be punishers. (See Myerson (1973) and Stubbs (1971) for additional data, and Marr (1979) for a review.)

The schedules that Neuringer and Chung (1967) used were: a) A variable-interval one-minute schedule (VI 1), where food or other reinforcer is delivered for the first response after the elapse of one of a random set of time intervals averaging 1 min; and b) A response-initiated Fixed Interval (FI) schedule, in which a response would start a 5-s timer; the first response after that 5-s elapsed was followed either by a 1 s chamber blackout (BO), or food. During BO, all illumination in the chamber was extinguished. Food would be delivered if the VI interval had elapsed. During food delivery, the chamber remained illuminated. The authors called this a *Percentage Reinforcement* (PR) schedule: About 85% of the PRs were followed by BO, the rest by food. Rather than punish responding, under the superimposed PR schedule the median response rate of the 3 pigeons *increased* from 46 to 114 responses per minute (r/m). When returned to the baseline schedule, rates reverted to 50 r/m. When the BO was increased to 7 s, rates increased to 134 r/m; for ¼ s BO, to 116; and for 0 s BO, back to 51 r/m. Blackout was necessary to elevate rates, and rate was an increasing function of the duration of the BO.

There followed four interesting control conditions. When food access took the place of BO, rates increased to 136 r/m, about same as for the 7 s BO. Measured in terms of response rate, BO presentation was as effective as food presentation. When non-contingent BOs occurred every 11 s, median rate fell to 62 r/m, essentially baseline level: The BOs in themselves did not elevate response rates. BOs after the 11th peck (a fixed-ratio, FR, schedule) after an interval from the VI schedule elapsed gave 73 r/m; an analogous PR schedule in which every 11th response produced either a BO or food gave 136 r/m. This last comparison is perhaps the most important: When food was delivered on exactly the same schedule as the blackouts, median rate was 136; when instead it came according to the VI, uncorrelated with the signal, response rates were halved. This was the case even though the last blackout in the FR control condition was typically closer in time to food than under the PR schedule. This effect has been replicated (Shull et al., 1975). The patterns of responding cued by the BOs must be the same as those that lead to food, or they lose their effectiveness in maintaining high response rates. Why should this be?

Davison and Baum (2006) and Reed and Hall (1989) offer interpretations that emphasize the "discriminative function of the stimulus": Signaling which response will likely produce food causes the effect. That is certainly part of the answer; but as certainly an incomplete part. Why should the last two conditions of the Neuringer and Chung study have given such different results? The animals had long ago learned that key-pecks are what it takes to produce food. There was no requirement to peck fast, or to show an FR-like break–and–run pattern of responses to get food. There was no requirement to slow down in the control conditions. Because BOs in the control conditions were almost always closer to food than under the PR schedules, they should have been be more powerful S[D]s, encouraging pecking. They were not. Finally, if the BOs had been long enough to elicit roosting, we might expect a conditioned suppression of rates involving BOs. Just the reverse was the case for the PR schedules.

Neuringer and Chung's first PR condition required approximately 700 reinforcers—more than 4000 PR elements—before rates had reached their stable maximum. That is long enough to make the response habitual. We know that it requires a reliable pattern of behavior before an outcome (as in the PR conditions, where food was always preceded by the elemental FI or FR) to make and keep a response habitual (Thrailkill et al., 2018). Goal-oriented outcome is sensitive to the value of the goal, habitual responding is not. The PR responding in this experiment was not sensitive to the value of the outcome, in that the BOs were S[−]s: predictors of non-reinforcement. That they did not punish is suggested by the increase in response rates with the duration of the BOs. When the regularity of behavior pattern and reinforcement was disrupted under the control conditions, high rates were no longer elicited by the BOs.

We must turn to BST and DPT to make sense of this anomaly. Pecking regularly yields a stimulus (BO) in the presence of which the same pattern of pecking would sometimes lead to the next module (consumption of grain). The BOs were both consequences of responding, and stimuli to release that pattern of responding. That the

consequence was negative 85% of the time was apparently of little consequence, as long as the integrity of the habit was maintained by the PR schedule. The animals had mastered a module. As such, responding will be goal-insensitive with the BO in effect, and goal-sensitive with no BO. Such insensitivity to negative consequences like extinction is central part of the next anomaly as well.

6.6. The paradoxes of intermittent reinforcement

Lloyd Humphreys first noted that continuously reinforced behavior extinguishes more quickly when reinforcement is withheld than does intermittently reinforced behavior (Humphreys, 1939). The effect was replicated and came to be known as *Humphreys' paradox* (Lattal et al., 2013), paradoxical because reinforcing every response should, it was thought, generate greater "response strength" as manifest in prolonged responding in extinction. This "partial reinforcement extinction effect" drove research for a generation, with many explanations tendered, such as *generalization decrement*–decreased frequency of reinforcement is more noticeable coming from continuous reinforcement than from intermittent reinforcement. Other conditions and measures show the reverse effect, as studied in behavioral momentum theory (Nevin, 1988). In addition, intermittent reinforcement on ratio schedules will often yield higher response rates than does continuous reinforcement, another problematic effect that foiled Skinner's development of his model *reflex reserve* (Killeen, 1988). As a final insult to theory, animals will often prefer an intermittent schedule of reinforcement to one with a fixed delay of equal mean. This finding scuttles models of optimal foraging in which the currency optimized is mean rate of reinforcement.

A sequence of experiments begun by Kendall (1974) initiated study of yet another anomaly related to intermittent schedules, now called *sub-optimal choice*. After training, pigeons, dogs, and in some cases rats and humans, will choose an option (Option 1, "S^D") giving a stimulus, say, 20% of the time, with that stimulus always signaling a preferred outcome (e.g., ability to eat), and give another stimulus signaling extinction (an eventual BO rather than food) the other 80% of the time. Option 2 ("no [differential] S^D") gives either of two stimuli with equal probability, both indicating a 50% probability of reinforcement. Option 2 will provide reinforcement half the time, whereas Option 1 will only provide it 20% of the time. After a number of sessions, most pigeons show indifference between the options, and eventually all come to show a reliable preference for the suboptimal option; Zentall (2016) calls the effect "paradoxical". Many hypotheses have been tendered, and tested in clever experiments (and largely rejected; see the reviews in McDevitt et al., 2016).

One obvious interpretation is that all that matters is the value (% reinforcement) of the best possible outcome of a wager, while the probability of obtaining it (and of obtaining instead a bad outcome, S– indicating no payoff, in its place) matters not at all. This is on the face of it what happens. In one experiment (Zentall and Stagner, 2011) pigeons strongly preferred a one-in-five chance of obtaining a signal for 10 pellets over a signal for a sure-thing 3 pellets. Such research provides a model for human gambling, wherein the quality of the best possible outcome (say, a million dollar payoff) controls betting, not the recognition that the probability of obtaining that outcome is less than a million-to-one (Zentall and Laude, 2013). Gamblers are essentially choosing between a negligible chance of a million dollars versus a sure win (saving) of the $2 that they do not spend on the ticket. What remains paradoxical is why those relative probabilities of a payoff are ignored, and why one very good S^D out of many bad ones on Option 1 should be so much more powerful than two pretty-good stimuli all of the time on Option 2. The current account provides a version of this (conditioned reinforcement) hypothesis, and an attempt to unravel this paradox of choice of intermittent schedules.

It is important to clear one's thinking of food consumption per se as reinforcing the choice of options. Think instead of the opportunity to engage in instrumental responses that reliably produce the opportunity

to eat as a key factor. Some data relevant to this claim were provided by Thrailkill and Bouton (2017a, 2017b), who showed that responding in the initial link of an heterogenous chain was more depressed when the responses in the second, terminal link were extinguished, than it was when the stimulus cuing that terminal link was extinguished. Whatever positive value the stimuli have accrued, it is as markers of the instrumental responses that they allow. The current rendering of BRT holds that it is the transition from choosing an option into the responses in the terminal link that motivates the choice response. Choosing Option 1 will on occasion deliver the terminal link whose responses pay off 100%, thus providing access to the activity most depressed below the level at which it would otherwise occur. If that outcome is made more probable (say, occurring 1/3 of the time rather than 1/5 of the time), there will be less response deprivation and preference will be less suboptimal (Zentall et al., 2019a); another mini-anomaly. This compounding of the paradox is not explained by simple conditioned reinforcement theory but is by BRT. Meanwhile, the signal for non-reinforcement is ignored (Fortes et al., 2016) as is the ensuing non-reward, just as it was in (Killeen et al., 2009). This differential salience of cues for reinforcement and extinction receives support in a model of trial-by-trial acquisition (Daniels and Sanabria, 2018).

Why do the differential rates of reinforcement on the two options not control choice? It is possible that those are overshadowed by the terminal link stimuli and responses. Pearce and Hall (1978) showed that brief stimuli paired with reinforcement could decrease the response rate even while demonstrating that they continued to function as conditioned reinforcers, prolonging responding in extinction. They attributed this effect to the "overshadowing of the response–reinforcer association by the formation of a stimulus–reinforcer association" p. 356). Williams (1999) showed that stimuli presented after a response (marking) can speed acquisition of conditioning with delayed reinforcement, whereas ones before reinforcement can retard or eliminate acquisition. He also showed that the ability to engage in an alternate, pretrained response between a target response and reinforcement blocks acquisition of the target response (Williams, 1975), even though such responses play a central role in maintaining habitual responses. These experiments suggest that blocking and overshadowing effects may suppress control by differential reinforcement in the sub-optimal choice paradigm.

The S^D for 100% reinforcement has incentive motivational properties because it leads to depressed contingent responding. Permitting the animal to switch out of the dead links (signaled by $S^−$) might generate the attention to them that would give them more weight in the initial choice. Permitting the animal to make instrumental responses for alternate low probability rewards in the otherwise unreinforced component might also decrease preference for that side (another paradox, recently demonstrated by Fortes et al., 2018).

There are interesting similarities between sub-optimal choice and quasi-reinforcement. In quasi-reinforcement, the better (the longer) the BO stimulus, the faster subjects responded in the component leading to it. They do so even though most of the time it leads to another $S^−$. In sub-optimal choice, the better the stimulus (viz., one predicting reinforcement with probability 1.0), the more likely subjects are to choose it in the component leading to it. They do so even though most of the time it leads to another $S^−$. In quasi-reinforcement, reduce the BO to 0 (no $S^−$), and there is no rate enhancement. In sub-optimal choice, the component with no differential stimuli is dispreferred. If the responding for the 100% component has become habitual, and that for the ambiguous link not, shifts in motivation should shift preference between them. If habitual responding for the 100% stimulus plays a role, changing the terminal links to variable-ratio schedules should decrease sub-optimal choice.

Reinforcing the importance of the incentive salience perspective of Section 6.3.1, the Zentall group has concluded that "in order to obtain suboptimal choice behavior, incentive salience alongside strong stimulus-reward predictive utility may be necessary; thus, maladaptive

decision-making can be driven more by the value attributed to stimuli imbued with incentive salience that reliably predict a reward rather than the reward itself" (Chow et al., 2017, p. 244). Timberlake (e.g., 1984; and 1994, Fig. 2) might paraphrase this as "instrumental behavior [choosing the signaled link] leading to stimuli that reliably allow contingent actions that are depressed below their set-point in that context will be reinforced by that transition". The apparent differences between pigeons and rats may be due to the ability of the stimuli used to effectively partition the depressed response, in its own context, from the context at large (Zentall, et al., 2019b).

7. Implications for non-anomalous behavior science

Premack, and Timberlake and Allison (1974), have argued that primary reinforcers comprise more probable, more depressed responses. Thrailkill and Bouton (2017a, 2017b) have shown that responses in links of a heterogenous behavior chain are the events that govern behavior in the earlier link. Why should conditioned reinforcers not also be actions? The essential heart of BRT was that reinforcers are contingent *responses* that are depressed below their set–point. Discriminative stimuli may be conditioned to the actions that they allow, if those are a step down the imminence gradient, rather than to an ultimate consummatory response. They may motivate those actions, and they may be approached, not as goods in themselves, but as cues that enable the next set of responses. Behaviors elicited at the various modules can clearly have the appearance of reinforcers (e.g., Cetinkaya and Domjan, 2006). In an important study, Schuster (1969) showed that conditioned reinforcers added to a patch (one terminal link of a concurrent-chain schedule) did not increase preference for that patch. Like the control conditions in the quasi-reinforcement experiment, those stimuli did not govern any distinctive behavior; if it is the behavior the stimuli release that constitute conditioned reinforcement, his results are what one would expect. Conversely, Grice (1948, p. 1) placed blocks in one length of alley that provided delayed reinforcement, and found that "when animals were forced to make characteristically different motor responses to the black and white stimuli, they learned at a significantly faster rate than animals which received equal delay, but made no such characteristically different motor adjustments" (cf. the differential outcome effect; Peterson and Trapold, 1980; Urcuioli, 2005). Stimuli that instigate different actions are stronger discriminative stimuli (Chow et al., 2017); and the responses that those stimuli instigate should be stronger reinforcers.

In his review of the recent literature on conditioned reinforcement Shahan (2010; also see Shahan, 2017) agreed with Davison and Baum's observation: "The most general principle, rather than a strengthening and weakening by consequences, may be that whatever events predict phylogenetically important … events, such as food…, will guide behavior into activities that produce fitness-enhancing events" (Davison et al., 2006, p. 281)[3] . Perhaps it is those activities themselves that reinforce the prior. Shahan concluded that conditioned reinforcers act more like discriminative stimuli than like reinforcers, as did Bindra (Bindra, 1978; Lajoie and Bindra, 1976), whose "motivational states" are Timberlake's *modes,* and whose "act-assemblies" are Timberlake's

modules. That it is the behavior occasioned by stimuli that empower them is suggested by the work of Thrailkill and Shahan (2014), who showed that a novel stimulus (CS2) occurring after a *trace* CS1 (one which appears and quickly disappears before onset of the US) was an effective conditioned reinforcer for new responses; whereas a novel stimulus paired with a *delay* CS1 (one which appears and stays on until onset of the US) were much less effective. The trace CS2 occasioned various (superstitious, intrusive, "derived") activities; stimuli paired with those behaviors were incentive motivators. The delay CS2 occasioned hopper oriented, goal-tracking consummatory responses—no more transitions necessary, no ability to condition instrumental responses when the animal was already at the threshold of consummation. This concept of the depressed actions as the reinforcers is a core principle of BRT. Vandbakk and associates (Vandbakk et al., 2018) showed that, ceteris paribus, discriminative stimuli [and the behavior they directed] were more effective conditioned reinforcers than Pavlovian CSs. Movement down the gradient to *actions* at lower levels reinforces actions that get animals there (Table 1).

8. Discussion

Table 1 summarizes the above anomalies, giving the background expectations which made them anomalous, the originating reference, more recent reviews, and the tentative resolution offered by Timberlake and associates' BRT and BST and Dickinson and associates' DPT. The thumbnails in the table do scant justice to the literature, or to the potential for resolution offered by the combined theories. Also lacking from the table is a column listing crucial experiments that will test the proposed resolutions.

Table 1 helps us to recognize the common elements among these phenomena. The weltanschauung that makes them anomalous in general is the assumption of rationality, economic sensibility or optimality (Bianchi and De Marchi, 2016); in all cases the animals are engaging in responses that are unnecessary to achieve the nominal goal, and in many cases interfere with the achievement of that goal. Because the deprivation schedule and primary reinforcer drive these phenomena, we assume that behavior will come to be efficiently organized, within the constraints of the capacities of the organism, to achieve those goals expeditiously. Instead, those reinforcers strike (or create) a tuned organism, which resonates with modules of activity that we have not shaped, and are not what we think we would do in similar situations.

In many cases instrumental activities take on a life of their own, becoming intrinsically reinforcing. This is probably at the point at which they become habitual; the action is engaged because it is reinforcing in that context, and not because it leads to the goal that originally motivated it. Indeed, it may be sufficiently salient that it blocks control by the ultimate consummatory response (Cronin, 1980), as often the case for misbehavior. This is what I call "canalization of the action". In some cases, the canalized actions have been shown to share features with addictive responses.

Table 1 hardly exhausts the anomalies in our field. Animals will respond to deliver painful electric shock to themselves when inaction would avoid the shocks (McKearney, 1968; Byrd, 1969). Animals will choose stimuli associated with food more often if access to those stimuli had required more responses, or more delay (summarized in Zentall, 2013. Access to those stimuli and the consummatory responses they permit after the delay was depressed below its baseline level in that context, suggesting its strength may have been enhanced through a BRT mechanism.). Some discriminated performances that should be very easy are very difficult (e.g., Reid et al., 2017; Urcuioli, 2006) and others that should be impossible are manifest (Pilley and Reid, 2011; McGaugh and LePort, 2014). In what its authors dubbed the *paradoxical incentive effect*, a study showed that increasing the magnitude of reinforcement from 1 to 2 pellets on small to moderate variable ratio schedules *decreased* response rates of rats; increasing further to 3 further decreased response rates (Bizo et al., 2001), even with possible

---

[3] This is a point with which Bill may disagree. "Skinner's notion of selection by consequences) presumes a single critical event, the presentation of a reinforcing consequence. …from a systems view this is not a simple causal manipulation. Presentation of the reinforcer is an act of the experimenter that constrains the entire functioning system, generating multiple causal sequences and typically producing multiple effects" (Timberlake, 1993, p. 107). I appreciate RIB pointing out this position and its inconsistency with the Skinnerian roots of Davison, Baum, and Shahan. Bill's observation, as uttered here, is, of course, consistent with the positions of Jenkins, Thrailkill and Shahan and many others cited in this paper, as well as for the authors of the citation diverting you to this footnote, as they have advocated molar behaviorisms, and sensitivity to, and guidance of, longer term correlations (e.g., Davison, 2017).

Table 1
Anomalies in behavioral science.

Anomaly	Expectation	Observation	Reference	Review	Resolution (?)
Misbehavior	Law of Effect is dominant	Species-specific responses intrude	Breland and Breland (1961)	Domjan (1983, 2008)	Variations in topography appropriate to released niche-specific actions
Contra-freeloading	Efficiency is optimized	Unnecessary work for food	Jensen (1963)	Osborne (1977); Inglis et al. (1997)	Canalization of actions
Activity-Based Anorexia	Primary needs take priority	Unnecessary activity to point of starvation	Spear and Hill (1962)	Epling and Pierce (1996a, 1996b)	Canalization of actions
Sign-tracking	Focus on goal not sign	Conditioning of motor actions in Pavlovian paradigm	Brown and Jenkins (1968)	Hearst and Jenkins (1974); Locurto et al. (1981)	Incentive motivation of search/appetitive modules
Quasi-reinforcement	S^- should inhibit, not excite, responding	Stimuli on same schedule as eating increase response rate	Neuringer and Chung (1967)	Marr (1979)	Canalization of actions
Sub-optimal choice	Economic rationality	S^+ overwhelms optimality of choice; S^- does not punish	Kendall (1974)	McDevitt et al. (2016); Zentall (2016)	Choice reinforced by releasing most depressed responding in S^+

satiation controlled for. There must be many anomalies in our field that remain unpublished; some perhaps unpublishable; and others just *harrumphed* at and then ignored. How about the transient increase in response rates at the start of extinction after maintenance on moderate to high rate reinforcement schedules (see, e.g., Figs. 4–6 of Killeen and Nevin, 2018), a bump that may conveniently be squashed by a logarithmic transformation of the *y*-axis. Anomalous, or merely strange? How about something that you have seen?

9. Envoi

The branching lattice-hierarchy appearing in many of Timberlake and student's papers, reproduced as Fig. 1 in Silva et al. (2019) and available elsewhere (google "W Timberlake" systems), epitomizes Bill's BST. The modes tune perception. Progress through the modules is reinforcing, increasing the probability of actions that allow it. It does this because in the context that cues or releases that module, the next module's actions are depressed below what they would be—until the instrumental responses are effective—and through selection they may come to be more quickly effective, to allow that reinforcing transit. BRT complements BST. The architecture is so obviously reasonable that it is easy to nod and pass on; yet harder to test. Testing it is what Timberlake and his students have creatively succeeded in doing over the years (e.g., Silva and Timberlake, 2005). Integrated with modern learning theory, such as Dickinson's dual-process theory, it offers new hypotheses about the reasons that drive animals to engage in behavior which, to our more rational minds attuned to cost/benefit ratios, seem anomalous. Together they can "rationalize"—provide satisfactory reasons for—behavioral excesses such as contra-freeloading, activity anorexia, and sign-tracking; and insufficiencies such as suboptimal choice, that seem otherwise irrational. We have Timberlake and students to thank for importing biological realism into the background psychological idealism that made species-specific behaviors seem anomalous. "The strength of the approach lies in the type of experiments it suggests, its ability to predict counter-intuitive outcomes, and its focus upon the evolved nature of learning" (Timberlake, 1983a; Timberlake, 1983b). Just so.

Author note

For the *Behavioural Processes* issue honoring William Timberlake.

Acknowledgements

Appendix A. Supplementary data

A dynamic visualization of the evolution of the transition matrix of MacNulty et al. (2007) describing wolf packs attacking elk herds may be fund at this doi:https://doi.org/10.1016/j.beproc.2019.103894. The states are A: Search; B: Approach; C: Watch; D; Attack Group; E: Attack Individual; F: Capture. The animation does not convey the time spent in each state. The visualization was made possible by the fascinating *SetOsa* blog of Victor Powell (http://setosa.io/blog/2014/07/26/markov-chains/).

References

Adams, C.D., Dickinson, A., 1981. Instrumental responding following reinforcer devaluation. Q. J. Exp. Psychol. 33, 109–121. https://doi.org/10.1080/14640748108400816.

Akins, C.K., Domjan, M., Gutierrez, G., 1994. Topography of sexually conditioned behavior in male Japanese quail (Coturnix japonica) depends on the CS-US interval. J. Exp. Psychol. Anim. Behav. Process. 20, 199–209. https://doi.org/10.1037/0097-

7403.20.2.199.

Alcaro, A., Panksepp, J., 2011. The SEEKING mind: primal neuro-affective substrates for appetitive incentive states and their pathological dynamics in addictions and depression. Neurosci. Biobehav. Rev. 35, 1805–1820. https://doi.org/10.1016/j. neubiorev.2011.03.002.

Alferink, L.A., Crossman, E.K., Cheney, C.D., 1973. Control of responding by a conditioned reinforcer in the presence of free food. Anim. Learn. Behav. 1, 38–40. https:// doi.org/10.3758/BF03198996.

Anselme, P., 2016. Motivational control of sign-tracking behaviour: a theoretical framework. Neurosci. Biobehav. Rev. 65, 1–20. https://doi.org/10.1016/j.neubiorev.2016. 03.014.

Badia, P., Ryan, K., Harsh, J., 1981. Choosing schedules of signaled appetitive events over schedules of unsignaled ones. J. Exp. Anal. Behav. 35, 187–195. https://doi.org/10. 1901/jeab.1981.35-187.

Baerends, G.P., 1976. The functional organization of behaviour. Anim. Behav. 24, 726–738. https://doi.org/10.1016/S0003-3472(76)80002-4.

Balleine, B.W., Garner, C., Gonzalez, F., Dickinson, A., 1995. Motivational control of heterogeneous instrumental chains. J. Exp. Psychol. Anim. Behav. Process. 21, 203–217. https://doi.org/10.1037/0097-7403.21.3.203.

Balsam, P.D., Graf, J.S., Silver, R., 1992. Operant and Pavlovian contributions to the ontogeny of pecking in ring doves. Dev. Psychobiol. 25, 389–410. https://doi.org/10. 1002/dev.420250602.

Balsam, P.D., Silver, R., 1994. Behavioral change as a result of experience: toward principles of learning and development. In: Hogan, J.A., Bolhuis, J.J. (Eds.), Causal Mechanisms of Behavioural Development. Cambridge University Press, Cambridge.

Baum, W.M., 1974. Chained concurrent schedules: reinforcement as situation transition. J. Exp. Anal. Behav. 22, 91–101. https://doi.org/10.1901/jeab.1974.22-91.

Baum, W.M., 2012. Rethinking reinforcement: allocation, induction, and contingency. J. Exp. Anal. Behav. 97, 101–124. https://doi.org/10.1901/jeab.2012.97-101.

Belke, T.W., Pierce, W.D., Welsh, T.M., 2018. Automatic reinforcement from operant wheel-running undermines temporal control by fixed-interval schedules of reinforcement. Behav. Processes 157, 91–101. https://doi.org/10.1016/j.beproc.2018. 09.003.

Bennett, A., Elman, C., 2006. Complex causal relations and case study methods: the example of path dependence. Political Anal. 14, 250–267. https://doi.org/10.1093/ pan/mpj020.

Berridge, K.C., 1996. Food reward: brain substrates of wanting and liking. Neurosci. Biobehav. Rev. 20, 1–25. https://doi.org/10.1016/0149-7634(95)00033-B.

Berridge, K.C., Fentress, J.C., Parr, H., 1987. Natural syntax rules control action sequence of rats. Behav. Brain Res. 23, 59–68.

Berridge, K.C., Robinson, T.E., 2003. Parsing reward. Trends Neurosci. 26, 507–513. https://doi.org/10.1016/S0166-2236(03)00233-9.

Bianchi, M., De Marchi, N., 2016. Economizing Mind, 1870-2015: When Economics and Psychology Met… Or Didn't. Duke University Press, Durham.

Bindra, D., 1978. How adaptive behavior is produced: a perceptual-motivational alternative to response-reinforcement. Behav. Brain Sci. 1, 41–91. https://doi.org/10. 1017/S0140525X00059380.

Bizo, L.A., Kettle, L.C., Killeen, P.R., 2001. Rats don't always respond faster for more food: the paradoxical incentive effect. Anim. Learn. Behav. 29, 66–78. https://doi.org/10. 3758/BF03192816.

Boakes, R.A., 1977. Performance on learning to associate a stimulus with positive reinforcement. In: David, H., Hurwitz, H.M.B. (Eds.), Operant-Pavlovian Interactions. Erlbaum, Hillsdale, NJ.

Boakes, R.A., 1997. Wheels, clocks, and anorexia in the rat. In: Bouton, M.E., Fanselow, M.S. (Eds.), Learning, Motivation, and Cognition: The Functional Behaviorism of Robert C. Bolles. American Psychological Association, Washington, DC.

Boakes, R.A., 2007. Self-starvation in the rat: running versus eating. Span. J. Psychol. 10, 251–257. https://doi.org/10.1017/S113874000000651X.

Boakes, R.A., Mills, K.J., Single, J.P., 1999. Sex differences in the relationship between activity and weight loss in the rat. Behav. Neurosci. 113, 1080–1089. https://doi.org/ 10.1037/0735-7044.113.5.1080.

Boakes, R.A., Poli, M., Lockwood, M.J., Goodall, G., 1978. A study of misbehavior: token reinforcement in the rat. J. Exp. Anal. Behav. 29, 115–134. https://doi.org/10.1901/ jeab.1978.29-115.

Bonardi, C., Robinson, J., Jennings, D., 2017. Can existing associative principles explain occasion setting? Some old ideas and some new data. Behav. Processes 137, 5–18. https://doi.org/10.1016/j.beproc.2016.07.007.

Bowers, R.I., 2017. Behavior systems. Encycl. Anim. Cogn. Behav. 1–8. https://doi.org/ 10.1007/978-3-319-47829-6_1232-1.

Bowers, R.I., 2018. A common heritage of behaviour systems. Behaviour 155, 415–442. https://doi.org/10.1163/1568539X-00003497.

Brackney, R.J., Cheung, T.H.C., Sanabria, F., 2017. A bout analysis of operant response disruption. Behav. Processes 141, 42–49. https://doi.org/10.1016/j.beproc.2017.04. 008.

Breland, K., Breland, M., 1961. The misbehavior of organisms. Am. Psychol. 16, 681–684. https://doi.org/10.1037/h0040090.

Breland, K., Breland, M., 1966. Animal Behavior.

Brown, P.L., Jenkins, H.M., 1968. Auto-shaping of the pigeon's key-peck. J. Exp. Anal. Behav. 11, 1–8. https://doi.org/10.1901/jeab.1968.11-1.

Burghardt, G.M., 2005. The Genesis of Animal Play: Testing the Limits. Mit Press.

Burghardt, G.M., 2019. A place for emotions in behavior systems research. Behav. Processes.

Burghardt, G.M., Bowers, R.I., 2017. From Instinct to Behavior Systems: an Integrated Approach to Ethological Psychology, APA Handbook of Comparative Psychology: Basic Concepts, Methods, Neural Substrate, and Behavior. American Psychological Association, Washington, DC, US.

Burghardt, G.M., Burkhardt Jr, R.W., 2018. Wallace Craig's Appetites and aversions as constituents of instincts: a centennial appreciation. J. Comp. Psychol. 132, 361. https://doi.org/10.1037/com0000155.

Burgos, J.E., Killeen, P.R., 2018. Suing for peace in the war against mentalism. Perspect. Behav. Sci. https://doi.org/10.1007/s40614-018-0169-2.

Buzsáki, G., 1982. The "where is it?" reflex: autoshaping the orienting response. J. Exp. Anal. Behav. 37, 461–484. https://doi.org/10.1901/jeab.1982.37-461.

Byrd, L.D., 1969. Responding in the cat maintained under response-independent electric shock and response-produced electric shock. J. Exp. Anal. Behav. 12, 1–10. https:// doi.org/10.1901/jeab.1969.12-1.

Cabrera, F., Sanabria, F., Jimenez, A.A., Covarrubias, P., 2013. An affordance analysis of unconditioned lever pressing in rats and hamsters. Behav. Processes 92, 36–46. https://doi.org/10.1016/j.beproc.2012.10.003.

Cabrera, F., Sanabria, F., Shelley, D., Killeen, P.R., 2009. The "lunching" effect: pigeons track motion towards food more than motion away from it. Behav. Processes 82, 229–235. https://doi.org/10.1016/j.beproc.2009.06.010.

Carrera, O., Fraga, A., Pellon, R., Gutierrez, E., 2014. Rodent model of activity-based anorexia. Curr. Protoc. Neurosci. 67 (9.47), 1–11. https://doi.org/10.1002/ 0471142301.ns0947s67.

Cetinkaya, H., Domjan, M., 2006. Sexual fetishism in a quail (Coturnix japonica) model system: test of reproductive success. J. Comp. Psychol. 120, 427–432. https://doi. org/10.1037/0735-7036.120.4.427.

Chow, J.J., Smith, A.P., Wilson, A.G., Zentall, T.R., Beckmann, J.S., 2017. Suboptimal choice in rats: incentive salience attribution promotes maladaptive decision-making. Behav. Brain Res. 320, 244–254. https://doi.org/10.1016/j.bbr.2016.12.013.

Colwill, R.M., Rescorla, R.A., 1988. The role of response reinforcer associations increases throughout extended instrumental training. Anim. Learn. Behav. 16, 105–111.

Corbit, L.H., Balleine, B.W., 2003. Instrumental and Pavlovian incentive processes have dissociable effects on components of a heterogeneous instrumental chain. J. Exp. Psychol. Anim. Behav. Process. 29, 99–106. https://doi.org/10.1037/0097-7403.29. 2.99.

Cordony, M.B.L., Chow, J.Y.L., Boakes, R.A., 2018. Motivation to run measured by progressive ratio tests: failure to support the addiction hypothesis for rats. Learn. Behav. 1–10. https://doi.org/10.3758/s13420-018-0348-8.

Craig, W., 1918. Appetites and Aversions as Constituents of Instincts 34. pp. 91–107.

Cronin, P.B., 1980. Reinstatement of postresponse stimuli prior to reward in delayed-reward discrimination learning by pigeons. Anim. Learn. Behav. 8, 352–358. https:// doi.org/10.3758/BF03199616.

Daniels, C.W., Sanabria, F., 2018. An associability decay model of paradoxical choice. J. Exp. Psychol. Anim. Learn. Cogn. 44, 258–271. https://doi.org/10.1037/ xan0000179.

Davey, G.C.L., 1989. Ecological Learning Theory. Routledge, London.

Davison, M., 2017. Killeen and Jacobs are not wrong. Behav. Anal. 40, 57–64. https://doi. org/10.1007/s40614-017-0118-5.

Davison, M., Baum, W.M., 2006. Do conditional reinforcers count? J. Exp. Anal. Behav. 86, 269–283. https://doi.org/10.1901/jeab.2006.56-05.

Dayan, P., Niv, Y., 2008. Reinforcement learning: the good, the bad and the ugly. Curr. Opin. Neurobiol. 18, 185–196. https://doi.org/10.1016/j.conb.2008.08.003.

Derman, R.C., Schneider, K., Juarez, S., Delamater, A.R., 2018. Sign-tracking is an expectancy-mediated behavior that relies on prediction error mechanisms. Learn. Mem. 25, 550–563. https://doi.org/10.1101/lm.047365.118.

Deutsch, J.A., 1960. The Structural Basis of Behavior. Univ. Chicago Press, Oxford, England.

Dezfouli, A., Lingawi, N.W., Balleine, B.W., 2014. Habits as action sequences: hierarchical action control and changes in outcome value. Philos. Trans. R. Soc. Lond. B Biol. Sci. 369. https://doi.org/10.1098/rstb.2013.0482.

Dickinson, A., 1985. Actions and habits: the development of behavioural autonomy. Philos. Trans. R. Soc. Lond. B Biol. Sci. 308, 67–78. https://doi.org/10.1098/rstb. 1985.0010.

Dickinson, A., 2016. Instrumental Conditioning Revisited: Updating Dual-Process Theory. Associative Learning and Cognition. Homage to Professor NJ Mackintosh. In Memoriam (1935-2015) 51. pp. 177–196.

Dickinson, A., Balleine, B.W., 1994. Motivational control of goal-directed action. Anim. Learn. Behav. 22, 1–18. https://doi.org/10.3758/BF03199951.

Dickinson, A., Balleine, B.W., Watt, A., Gonzalez, F., Boakes, R.A., 1995. Motivational control after extended instrumental training. Anim. Learn. Behav. 23, 197–206. https://doi.org/10.3758/BF03199935.

Domjan, M., 1983. Biological constraints on instrumental and classical conditioning: implications for general process theory. In: Bower, G.H. (Ed.), The Psychology of Learning and Motivation: Advances in Research and Theory. Academic Press, New York.

Domjan, M., 1994. Formulation of a behavior system for sexual conditioning. Psychon. Bull. Rev. 1, 421–428. https://doi.org/10.3758/BF03210946.

Domjan, M., 2005. Pavlovian conditioning: a functional perspective. Annu. Rev. Psychol. 56, 179–206. https://doi.org/10.1146/annurev.psych.55.090902.141409.

Domjan, M., 2008. Adaptive specializations and generality of the laws of classical and instrumental conditioning. In: Menzel, R. (Ed.), Learning Theory and Behavior. Elsevier, Oxford.

Domjan, M., Galef, B.G., 1983. Biological constraints on instrumental and classical conditioning: retrospect and prospect. Anim. Learn. Behav. 11, 151–161. https://doi. org/10.3758/BF03199643.

Donahoe, J.W., 1977. Some implications of a relational principle of reinforcement. J. Exp. Anal. Behav. 27, 341–350. https://doi.org/10.1901/jeab.1977.27-341.

Downing, K., Neuringer, A.J., 1976. Autoshaping as a function of prior food presentations. J. Exp. Anal. Behav. 26, 463–469. https://doi.org/10.1901/jeab.1976.26-463.

Dwyer, D.M., Boakes, R.A., 1997. Activity-based anorexia in rats as failure to adapt to a

feeding schedule. Behav. Neurosci. 111, 195–205. https://doi.org/10.1037/0735-7044.111.1.195.

Edwards, T.L., Lotfizadeh, A.D., Poling, A., 2019. Motivating operations and stimulus control. J. Exp. Anal. Behav. https://doi.org/10.1002/jeab.516.

Eisenberger, R., Karpman, M., Trattner, J., 1967. What is the necessary and sufficient condition for reinforcement in the contingency situation? J. Exp. Psychol. 74, 342–350. https://doi.org/10.1037/h0024719.

Epling, W.F., Pierce, W.D., 1996a. Activity Anorexia: Theory, Research, and Treatment. Lawrence Erlbaum Associates, Mahwah, NJ.

Epling, W.F., Pierce, W.P., 1996b. An Overview of Activity Anorexia, Activity Anorexia: Theory, Research and Treatment. Lawrence Erlbaum Associates, Mahwah, NJ.

Esmoris-Arranz, F.J., Pardo-Vazquez, J.L., Vazquez-Garcia, G.A., 2003. Differential effects of forward or simultaneous conditioned stimulus-unconditioned stimulus intervals on the defensive behavior system of the Norway rat (Rattus norvegicus). J. Exp. Psychol. Anim. Behav. Process. 29, 334–340. https://doi.org/10.1037/0097-7403.29.4.334.

Fanselow, M.S., 1994. Neural organization of the defensive behavior system responsible for fear. Psychon. Bull. Rev. 1, 429–438. https://doi.org/10.3758/BF03210947.

Fanselow, M.S., Lester, L.S., 1988. A functional behavioristic approach to aversively motivated behavior: predatory imminence as a determinant of the topography of defensive behavior. In: Bolles, R.C., Beecher, M.D. (Eds.), Evolution and Learning. Lawrence Erlbaum and Associates, Inc., Hillsdale, NJ.

Fentress, J.C., 1983. Ethological models of hierarchy and patterning of species-specific behavior. In: Teitelbaum, P., Satinoff, E. (Eds.), Handbook of Behavioral Neurobiology. Plenum Press, New York.

Fenz, W.D., Epstein, S., 1967. Gradients of physiological arousal in parachutists as a function of an approaching jump. Psychosom. Med. 29, 33–51.

Fortes, I., Pinto, C., Machado, A., Vasconcelos, M., 2018. The paradoxical effect of low reward probabilities in suboptimal choice. J. Exp. Psychol. Anim. Learn. Cogn. 44, 180–193. https://doi.org/10.1037/xan0000165.

Fortes, I., Vasconcelos, M., Machado, A., 2016. Testing the boundaries of "paradoxical" predictions: pigeons do disregard bad news. J. Exp. Psychol. Anim. Learn. Cogn. 42, 336–346. https://doi.org/10.1037/xan0000114.

Fraser, K.M., Holland, P.C., 2019. Occasion setting. Behav. Neurosci. 133, 145–175. https://doi.org/10.1016/j.beproc.2017.02.009.

Freeman, K.B., Riley, A.L., 2009. The origins of conditioned taste aversion learning: a historical analysis. In: Riley, S., Schachtman, T.R. (Eds.), Conditioned Taste Aversion: Behavioral and Neural Processes. Oxford Univ., Oxford, UK.

Fujikawa, K., Fujikawa, K., Suzuki, H., 2004. Path Integrals and Quantum Anomalies. Oxford University Press on Demand.

Gallistel, C.R., 1981. Précis of Gallistel's The organization of action: A new synthesis. Behav. Brain Sci. 4, 609–619. https://doi.org/10.1017/S0140525X00000480.

Garcia, J., 1981. Tilting at the paper mills of academe. Am. Psychol. 36, 149–158. https://doi.org/10.1037/0003-066X.36.2.149.

Gawley, D.J., Timberlake, W., Lucas, G.A., 1987. System-specific differences in behavior regulation: overrunning and underdrinking in molar nondepriving schedules. J. Exp. Psychol. Anim. Behav. Process. 13, 354–365. https://doi.org/10.1037/0097-7403.13.4.354.

Gibbon, J., Baldock, M.D., Locurto, C.M., Gold, L., Terrace, H.S., 1977. Trial and intertrial durations in autoshaping. J. Exp. Psychol. Anim. Behav. Process. 3, 264–284. https://doi.org/10.1037/0097-7403.3.3.264.

Gibbon, J., Balsam, P., 1981. Spreading association in time. In: Locurto, C.M., Terrace, H.S., Gibbon, J. (Eds.), Autoshaping and Conditioning Theory. Academic Press, New York.

Graybiel, A.M., 2008. Habits, rituals, and the evaluative brain. Annu. Rev. Neurosci. 31, 359–387. https://doi.org/10.1146/annurev.neuro.29.051605.112851.

Grice, G.R., 1948. The relation of secondary reinforcement to delayed reward in visual discrimination learning. J. Exp. Psychol. 38, 1–16. https://doi.org/10.1037/h0061016.

Gygax, L., 2017. Wanting, liking and welfare: the role of affective states in proximate control of behaviour in vertebrates. Ethology 123, 689–704. https://doi.org/10.1111/eth.12655.

Hanson, S.J., Timberlake, W., 1983. Regulation during Challenge: a general model of learned performance under schedule constraint. Psychol. Rev. 90, 261–282. https://doi.org/10.1037/0033-295X.90.3.261.

Hearst, E., Jenkins, H.M., 1974. Sign-tracking: the Stimulus-reinforcer Relation and Directed Action. The Psychonomic Society, Austin, TX.

Hebb, D.O., Williams, K., 1946. A method of rating animal intelligence. J. Gen. Psychol. 34, 59–65. https://doi.org/10.1080/00221309.1946.10544520.

Hebebrand, J., Albayrak, O., Adan, R., Antel, J., Dieguez, C., de Jong, J., Leng, G., Menzies, J., Mercer, J.G., Murphy, M., van der Plasse, G., Dickson, S.L., 2014. "Eating addiction", rather than "food addiction", better captures addictive-like eating behavior. Neurosci. Biobehav. Rev. 47, 295–306. https://doi.org/10.1016/j.neubiorev.2014.08.016.

Herrnstein, R.J., 1977. The evolution of behaviorism. Am. Psychol. 32, 593–603.

Heth, C.D., Warren, A., 1978. Response deprivation and response satiation as determinants of instrumental performance: some data and theory. Anim. Learn. Behav. 6, 294–300. https://doi.org/10.3758/BF03209617.

Hinde, R.A., 1982. Ethology: Its Nature and Relation with Other Sciences. Oxford University Press, New York.

Hogan, J.A., 1988. Cause and function in the development of behavior systems. In: Blass, E.M. (Ed.), Developmental Psychobiology and Behavioral Ecology. Plenum Press, New York.

Hogan, J.A., 1999. Structure and development of behavior systems. Dev. Anim. Behav. A Reader 300.

Hogan, J.A., 2017. The Study of Behavior: Organization, Methods, and Principles. Cambridge University Press.

Hogan, J.A., Bolhuis, J.J., 2009. Causal Mechanisms of Behavioural Development. Cambridge University Press, Cambridge, England.

Holland, P.C., 1977. Conditioned stimulus as a determinant of the form of the Pavlovian conditioned response. J. Exp. Psychol. Anim. Behav. Process. 3, 77–104. https://doi.org/10.1037/0097-7403.3.1.77.

Holland, P.C., 1980. CS–US interval as a determinant of the form of Pavlovian appetitive conditioned responses. J. Exp. Psychol. Anim. Behav. Process. 6, 155–174. https://doi.org/10.1037/0097-7403.6.2.155.

Holland, P.C., 1992. Occasion setting in Pavlovian conditioning. In: Medin, D.L. (Ed.), The Psychology of Learning and Motivation.

Horney, J., Fantino, E., 1984. Choice for conditioned reinforcers in the signaled absence of primary reinforcement. J. Exp. Anal. Behav. 41, 193–201. https://doi.org/10.1901/jeab.1984.41-193.

Humphreys, L.G., 1939. The effect of random alternation of reinforcement on the acquisition and extinction of conditioned eyelid reactions. J. Exp. Psychol. 25, 141–156. https://doi.org/10.1037/h0058138.

Inglis, I.R., Forkman, B., Lazarus, J., 1997. Free food or earned food? A review and fuzzy model of contrafreeloading. Anim. Behav. 53, 1171–1191. https://doi.org/10.1006/anbe.1996.0320.

Jacobs, K.W., Morford, Z.H., King, J.E., 2019. Disequilibrium in behavior analysis: a disequilibrium theory redux. Behav. Processes. https://doi.org/10.1016/j.beproc.2019.02.006.

Jacobs, K.W., Morford, Z.H., King, J.E., Hayes, L.J., 2017. Predicting the effects of interventions: a tutorial on the disequilibrium model. Behav. Anal. Pract. https://doi.org/10.1007/s40617-017-0176-x.

James, W., 1890/1983. Principles of Psychology. Harvard University Press, Cambridge, MA.

Jenkins, H.M., Barnes, R.A., Barrera, F.J., 1981. Why autoshaping depends on trial spacing. In: Locurto, C.M., Terrace, H.S., Gibbon, J. (Eds.), Autoshaping and Conditioning Theory. Academic Press, New York.

Jenkins, H.M., Barrera, F.J., Ireland, C., Woodside, B., 1978. Signal-centered action patterns of dogs in appetitive classical conditioning. Learn. Motiv. 9, 272–296. https://doi.org/10.1016/0023-9690(78)90010-3.

Jensen, G.D., 1963. Preference for bar pressing over "freeloading" as a function of number of rewarded presses. J. Exp. Psychol. 65, 451–454. https://doi.org/10.1037/h0049174.

Kagan, J., Berkun, M., 1954. The reward value of running activity. J. Comp. Physiol. Psychol. 47, 108. https://doi.org/10.1037/h0058877.

Kamil, A.C., 1987. A synthetic approach to the study of animal intelligence. Dienstbier, R.A., Leger, D.W. (Eds.), NEbraska Symposium on Motivation 1987 Vol. 35.

Kanarek, R.B., D'Anci, K.E., Jurdak, N., Mathes, W.F., 2009. Running and addiction: precipitated withdrawal in a rat model of activity-based anorexia. Behav. Neurosci. 123, 905–912. https://doi.org/10.1037/a0015896.

Kendall, S.B., 1974. Preference for intermittent reinforcement. J. Exp. Anal. Behav. 21, 463–473. https://doi.org/10.1901/jeab.1974.21-463.

Killeen, P.R., 1975. On the temporal control of behavior. Psychol. Rev. 82, 89–115. https://doi.org/10.1037/h0076820.

Killeen, P.R., 1978. Superstition: a matter of bias, not detectability. Science 199, 88–90. https://doi.org/10.1126/science.199.4324.88.

Killeen, P.R., 1988. The reflex reserve. J. Exp. Anal. Behav. 50, 319–331. https://doi.org/10.1901/jeab.1988.50-319.

Killeen, P.R., 1989. Behavior as a trajectory through a field of attractors. In: Brink, J.R., Haden, C.R. (Eds.), The Computer and the Brain: Perspectives on Human and Artificial Intelligence. Elsevier, Amsterdam.

Killeen, P.R., 1992. Mechanics of the animate. J. Exp. Anal. Behav. 57, 429–463. https://doi.org/10.1901/jeab.1992.57-429.

Killeen, P.R., 2007. Replication statistics. In: Osborne, J.W. (Ed.), Best Practices in Quantitative Methods. Sage, Thousand Oaks, CA.

Killeen, P.R., 2011. Models of trace decay, eligibility for reinforcement, and delay of reinforcement gradients, from exponential to hyperboloid. Behav. Processes 87, 57–63. https://doi.org/10.1016/j.beproc.2010.12.016.

Killeen, P.R., 2013. The structure of scientific evolution. Behav. Anal. 36, 325–344.

Killeen, P.R., 2014a. Pavlov + skinner = Premack. Int. J. Comp. Psychol. 27, 544–568 escholarship.org/uc/item/1v20v1dg.

Killeen, P.R., 2014b. A theory of behavioral contrast. J. Exp. Anal. Behav. 102, 363–390. https://doi.org/10.1002/jeab.107.

Killeen, P.R., 2019a. In: Zilio, D. (Ed.), The Behavioral Psychophysics of SS Stevens. Behaviorisms, Brazil.

Killeen, P.R., 2019b. Predict, control, and replicate to understand: how statistics can foster the fundamental goals of science. Perspect. Behav. Sci. 109–132. https://doi.org/10.1007/s40614-018-0171-8.

Killeen, P.R., Hall, S.S., Reilly, M.P., Kettle, L.C., 2002. Molecular analyses of the principal components of response strength. J. Exp. Anal. Behav. 78, 127–160. https://doi.org/10.1901/jeab.2002.78-127.

Killeen, P.R., Jacobs, K.W., 2017a. Coal is not black, snow is not white, food is not a reinforcer: The roles of affordances and dispositions in the analysis of behavior. Behav. Anal. 40, 17–38. https://doi.org/10.1007/s40614-016-0080-7.

Killeen, P.R., Jacobs, K.W., 2017b. The modulated contingency. Behav. Anal. 1–20. https://doi.org/10.1007/s40614-017-0101-1.

Killeen, P.R., Nevin, J.A., 2018. The basis of behavioral momentum in the nonlinearity of strength. J. Exp. Anal. Behav. 109, 4–32. https://doi.org/10.1002/jeab.304.

Killeen, P.R., Pellon, R., 2013. Adjunctive behaviors are operants. Learn. Behav. 41, 1–24. https://doi.org/10.3758/s13420-012-0095-1.

Killeen, P.R., Sanabria, F., Dolgov, I., 2009. The dynamics of conditioning and extinction. J. Exp. Psychol. Anim. Behav. Process. 35, 447–472. https://doi.org/10.1037/a0015626.

Klatt, K.P., Morris, E.K., 2001. The premack principle, response deprivation, and establishing operations. Behav. Anal. 24, 173–180 PMID: 22478362.

Labajos, M.J., Pellón, R., 2018. Activity-Based Anorexia and Food Schedule Induction. Handbook of Famine, Starvation, and Nutrient Deprivation: From Biology to Policy. pp. 1–15.

Lajoie, J., Bindra, D., 1976. An interpretation of autoshaping and related phenomena in terms of stimulus-incentive contingencies alone. Can. J. Psychol. 30, 157–172. https://doi.org/10.1037/h0082057.

Lattal, K.A., St Peter, C., Escobar, R., 2013. Operant extinction: elimination and generation of behavior. APA Handbook Behav. Anal. 2, 77–107.

Lieberman, D.A., Davidson, F.H., Thomas, G.V., 1985. Marking in pigeons: the role of memory in delayed reinforcement. J. Exp. Psychol. Anim. Behav. Process. 11, 611–624. https://doi.org/10.1037/0097-7403.11.4.611.

Lieberman, D.A., McIntosh, D.C., Thomas, G.V., 1979. Learning when reward is delayed: a marking hypothesis. J. Exp. Psychol. Anim. Behav. Process. 5, 224–242. https://doi.org/10.1037/0097-7403.5.3.224.

Locurto, C.M., Terrace, H.S., Gibbon, J., 1981. Autoshaping and Conditioning Theory. Academic Press, New York.

Lorenz, K., 1950. The Comparative Method in Studying Innate Behaviour Patterns, Symposia of the Society for Experimental Biology. Cambridge University Press, Cambridge.

MacNulty, D.R., Mech, L.D., Smith, D.W., 2007. A proposed ethogram of large-carnivore predatory behavior, exemplified by the wolf. J. Mammal. 88, 595–605.

Maier, N.R.F., Schneirla, T.C., 1935. Principles of Animal Psychology. McGraw-Hill Book Company, New York 529 pp.

Marley, E., Morse, W.H., 1966. Operant conditioning in the newly hatched chicken. J. Exp. Anal. Behav. 9, 95–103. https://doi.org/10.1901/jeab.1966.9-95.

Marr, M., 1979. Second-order schedules and the generation of unitary response sequences. In: Zeiler, M.D., Harzem, P. (Eds.), Advances in the Analysis of Behavior. John Wiley & Sons, New York.

Mather, J.G., 1981. Wheel-running activity: a new interpretation. Mamm. Rev. 11, 41–51. https://doi.org/10.1111/j.1365-2907.1981.tb00246.x.

McDevitt, M.A., Dunn, R.M., Spetch, M.L., Ludvig, E.A., 2016. When good news leads to bad choices. J. Exp. Anal. Behav. 105, 23–40. https://doi.org/10.1002/jeab.192.

McFall, R.M., Allison, J., Viken, R.J., Timberlake, W., 2019. Response disequilibrium therapy. Clin. Psychol. Sci in press.

McGaugh, J.L., LePort, A., 2014. Rememberance of all things past. Sci. Am. 41–45.

McKearney, J.W., 1968. Maintenance of responding under a fixed-interval schedule of electric shock-presentation. Science 160, 1249–1251. https://doi.org/10.1901/jeab.1969.12-1.

Meijer, J.H., Robbers, Y., 2014. Wheel running in the wild. Proceedings of the Royal Society B: Biological Sciences 281. https://doi.org/10.1098/rspb.2014.0210.

Moore, J., 1975. On the principle of operationism in a science of behavior. Behaviorism 3, 120–138.

Morrison, S.E., Bamkole, M.A., Nicola, S.M., 2015. Sign tracking, but not goal tracking, is resistant to outcome devaluation. Front. Neurosci. 9, 468. https://doi.org/10.3389/fnins.2015.00468.

Moser, J., Burla, J.-B., Gygax, L., 2019. Executing specific foraging behaviours does not represent a general goal state of foraging in dry sows (*Sus scrofa*). Behav. Processes 164, 115–122. https://doi.org/10.1016/j.beproc.2019.05.005.

Myerson, J., 1973. Trace Stimulus Control in Second-Order Schedules. Department of Psychology, Arizona State Universiey, Tempe, AZ.

Neuringer, A.J., 1969. Animals respond for food in the presence of free food. Science 166, 399–401. https://doi.org/10.1126/science.166.3903.399.

Neuringer, A.J., 1970. Many responses per food reward with free food present. Science 169, 503–504. https://doi.org/10.1126/science.169.3944.503.

Neuringer, A.J., Chung, S.H., 1967. Quasi-reinforcement: control of responding by a percentage-reinforcement schedule. J. Exp. Anal. Behav. 10, 45–54. https://doi.org/10.1901/jeab.1967.10-45.

Nevin, J.A., 1988. Behavioral momentum and the partial reinforcement effect. Psychol. Bull. 103, 44–56. https://doi.org/10.1037/0033-2909.103.1.44.

Norman, D.A., 1981. Categorization of action slips. Psychol. Rev. 88, 1–15. https://doi.org/10.1037/0033-295X.88.1.1.

Ohman, A., Flykt, A., Esteves, F., 2001. Emotion drives attention: detecting the snake in the grass. J. Exp. Psychol. Gen. 130, 466–478. https://doi.org/10.1037/AXJ96-3445.130.3.466.

Osborne, S.R., 1977. The free food (contrafreeloading) phenomenon: a review and analysis. Anim. Learn. Behav. 5, 221–235. https://doi.org/10.3758/BF03209232.

Panksepp, J., 1989. The neurobiology of emotions: of animal brains and human feelings. In: Wagner, H.L., Manstead, A. (Eds.), Handbook of Social Psychophysiology. Wiley, New York.

Pearce, J.M., Hall, G., 1978. Overshadowing the instrumental conditioning of a lever-press response by a more valid predictor of the reinforcer. J. Exp. Psychol. Anim. Behav. Process. 4, 356. https://doi.org/10.1037/0097-7403.4.4.356.

Pellis, S.M., Burghardt, G.M., Palagi, E., Mangel, M., 2015. Modeling play: distinguishing between origins and current functions. Adapt. Behav. 23, 331–339.

Pellis, S.M., Pellis, V.C., Pelletier, A., Leca, J.-B., 2019. Is play a behavior system, and, if so, what kind? Behav. Processes.

Pellon, R., Killeen, P.R., 2015. Responses compete and collaborate, shaping each others' distributions: commentary on Boakes, Patterson, Kendig, and Harris (2015). J. Exp. Psychol. Anim. Learn. Cogn. 41, 444–451. https://doi.org/10.1037/xan0000067.

Perkins Jr, C.C., 1968. An analysis of the concept of reinforcement. Psychol. Rev. 75, 155–172 PMID: 4870551.

Peterson, G.B., Trapold, M.A., 1980. Effects of altering outcome expectancies on pigeons' delayed conditional discrimination performance. Learn. Motiv. 11, 267–288. https://doi.org/10.1016/0023-9690(80)90001-6.

Pierce, W.D., Epling, W.F., 1996. Theoretical developments in activity anorexia. In: Epling, W.F., Pierce, W.D. (Eds.), Activity Anorexia: Theory, Research, and Treatment. Lawrence Erlbaum Associates, Mahwah, NJ.

Pilley, J.W., Reid, A.K., 2011. Border collie comprehends object names as verbal referents. Behav. Processes 86, 184–195. https://doi.org/10.1016/j.beproc.2010.11.007.

Platt, J.R., 1972. Amount of training, deprivation, and variability of chain length as determinants of response-velocity gradients in homogeneous chains. J. Exp. Psychol. 92, 191–197. https://doi.org/10.1037/h0032082.

Premack, D., 1959. Toward empirical behavior laws. I. positive reinforcement. Psychol. Rev. 66, 219–233. https://doi.org/10.1037/h0040891.

Premack, D., 1965. Reinforcement theory. In: Levine, D. (Ed.), Nebraska Symposium on Motivation. University of Nebraska Press, Lincoln.

Premack, D., 1971. Catching up with common sense, or two sides of a generalizaiton: reinforcement and punishment. In: Blaser, R. (Ed.), The Nature of Reinforcement. Academic Press, New York.

Premack, D., 2010. Why humans are unique: three theories. Perspect. Psychol. Sci. 5, 22–32. https://doi.org/10.1177/1745691609356782.

Premack, D., Woodruff, G., 1978. Does the chimpanzee have a theory of mind? Behav. Brain Sci. 1, 515–526. https://doi.org/10.1017/S0140525X00076512.

Rachlin, H., Battalio, R., Kagel, J., Green, L., 1981. Maximization theory vindicated. Behav. Brain Sci. 4, 405–417. https://doi.org/10.1017/S0140525X00009602.

Ramstead, M.J., Kirchhoff, M.D., Constant, A., Friston, K.J., 2019. Multiscale integration: beyond internalism and externalism. Synthese 1–30. https://doi.org/10.1007/s11229-019-02115-x.

Reed, P., Hall, G., 1989. The quasi-reinforcement effect: the influence of brief stimuli uncorrelated with reinforcement on variable ratio schedules. Learn. Motiv. 20, 242–261. https://doi.org/10.1016/0023-9690(89)90007-6.

Reid, A.K., Futch, S.E., Ball, K.M., Knight, A.G., Tucker, M., 2017. Assessment of progressively delayed prompts on guided skill learning in rats. Learn. Behav. 45, 62–75. https://doi.org/10.3758/s13420-016-0241-2.

Robinson, T.E., Berridge, K.C., 2000. The psychology and neurobiology of addiction: an incentive-sensitization view. Addiction 95 (Suppl. 2), S91–117. https://doi.org/10.1046/j.1360-0443.95.8s2.19.x.

Routtenberg, A., Kuznesof, A.W., 1967. Self-starvation of rats living in activity wheels on a restricted feeding schedule. J. Comp. Physiol. Psychol. 64, 414–421. https://doi.org/10.1037/h0025205.

Sanabria, F., Daniels, C., Gupta, T., Dos Santos, C., 2019. A computational implementation of the behavior systems approach to learning, motivation, and cognition. Behav. Processes in press.

Sanabria, F., Sitomer, M.T., Killeen, P.R., 2006. Negative automaintenance omission training is effective. J. Exp. Anal. Behav. 86, 1–10. https://doi.org/10.1901/jeab.2006.36-05.

Saunders, B.T., Robinson, T.E., 2011. Individual variation in the motivational properties of cocaine. Neuropsychopharmacology 36, 1668–1676. https://doi.org/10.1038/npp.2011.48.

Sawisch, L.P., Denny, M.R., 1973. Reversing the reinforcement contingencies of eating and keypecking behaviors. Anim. Learn. Behav. 1, 189–192. https://doi.org/10.3758/BF03199072.

Schuster, R., 1969. A functional analysis of conditioned reinforcement. In: Hendry, D. (Ed.), Conditioned Reinforcement. The Dorsey Press, Homewood, IL.

Segal, E.F., 1972. Induction and the provenance of operants. In: Gilbert, R.M., Millenson, J.R. (Eds.), Reinforcement: Behavioral Analyses. Academic Press, New York.

Seligman, M.E., Hager, J.L., 1972. Biological Boundaries of Learning. Appleton-Century-Crofts, East Norwalk, CT, US.

Shahan, T.A., 2010. Conditioned reinforcement and response strength. J. Exp. Anal. Behav. 93, 269–289. https://doi.org/10.1901/jeab.2010.93-269.

Shahan, T.A., 2017. Moving beyond reinforcement and response strength. Behav. Anal. 40, 107–121. https://doi.org/10.1007/s40614-017-0092-y.

Sherwin, C., 1998. Voluntary wheel running: a review and novel interpretation. Anim. Behav. 56, 11–27. https://doi.org/10.1006/anbe.1998.0836.

Shettleworth, S.J., 1988. Foraging as operant behavior and operant behavior as foraging: what have we learned. In: Bower, G.H. (Ed.), The Psychology of Learning and Motivation: Advances in Research and Theory. Academic Press, New York.

Shettleworth, S.J., 1994a. Commentary: what are behavior systems and what use are they? Psychon. Bull. Rev. 1, 451–456. https://doi.org/10.3758/BF03210949.

Shettleworth, S.J., 1994b. The varieties of learning in development: toward a common framework. In: Hogan, J.A. (Ed.), Causal Mechanisms of Behavioural Development. Cambridge University Press, Cambridge, UK.

Shettleworth, S.J., 2009. Cognition, Evolution, and Behavior. Oxford University Press.

Shull, R.L., 2004. Bouts of responding on variable-interval schedules: effects of deprivation level. J. Exp. Anal. Behav. 81, 155–167. https://doi.org/10.1901/jeab.2004.81-155.

Shull, R.L., Guilkey, M., Witty, W., 1975. Brief stimuli paired with food: comparison of two schedule arrangements. Psychol. Rep. 36, 287–294. https://doi.org/10.2466/pr0.1975.36.1.287.

Silva, F.J., Timberlake, W., Gont, R.S., 1998. Spatiotemporal characteristics of serial CSs and their relation to search modes and response form. Anim. Learn. Behav. 26, 299–312. https://doi.org/10.3758/BF03199223.

Silva, K.M., Silva, F.J., Machado, A., 2019. The evolution of the behavior systems framework and its connection to interbehavioral psychology. Behav. Processes 158, 117–125. https://doi.org/10.1016/j.beproc.2018.11.001.

Silva, K.M., Timberlake, W., 1997. A behavior systems view of conditioned states during long and short CS–US intervals. Learn. Motiv. 28, 465–490.

Silva, K.M., Timberlake, W., 1998. A behavior systems view of responding to probe stimuli during an interfood clock. Anim. Learn. Behav. 26, 313–325. https://doi.org/

10.3758/BF03199224.

Silva, K.M., Timberlake, W., 2005. A behavior systems view of the organization of mul- tiple responses during a partially or continuously reinforced interfood clock. Learn. Behav. 33, 99–110. https://doi.org/10.3758/BF03196054.

Skinner, B.F., 1945/1984. The operational analysis of psychological terms. In: Catania, A.C., Harnad, S. (Eds.), The Operant Behaviorism of B. F. Skinner, Behavioral and Brain Sciences.

Smedley, E.B., Smith, K.S., 2018. Evidence of structure and persistence in motivational attraction to serial Pavlovian cues. Learn. Mem. 25, 78–89. https://doi.org/10.1101/lm.046599.117.

Smith, K.S., Graybiel, A.M., 2013. A dual operator view of habitual behavior reflecting cortical and striatal dynamics. Neuron 79, 361–374. https://doi.org/10.1016/j.neuron.2013.05.038.

Smith, K.S., Graybiel, A.M., 2014. Investigating habits: strategies, technologies and models. Front. Behav. Neurosci. 8, 39. https://doi.org/10.3389/fnbeh.2014.00039.

Smolin, L., 2006. The Trouble With Physics. Haughton Mifflin Company, Boston.

Sosa, R., dos Santos, C.V., Flores, C., 2011. Training a new response using conditioned reinforcement. Behav. Processes 87, 231–236.

Spear, N.N., Hill, W.F., 1962. Methodological note: weight loss in rats living in running wheel cages. Psychol. Rep. 11, 437–438.

Staddon, J.E.R., 1979. Operant behavior as adaptation to constraint. J. Exp. Psychol. Gen. 108, 48–67. https://doi.org/10.1037/0096-3445.108.1.48.

Staddon, J.E.R., 1999. Theoretical behaviorism. In: O'Donohue, W., Kitchener, R. (Eds.), Handbook of Behaviorism. Academic Press, San Diego, CA.

Staddon, J.E.R., 2013. Limits to Action: the Allocation of Individual Behavior. Academic Press.

Staddon, J.E.R., 2017. Theoretical behaviorism. Behav. Philos. 45, 27–44.

Stegner, W., 2002. Crossing to Safety. Modern Library.

Stevens, S.S., 1935. The operational definition of psychological concepts. Psychol. Rev. 42, 517–527.

Stiers, M., Silberberg, A., 1974. Lever-contact responses in rats: automaintenance with and without a negative response-reinforcer dependency. J. Exp. Anal. Behav. 22, 497–506. https://doi.org/10.1901/jeab.1974.22-497.

Stigler, S.M., 1999. Statistics on the Table. Harvard University Press, Cambridge.

Stubbs, D.A., 1971. Second-order schedules and the problem of conditioned reinforce- ment. J. Exp. Anal. Behav. 16, 289–313. https://doi.org/10.1901/jeab.1971.16-289.

Su, H., Liu, W., Xu, H., Yang, J., Su, B., Zhang, X., Wang, R., Li, Y., 2018. Introducing chicken farming into traditional ruminant-grazing dominated production systems for promoting ecological restoration of degraded rangeland in northern China. Land Degrad. Dev. 29, 240–249.

Suboski, M.D., 1990. Releaser-induced recognition learning. Psychol. Rev. 97, 271–284. https://doi.org/10.1037/0033-295X.97.2.271.

System_Dynamics, 2019. System dynamics. Wikipedia, The Free Encyclopedia.

Systems_Biology, 2019. Wikipedia, the Free Encyclopedia. https://en.wikipedia.org/w/index.php?title=Systems_biology&oldid=801391167.

Teigen, K.H., 2002. One hundred years of laws in psychology. Am. J. Psychol. 115, 103–118. https://doi.org/10.2307/1423676.

Terhune, J.G., Premack, D., 1970. On the proportionality between the probability of not- running and the punishment effect of being forced to run. Learn. Motiv. 1, 141–149. https://doi.org/10.1016/0023-9690(70)90080-9.

Thomas, J.R., 1966. Avoidance of a return to the first component of a chain from the terminal component. J. Exp. Anal. Behav. 9, 435–441. https://doi.org/10.1901/jeab.1966.9-435.

Thorndike, E.L., 1913. Educational Psychology Vol II: The Psychology of Learning. Teachers College Press, Columbia University, New York.

Thrailkill, E.A., Bouton, M.E., 2016. Extinction and the associative structure of hetero- geneous instrumental chains. Neurobiol. Learn. Mem. 133, 61–68. https://doi.org/10.1016/j.nlm.2016.06.005.

Thrailkill, E.A., Bouton, M.E., 2017a. Effects of outcome devaluation on instrumental behaviors in a discriminated heterogeneous chain. J. Exp. Psychol. Anim. Learn. Cogn. 43, 88–95. https://doi.org/10.1037/xan0000119.

Thrailkill, E.A., Bouton, M.E., 2017b. Factors that influence the persistence and relapse of discriminated behavior chains. Behav. Processes 141, 3–10. https://doi.org/10.1016/j.beproc.2017.04.009.

Thrailkill, E.A., Shahan, T.A., 2014. Temporal integration and instrumental conditioned reinforcement. Learn. Behav. 42, 201–208. https://doi.org/10.3758/s13420-014-0138-x.

Thrailkill, E.A., Trask, S., Vidal, P., Alcala, J.A., Bouton, M.E., 2018. Stimulus control of actions and habits: a role for reinforcer predictability and attention in the develop- ment of habitual behavior. J. Exp. Psychol. Anim. Learn. Cogn. 44, 370–384. https://doi.org/10.1037/xan0000188.

Timberlake, W., 1980. A Molar Equilibrium Theory of Learned Performance, Psychology of Learning and Motivation. Elsevier.

Timberlake, W., 1983b. Rats' responses to a moving object related to food or water: a behavior-systems analysis. Anim. Learn. Behav. 11, 309–320. https://doi.org/10.3758/BF03199781.

Timberlake, W., 1983a. The functional organization of appetitive Behavior: behavior systems and learning. In: Zeiler, M.D., Harzem, P. (Eds.), Advances in Analysis of Behavior. Wiley, New York.

Timberlake, W., 1984. Behavior regulation and learned performance: some mis- apprehensions and disagreements. J. Exp. Anal. Behav. 41, 355–375. https://doi.org/10.1901/jeab.1984.41-355.

Timberlake, W., 1993. Behavior systems and reinforcement: an integrative approach. J. Exp. Anal. Behav. 60, 105–128. https://doi.org/10.1901/jeab.1993.60-105.

Timberlake, W., 1994. Behavior systems, associationism, and Pavlovian conditioning. Psychon. Bull. Rev. 1, 405–420. https://doi.org/10.3758/BF03210945.

Timberlake, W., 1995. Reconceptualizing reinforcement - a causal, system approach to reinforcement and behavior-change. In: Odonohue, W., Krasner, L. (Eds.), Theories of Behavior Therapy. Amer Psychological Assoc, 750 First Street NE, Washington, DC 20002.

Timberlake, W., 1999. Biological behaviorism. In: O'Donahue, W., Kitchener, R. (Eds.), Handbook of Behaviorism. Academic Press, New York.

Timberlake, W., 2000. Motivational modes in behavior systems. In: Mowrer, R.R., Klein, S.B. (Eds.), Handbook of Contemporary Learning Theories. Erlbaum Associates, Mawah, NJ.

Timberlake, W., 2001. Integrating niche-related and general process approaches in the study of learning. Behav. Processes 54, 79–94. https://doi.org/10.1016/S0376-6357(01)00151-6.

Timberlake, W., Allison, J., 1974. Response deprivation: an empirical approach to in- strumental performance. Psychol. Rev. 81, 146–164. https://doi.org/10.1037/h0036101.

Timberlake, W., Fanselow, M.S., 1994. Introduction to: symposium on behavior systems- learning, neurophysiology, and development. Psychon. Bull. Rev. 1, 403–404.

Timberlake, W., Farmer-Dougan, V.A., 1991. Reinforcement in applied settings: figuring out ahead of time what will work. Psychol. Bull. 110, 379–391. https://doi.org/10.1037/0033-2909.110.3.379.

Timberlake, W., Grant, D.L., 1975. Auto-shaping in rats to the presentation of another rat predicting food. Science 690–692. https://doi.org/10.1126/science.190.4215.690.

Timberlake, W., Lucas, G.A., 1985. The basis of superstitious behavior: chance con- tingency, stimulus substitution, or appetitive behavior? J. Exp. Anal. Behav. 44, 279–299. https://doi.org/10.1901/jeab.1985.44-279.

Timberlake, W., Lucas, G.A., 1989. Behavior systems and learning: from misbehavior to general principles. In: Klein, S.B., Mowrer, R.R. (Eds.), Contemporary Learning Theories: Instrumental Conditioning Theory and the Impact of Constraints on Learning. Erlbaum, Hillsdale, NJ.

Timberlake, W., Silva, F.J., 1994. Observation of behavior, inference of function, and the study of learning. Psychon. Bull. Rev. 1, 73–88. https://doi.org/10.3758/bf03200762.

Timberlake, W., Wahl, G., King, D.A., 1982. Stimulus and response contingencies in the misbehavior of rats. J. Exp. Psychol. Anim. Behav. Process. 8, 62–85. https://doi.org/10.1037/0097-7403.8.1.62.

Timberlake, W., Wozny, M., 1979. Reversibility of reinforcement between eating and running by schedule changes: a comparison of hypotheses and models. Anim. Learn. Behav. 7, 461–469. https://doi.org/10.3758/BF03209702.

Tinbergen, N., 1952. "Derived" activities; their causation, biological significance, origin, and emancipation during evolution. Q. Rev. Biol. 27, 1–32.

Tinsley, M.R., Timberlake, W., Sitomer, M., Widman, D.R., 2002. Conditioned inhibitory effects of discriminated Pavlovian training with food in rats depend on interactions of search modes, related repertoires, and response measures. Anim. Learn. Behav. 30, 217–227. https://doi.org/10.3758/BF03192831.

Tomie, A., Grimes, K.L., Pohorecky, L.A., 2008. Behavioral characteristics and neuro- biological substrates shared by Pavlovian sign-tracking and drug abuse. Brain Res. Rev. 58, 121–135. https://doi.org/10.1016/j.brainresrev.2007.12.003.

Tomie, A., Morrow, J., 2018. Sign-Tracking and Drug Addiction. Maize Books.

Urcuioli, P.J., 2005. Behavioral and associative effects of differential outcomes in dis- crimination learning. Learn. Behav. 33, 1–21. https://doi.org/10.3758/BF03196047.

Urcuioli, P.J., 2006. When discrimination fails (or at least falters). J. Exp. Psychol. Anim. Behav. Process. 32, 359–370. https://doi.org/10.1037/0097-7403.32.4.359.

van den Berg, L., Schilder, M.B., Knol, B.W., 2003. Behavior genetics of canine aggression: behavioral phenotyping of golden retrievers by means of an aggression test. Behav. Genet. 33, 469–483.

Vandbakk, M., Olaff, H.S., Holth, P., 2018. Conditioned reinforcement: the effectiveness of stimulus—stimulus pairing and operant discrimination procedures. Psychol. Rec. 1–15. https://doi.org/10.1007/s40732-018-0318-8.

Wasserman, E.A., 1973. The effect of redundant contextual stimuli on autoshaping the pigeon's keypeck. Anim. Learn. Behav. 1, 198–206. https://doi.org/10.3758/BF03199074.

Wasserman, E.A., Franklin, S.R., Hearst, E., 1974. Pavlovian appetitive contingencies and approach versus withdrawal to conditioned stimuli in pigeons. J. Comp. Physiol. Psychol. 86, 616–627. https://doi.org/10.1037/h0036171.

Wasserman, E.A., Hunter, N., Gutowski, K., Bader, S., 1975. Autoshaping chicks with heat reinforcement: the role of stimulus-reinforcer and response-reinforcer relations. J. Exp. Psychol. Anim. Behav. Process. 104, 158–169. https://doi.org/10.1037/0097-7403.1.2.158.

Watson, P., de Wit, S., 2018. Current limits of experimental research into habits and future directions. Curr. Opin. Behav. Sci. 20, 33–39. https://doi.org/10.1016/j.cobeha.2017.09.012.

Watson, P., Wiers, R.W., Hommel, B., de Wit, S., 2014. Working for food you don't desire. Cues interfere with goal-directed food-seeking. Appetite 79, 139–148. https://doi.org/10.1016/j.appet.2014.04.005.

Williams, B.A., 1975. The blocking of reinforcement control. J. Exp. Anal. Behav. 24, 215–225.

Williams, B.A., 1991. Marking and bridging versus conditioned reinforcement. Anim. Learn. Behav. 19, 264–269. https://doi.org/10.3758/bf03197885.

Williams, B.A., 1999. Associative competition in operant conditioning: blocking the re- sponse-reinforcer association. Psychon. Bull. Rev. 6, 618–623.

Williams, D.R., Williams, H., 1969. Auto-maintenance in the pigeon: sustained pecking despite contingent non-reinforcement. J. Exp. Anal. Behav. 12, 511–520. https://doi.org/10.1901/jeab.1969.12-511.

Wood, W., Rünger, D., 2016. Psychology of habit. Annu. Rev. Psychol. 67, 289–314.

Zabegalov, K.N., Khatsko, S.L., Lakstyga, A.M., Demin, K.A., Cleal, M., Fontana, B.D., McBride, S.D., Harvey, B.H., de Abreu, M.S., Parker, M.O., 2019. Abnormal repetitive

behaviors in zebrafish and their relevance to human brain disorders. Behav. Brain Res. 367, 101–110. https://doi.org/10.1016/j.bbr.2019.03.044.

Zapata, A., Minney, V.L., Shippenberg, T.S., 2010. Shift from goal-directed to habitual cocaine seeking after prolonged experience in rats. J. Neurosci. 30, 15457–15463. https://doi.org/10.1523/JNEUROSCI.4072-10.2010.

Zentall, T.R., 2013. Animals prefer reinforcement that follows greater effort: justification of effort or within-trial contrast? Comp. Cogn. Behav. Rev. 8, 60–77. https://doi.org/10.3819/ccbr.2013.80004.

Zentall, T.R., 2016. Resolving the paradox of suboptimal choice. J. Exp. Psychol. Anim. Learn. Cogn. 42, 1–14. https://doi.org/10.1037/xan0000085.

Zentall, T.R., Andrews, D.M., Case, J.P., 2019a. Contrast between what is expected and what occurs increases pigeon's suboptimal choice. Anim. Cogn. https://doi.org/10.1007/s10071-018-1223-x.

Zentall, T.R., Laude, J.R., 2013. Do pigeons gamble? I wouldn't bet against it. Curr. Dir. Psychol. Sci. 22, 271–277. https://doi.org/10.1177/0963721413480173.

Zentall, T.R., Smith, A.P., Beckmann, J., 2019b. Differences in rats and pigeons

suboptimal choice may depend on where those stimuli are in their behavior system. Behav. Processes 159, 37–41. https://doi.org/10.1016/j.beproc.2018.11.012.

Zentall, T.R., Stagner, J., 2011. Maladaptive choice behaviour by pigeons: an animal analogue and possible mechanism for gambling (sub-optimal human decision-making behaviour). Proc. R. Soc. B Biol. Sci. 278, 1203–1208. https://doi.org/10.1098/rspb.2010.1607.

Zimbardo, P.G., 1973. A field experiment in auto shaping. In: Ward, C. (Ed.), Vandalism. Architectural Press, London.

Zimmerman, J., 1963. Technique for sustaining behavior with conditioned reinforcement. Science 142, 682–684. https://doi.org/10.1126/science.142.3593.682.

Zimmerman, J., 1969. Meanwhile… back at the key: maintenance of behavior by conditioned reinforcement and response-independent primary reinforcement. In: Hendry, D.P. (Ed.), Conditioned Reinforcement. The Dorsey Press, Homewood, IL.

Zuriff, G.E., 1979. Ten inner causes. Behaviorism 7, 1–8. http://www.jstor.org/stable/27758925.