



Pavlov + Skinner = Premack

Peter R. Killeen

Arizona State University

Behavior is a sequence of actions. Premackian conditioning occurs when one of those actions permits an animal to engage in more biologically potent positive responses—reinforcement—or forces them to engage in less positive (or negative) responses — punishment. Signals of the transition from one class of actions to another focus the instrumental responses in the first class and inform the contingent responses in the second class. The signals may be innate (USs) or learned (sign-learning); excitatory (leading to more positive actions) or inhibitory (leading to less positive actions). The potency of an action has phylogenetic origins, but may be conditioned by proximity to more potent responses, such as consummation of a reinforcer. With practice instrumental responses may take on increased strength, and in some cases become motivationally autonomous—become habits. Stimuli or responses that signal the availability of more positive actions may become incentive motivators that animals will approach. Discriminative stimuli do not have intrinsic value as reinforcers, but only the value derived from the responses that they release. These forces bend an animal’s trajectory through its stimulus-action-time context into a path that leads more directly to positive actions. The association of actions (conditioned responses, operants, and observing responses) with actions of different potency (ultimately unconditioned responses or consummatory behavior) is the primary association in Premackian conditioning. All other types of conditioning may be interpreted as instances of such Premackian conditioning.

It is the thesis of this paper that all conditioning is Premackian. Thousands of papers have been written about the relations between classical and neoclassical conditioning, suggesting or rejecting attempts at unification. “Belief in the classical-instrumental distinction may have impeded as much as stimulated research” (Hearst, 1975, p. 218). Classical and neoclassical procedures are different, and the response that is measured is typically different. The changes in behavior that issue from both kinds of procedures are instances of Premackian conditioning; conditioning of the signs that signal a change in modality and the actions that are relevant to them, and conditioning of the particular actions that will lead to a better state of affairs.

Premackian Conditioning

In a stroke of genius, Premack (1959, 1965, 1971) noted the obvious and elevated it into a principle. He noted that it is not stimuli that reinforce actions, but rather actions that reinforce actions (see the appendix for a glossary of terms). Grain doesn't reinforce pecking, nor do Noyes pellets reinforce lever pressing. Eating reinforces both. Most of the things we work for are the opportunities to expend time and energy in doing things that we enjoy: cooking, driving a car, listening to music, reading a book, playing golf. In none of those cases is the termination of the activity the goal of the activity. Were I to tell you that the butler did it, I would not concentrate your pleasure, but rather devastate it. "People learn to ration rewards — sipping wines, delaying orgasm, and so on...having learned that the race to consummation is only a short-term preference for the smaller total utility" (Herrnstein, 1988, p. 572). As is often true of travel, it is the journey that is the destination, the destination an excuse for the journey. The act of sipping and the act of making love become ends in themselves, as the child grows into the connoisseur. This is one of the key insights of molar behaviorism (Baum, 2004, 2012; Rachlin, 1992, 2014). The perspective taken in this paper shares that aspect of molar behaviorism, but is perhaps more akin to the multiscaled approach toward which Shimp has been struggling over the decades (Pitts, 2013; Shimp, 2013). It bears strong affinities to Denny's (e.g., Denny & Adelman, 1955) and Donahoe's (e.g., 2014) elicitation theories.

In his many seminal experiments, Premack (Hundt & Premack, 1963; Premack, 1963; Terhune & Premack, 1970) and colleagues demonstrated that the reinforcing relation between two actions such as wheel-running and drinking could be established and reversed by differential histories of availability of those actions. In the popular formulation of the *Premack principle*, one action (the *contingent* behavior) would reinforce another (the *instrumental* action) if: 1) The instrumental action was required for the ability to emit the contingent behavior, and 2) The contingent behavior occurred with higher probability than the instrumental action. Conversely, the contingent behavior would punish the instrumental action if it were forced to occur at a higher than baseline probability. This hypothesis was later refined (Allison & Timberlake, 1974; Eisenberger, Karpman, & Trattner, 1967; Timberlake & Allison, 1974; Timberlake & Wozny, 1979) by showing that for an action to be a reinforcer, it must be depressed below the level at which it would normally occur in that situation. Contra Premack, a low probability behavior can actually reinforce a higher probability action as long as the former is occurring below the level at which it would if unconstrained. Of course such response suppression automatically increases the local probability of the contingent behavior, and the original principle is saved if we read *probability* as referring to probability *in that context*, not in some unconstrained baseline context.

The reason that the rate of a behavior may be depressed is because of experimental or environmental constraints placed on the animal. A standard way of making a pellet reinforcing is to deprive the animal of food, thus depressing the rate of eating, and at the same time increasing its probability of occurrence when that is unbuckled. The action of eating the pellet will then reinforce the instrumental act of lever pressing for it. These constraints are thus motivational or "establishing" operations (see Klatt & Morris (2001), who draw out this relation to establishing operations after a nice summary of the above issues). To be plain, the animal's behavior is depressed because something is not letting it do what it wants to do, and it will work (the instrumental action) in order to do what it wants to do (the contingent behavior). Because there is decreasing marginal utility in doing what it wants, and because there is disutility in working, the animal will usually settle for a bit less of what it wants in order to not work quite so hard to get it. These observations opened the door to the paradigm of behavioral economics (Allison, 1983), assaying different models that minimized deviation from an optimal combination of actions (e.g., Hanson & Timberlake, 1983; Staddon, 1979), or maximized a set of utilities of a package of actions tied together by experimental constraints (e.g., Rachlin, Battalio, Kagel, & Green, 1981), or engaged other economic analyses (e.g., Hursh, 1984). Whereas Premack's observations of this tradeoff were seminal, behavioral economics is not central to the ideas pursued here.

What Rats Really Want

The key idea of Premackian conditioning is that actions, and the opportunity to make them, reinforce other actions. The subjects in laboratory studies of behavior typically want to eat or drink because we made them hungry or thirsty. They want to run because we keep them in cages. But in the wild they are freer to indulge a larger variety of motivations with a larger variety of actions. “We begin to conceive of behavior, which we have always thought of as highly modifiable, as consisting of a lot of fixed packages, software programs as it were.” (Bolles, 1983, p. 43). Bolles’s observation echoed Premack’s (1965, p. 125); “We may view the fetal rat as moving down an assembly line, receiving first a motor from one bin, a little later on a motor from another bin, the bins marked ‘eat’, ‘drink’, ‘run’, etc.’.” We return to the operating characteristics of these motors in a subsequent section. A picture of what some of these packages might be is shown in Figure 1, adapted from Timberlake and Lucas (1989); Hogan (1994, 2014) erected a similar framework. Catania (2013) also described behavior hierarchies, and Hull (1934; 1943) used them in a fashion similar to mine. For excellent overviews, see Gallistel (1980; 1981).

Timberlake’s Rats

Few behaviorists have provided a better model of situating modern behaviorism within its proper larger setting of the ecology of the organism than has Bill Timberlake (see, e.g., Timberlake, 1993, 2001). Figure 1 is one branch of a many-limbed ethogram of actions that the lab rat’s country cousin might engage over the course of the night. At the rightmost level the descriptions are objective, and where ambiguity arises, researchers will link them to operational definitions. The molar levels to the left are more interpretative, and clarify the context in which the actions occur. Thus locomotion could occur equally well as a defensive action in an escape mode, and grabbing and biting in an aggressive or defensive encounter. Timberlake calls the levels *subsystem* (here the highest level of Predation), *mode*, *module* and *action*. Notice that as drawn in Figure 1, the actions at the top such as locomote can occur with the least environmental support. They are least constrained by context. As one moves down the column, the actions need more and more support. They are more constrained. To swallow prey requires the step above it, and so on. It is the thesis of this paper that operations that move an animal into a particular subsystem or mode are *motivational*. When in that subsystem, the ability to engage in an action in the right column will reinforce actions that lie above it in that column, and be reinforced by actions that lie below it in that column. (This is my (1992), not Timberlake’s, extension of this diagram.)

In the above quote, Bolles went on: “These preformed packages can be shifted around from one application, or object, to another” (Bolles, 1983, p. 43) – from predation to escape subsystems, for example. The next year Gould and Marler (1984, p. 66) amplified that observation: “Such a strategy, which involves building up complex motor behavior out of a ‘library’ of innate elements, has obvious advantages for certain tasks.” But the point of embedding such motor acts in an action hierarchy such as shown in Figure 1 is to emphasize that their character depends profoundly on context. Immobility can be freezing to elude the sensors of predators, or camouflage to elude the sensors of prey; a lion’s hold and bite can mean death to an antelope, copulation to a lioness. Context and physiology provide incentive motivation to keep behavior on track. Although there is press for movement down the action chain, only a limited amount of skipping over intermediate steps may be possible. That is because actions high on the hierarchy may not provide releasing stimuli for those much lower on it. Courtship rituals of animals provide many examples that can be read in the classics of ethology (e.g., Kruijt, 1964; Tinbergen, 1961). Movement into a motivational subsystem, and down through the hierarchy of mating responses, intensifies motivation to mate: “the courtship situation, and more especially the pre-coition situation, is characterized by a strong inner mating urge, but as long as the partner’s final signal has not been given, its impulses cannot find an outlet through the mating act” (Tinbergen, 1961, p. 115). The energy spills over to a variety of creative courtship actions, some of which may convince the blushing but curious bride. But this is the situation all the way down the action hierarchy – courtship rituals are checklists, and if your stimuli do not quite release the next action pattern in the potential mate, you become stymied at that spot, left pressing at the bar.

It is in the permutations of actions within modules that stochasticity arises. There are many ways to travel, socialize, investigate, and so on. “Attempting to ‘zoom in’ on a unit always just yields new, smaller, units...Behavior appears very fractal in our behavioral microscopes” (Brembs, personal communication, September 13, 2014). When actions within modules bring no progression down the chain, they habituate, fatigue, or weaken through short-term extinction. Other actions or action blends arise to take their place, which in particular depending on recent histories of reinforcement and fatigue, and environmental affordances (Pryor, Haag, & O'Reilly, 1969). Variability in what to try next can itself become an action – an operant (Neuringer, 2002). “That being said, there probably are ‘basins of attraction’ which limit the choice of actions to a specific subset, or help shape the probabilities with which certain behaviors can be emitted and strung together” (Brembs, personal communication, September 13, 2014).

Dalliance

Each of the actions in Figure 1 can be a source of pleasure – a reinforcer – and many can be modifiable action patterns. Cats notoriously play with their prey. Actions of tracking, cutting-off and grabbing that are more effective in leading to a hold are strengthened – reinforced. *Innate elements* are modified, sequenced and optimized, making cats connoisseurs of the kill. It is helpful “to recognize a continuum from aspects of behavior that are environmentally stable ... to those that are environmentally labile (i.e., varying greatly with environmental influences)” (Hinde, 1982, p. 86). Selective breeding can stall movement down the action chain – sporting dogs that hold a point on prey have a broken link in the pause-lunge-size-kill-eat chain. It is as though the key releaser for *lunge* no longer turns the lock. Retrievers got stuck on a later link, as did herding dogs. The hammer of ontogeny and the anvil of phylogeny between them forge the chain of actions, the points at which selective forces impinge, and how they reverberate to other levels of the hierarchy. Learning is one of the modern techniques in nature’s foundry, making variants on a response that release more potent action more likely in the future. This is Premackian reinforcement.

An even more modern innovation that helps to avoid impasses is the shift in control from releasing stimuli to the virtual stimuli of human imagination. Along with virtual enactment of actions, they permit movement down imaginary action chains, to emulate and evaluate alternate action plans, or merely to delight in the possibilities of alternate journeys. Consciousness, it may be, is merely the flag that marks these enactments as virtual, guarding us from confusion of ideation with action.

Pavlov’s Dogs

Whereas the stimuli that release movement down the chain of action are sometimes innate sign stimuli and the responses to them invariant modal-action patterns, at other times they are more flexible, and constitute the elements of learned action patterns. Pavlov’s dogs did more than salivate when he conditioned them; when untethered they approached the CS+ and then the US sites, and “showed individually distinctive action patterns to the CS+, in some cases suggestive of soliciting, in another of sight-pointing” (Jenkins, Barrera, Ireland, & Woodside, 1978, p. 272). Pavlov’s laboratory studied digestive processes, and it was natural for him to focus on salivation; subsequently researchers elevated the similarity of a response to a UCR to be a criterion for calling that response a CR (e.g., Gormezano & Kehoe, 1975). Whereas behaviorists have largely concerned themselves with the rate of emission of simple glandular secretions or simple limb movements as functions of stimulating conditions, developmental biologists have concerned themselves with the topography of movements, and the coordinated development of higher-level units (see, e.g., Hinde, 1982; Hogan & Bolhuis, 2009; Timberlake, 1983). Peter Balsam is among the rare behaviorists who have contributed to both fields of

investigation. He and his colleagues (e.g., Balsam, Graf, & Silver, 1992) have shown, inter alia, that pairing of the sight of seeds with food is adequate to induce some pecking at seed by squab, but experience in handling

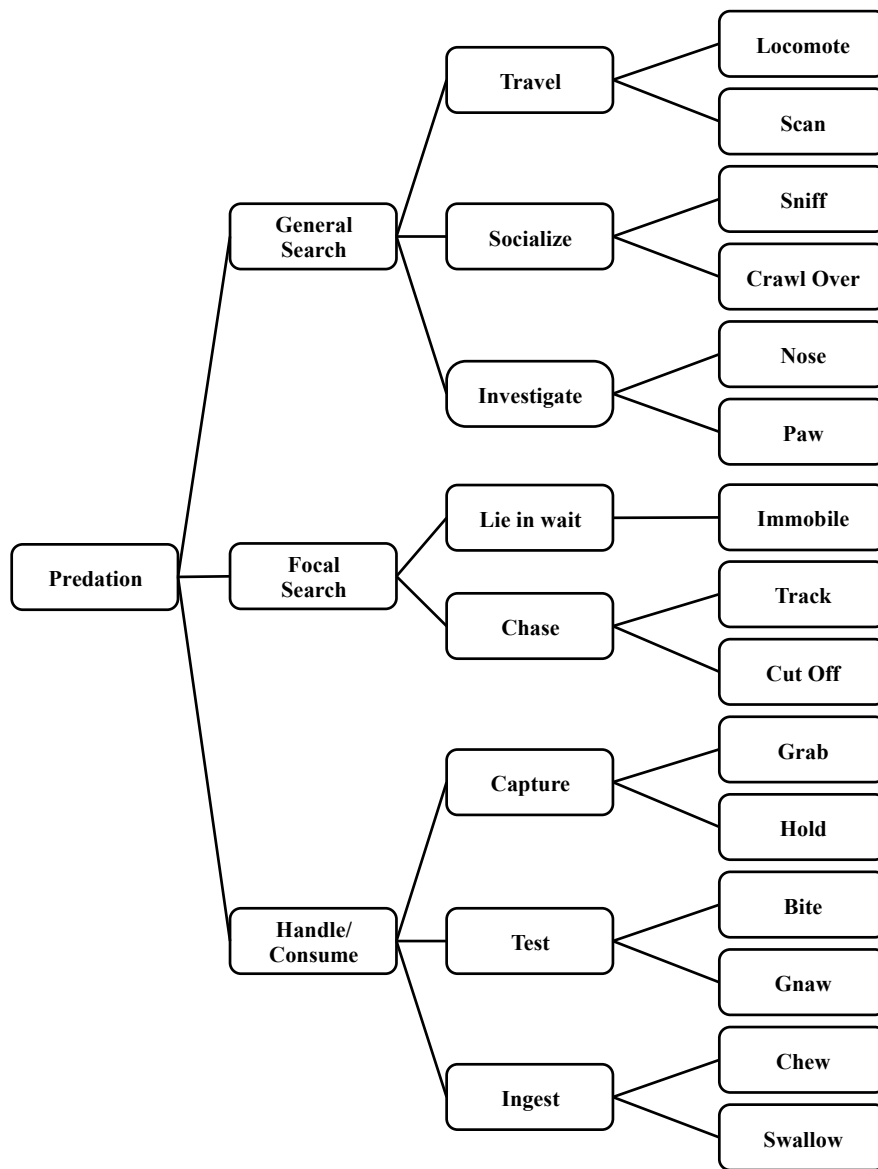


Figure 1. A hierarchy of action patterns, adapted from Timberlake and Lucas (1990). The conceptual categories to the left provide an intuitive organization of the actions on the right. When an organism is in another mode than predation (say, nesting), a different set of actions with a different ordering prevails. The actions are ordered according to their biological potency. In the presence of stimuli that can release (provide affordance for) an action, an animal engages in that action. The action will often introduce stimuli that release actions lower on the chain. It will be attracted to those actions, requiring constraint to keep it at higher levels. Letting the animal approach the lower levels (viz. approach the stimuli that release those actions) converts the potential energy of that gradient into the kinetic energy of motion and conditioning.

the seeds was necessary before they attained proficiency, and pairing with positive ingestional consequences necessary for development of normal seed pecking (see also Hogan, 1994). Balsam and Silver (1994) provide an invaluable review of the development and change of patterns of behavior that bridge the experiments of behaviorists and ethologists. The theme of elicitation of a situationally appropriate action pattern, and the

refinement of it by reinforcement processes, is ubiquitous in this literature (and others, e.g., Killeen & Pellón, 2013; Segal, 1972).

Whereas the use of arbitrary stimuli as the CS is convenient for laboratory research, stimuli that have some relevance to the subsystem and mode of the animal often have different effects than tones and lights (Domjan, 2005), among them: faster learning and much longer delay gradients for taste/poison conditioning; taste-odor potentiation and contra-blocking; resistance to blocking and sensitization rather than habituation with repeated nonreinforced exposure to particular cues in sexual conditioning. Domjan also notes differential effectiveness of different classes of stimuli in fear conditioning. The use of Med-PC stimuli, like the use of college sophomores in psychological research, restricts the generalizations we can make about the associative processes. As Weisman (1977, p. 92) trenchantly noted, the “constraints” on learning may have to be divided between the subject and the experimenter. The increasing use of highly constrained experimental chambers limits what we can learn about behavior. After reviewing a number of other ecological effects in conditioning, Boakes ended his symposium contribution (Boakes, 1977) “on a more positive note... one very definite lesson: we should resist the temptation of taking from the shelf equipment that can be immediately plugged in” (p. 95).

In their study of signal-centered action patterns in dogs, Jenkins and colleagues (Jenkins et al., 1978) proposed that the CS-US episode mimics a naturally occurring sequence for which preorganized action patterns exist. The CS “substitutes for a natural signal, not for the object being signaled as in the Pavlovian concept of substitution” (p. 292). In a section entitled “Conditional Release”, they note the number of “otherwise puzzling” aspects of Pavlovian conditioning that this hypothesis explains. They conclude that [the topographies of] signal-centered 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. On this latter point, see Stokes & Balsam, 1991). This treatment is essentially the same as Timberlake’s *appetitive structure* hypothesis (see, e.g., Timberlake, Wahl, & King, 1982). Bolles (1970) and Fanselow (e.g., Fanselow & Lester, 1988) provide parallel accounts in the case of aversive contexts. In such contexts, it is access to behaviors that will terminate, postpone, or remove an aversive situation that reinforce actions leading to them (Denny, 1991). Konorski (1967) also distinguished levels within the response chain, but lumped them into preparatory and consummatory responses, similar to the ethologists’ appetitive and consummatory responses. Balsam and associates (Balsam, Deich, Ohyama, & Stokes, 1998) found that “This mechanism [reinforcement by action chains] turned out to play a crucial role in the development of pecking in the doves” (p. 412) that they studied.

Depending on its nature (Holland, 1977) and its temporal and spatial relation to the US, the CS may release different elements of the response chain (Holland, 1980; F. J. Silva, Timberlake, & Koehler, 1996; K. M. Silva & Timberlake, 1997), some of which may become autonomously attractive in their own right. With such conditioning, a limited vertical reordering of the action chain is possible. Animals may hunt when satiated (Morgan, 1974), work for food when work is unnecessary (Neuringer, 1970b), work for it in the presence of free food (Mason, Bateson, & Bean, 1999; Neuringer, 1970a; Osborne, 1978) and misbehave in general (Boakes, Poli, Lockwood, & Goodall, 1978; Timberlake et al., 1982). (Some of these phenomena have been shown to be reversible by Premackian constraint: see Sawisch and Denny, 1973.) Such “misbehavior” should not happen if the response were instrumental – goal directed – rather than released by the CS. But that may be due to a misperception of the animal’s goals by the experimenter. The categories of elicited, acquired, emitted/operant are neither exclusive nor fixed. The intrinsic attractiveness of links on the action chain may wax and wane. The pleasure of the chase will generally abate if never consummated, yet has its undeniable charms, adding its own utility to the whole chain. Behavior may proceed, as Tinbergen noted, in ignorance of its effect; but if its effect is to occasion a stimulus that releases an action of greater biological potency – one lower on a response chain – it may be shaped by that lucky happenstance (Timberlake & White, 1990). In cases where those more potent actions are already being released by concurrent stimuli, the conditioning of other aspects of the stimulus complex as signs will be blocked (Donahoe, 2014); the animal is not hung up there, it is already moving on.

Getting Turned On: Incentive Motivation

Recall Tinbergen's herring gull on the beach. It is not generally, at least as far as observation can tell, preoccupied with sex. The presence of appropriate hormones and the sight of an eligible partner will induce the first steps into the mating hierarchy – they move the gull into a particular mode of the mating system. And then the chain reaction will continue – or not – depending on the target's level of appropriate hormones and the power of the stimulation of the releasing stimuli provided by the initiator's courtship actions. Even if descent to the bottom is not possible, entering and moving down the chain can reinforce instrumental responses (for review in the case of sexual conditioning, see Crawford, Holloway, & Domjan, 1993). *L'appétit vient en mangeant*; thus appetizers – or more formally, incentive motivators. Once situated in a mode, the relevant actions are primed, and released by the appropriate sign. “I regard as the most important breakthrough of all of our attempts to understand animal and human behavior the recognition of the following fact: [Cells are tonically excited, and held in check by inhibitory cells.] It is this inhibiting cell which is influenced by the receptor and ceases its inhibitory activity at the biologically ‘right’ moment” (Lorenz, 1973¹). The present thesis modifies this insight only to suggest that the excitation is not being “permanently produced”, as Lorenz claims, but rather produced when the animal moves into a particular mode or module, increasing the probability of relevant actions, to then be released by their pertinent signs. The gull does not permanently have sex on its mind, but only when the season is right and the potential mate in sight; only then may the dance begin.

In a systematic series of experiments (summarized in Holland, 1992), Peter Holland studied the conditioning that occurred in the occasion setting paradigm. In this design, a CS is predictive of both a US and of its absence. Only in the presence of a prior stimulus – the *occasion setter* – is it a CS+. The occasion setter elicits some of the response topographies characteristic of both itself and of the CS, but it becomes *more* potent as the delay to the US increases (complementary to the usual and observed *decrease* in associability of the CS as a function of delay to the US). He speculated about the role of the occasion-setter in multiple memory systems, and in a hierarchical control system. Although he considered the evidence non-committal about these possibilities (but see Holland, 1990), it would be consistent with our treatment if the role of the occasion-setter were to establish the operation of one of the modes or modules in Figure 1, potentiating the possibility that any of the several appropriate actions would occur, and focusing the animal's attention by priming it for the onset of the CS – which would further release those actions. It is an *incentive motivator*. Consistent with this hypothesis, Timberlake (2001) has suggested that motivational operations function to move an animal into a mode relevant to their history and immediate needs; as has Hogan (1988).

The Role of the US

Pairing of CS with US makes the CS a conditioned sign for US-relevant (and CS-relevant) behavior. But is it the pairing of CS with US that matters, or is it the pairing of CS with unconditional response – the UR – that matters? In an elegant series of experiments with throat movements of pigeons elicited by injection of water, Donahoe and Vegas (2004) were able to temporally unconfound the relation of the CS to the US and UR, and found that the CS-UR relationship was the critical factor in this preparation. It was the biologically potent UR that increased the less-potent CR as a function of its temporal relation to it, and to the CS that elicited it. While these investigators showed that the proximity of the UR affected the strength of the CR, Domjan (2005) showed the many ways in which conditioning can change the intensity and nature of the UR.

¹ I thank Dr. Brembs for directing me to this.

The role of the US in general may be to elicit the action sequence of the UR, which is the biologically potent event in the conditioning process. If so, then classical conditioning is an instance of Premackian conditioning.

To distinguish the operations of making a CS a predictor of the US (i.e., Pavlovian sign-conditioning) from its behavioral results (the CRs), Brembs calls sign-conditioning *world-learning*, which he defines as “the process of assigning value to sensory stimuli” (Colomb & Brembs, 2010, p. 142). The nominal distinction between process and outcome was overdue. An alternative to Brembs’s name for the result would be “sign-learning”, closer to Pavlov’s “signalization”, and to our present hypothesis (Brembs & Heisenberg, 2000; Pavlov, 1927). Stimuli that signal the opportunity to engage in more potent responses become conditioned signs; they are approached if the transit to the next link is delayed, and abandoned when the more potent action becomes available. “Approach and withdrawal are the only empirical objective terms applicable to all motivated behavior in all animals” (Schneirla, 1959, p. 1), describing the key hypothesis of the present theory, movement through the links of action chains, guided by signs to more potent links.

Exceptions Prove the Rule

Sensory preconditioning provides prima-facie evidence against the hypothesis that cue-conditioning is Premackian in nature, as the response controlled by the stimuli are chosen to be not particularly biologically potent, and to not elicit a substantial UCR. Most stimuli do, however, elicit orienting responses, and these may mediate sensory conditioning. Whereas these are of modest magnitude, so also is the size of the sensory preconditioning effect (Mackintosh, 1974; Weisman, 1977). Pre-exposure, which habituates the orienting response, undermines subsequent conditioning (Lubow, 1989; Rescorla, 1971). Mackintosh (pp. 92-93) also noted numerous failures to demonstrate conditioned responses to a CS even when the UCS elicited a reliable UCR. But he went on to observe that “animals in both types of experiment [sensory preconditioning and apparent non-conditioning despite UCRs, may] learn that one stimulus follows another, but the formation of such an association is not sufficient to produce a CR to the first stimulus” (p. 96)”. That may wait upon appropriate motivation.

It is possible to reinforce actions with intracranial self-stimulation (ICSS), for which no consummatory response is necessary. Many consummatory responses are nonetheless elicited by ICSS to the hypothalamus, and the reinforcing strength of the stimulation is greater when affordance (e.g., food or water) is provided for them. Because ICSS short-circuits the natural reward system, the lack of a natural consummatory response for it despite its potent reinforcement strength provides only ambiguous evidence against the present thesis. The same may be said of the optogenetic stimulation of dopamine mediated reinforcement circuits. My belief that I cannot start my car with a dead battery is invalidated – but hardly disabused – if you open the hood and jump-start it from your battery.

What consummatory response maintains escape from noxious stimulation? The ability to relax in a safe environment is certainly more attractive than the pained behavior in a shock compartment (Rachlin, 1985). The longer the time to relax, the greater the attraction (Denny & Weisman, 1964; Reynierse & Rizley, 1970), consistent with Denney’s (Denny & Adelman, 1955) elicitation theory of escape/avoidance.

Skinner’s Rats

Whereas the functionality of classical conditioning is now well established (Domjan, 2005; Hollis, 1983; Turkkan, 1989), the functionality of neoclassical – operant – conditioning is its defining characteristic. An operant is a movement that is modified by its antecedents and consequences. Those factors change the location, topography and timing of the movements. We may situate the operant a multidimensional space, one whose dimensions include time, location, and sensory attributes. The topography of that landscape is dictated

by phylogeny, and contoured by early experience and learning (Cairns, Garipey, & Hood, 1990; Gottlieb, 2007). The loci of potent actions are attractors, sign-learning marks routes to those points, and sign-tracking carries the animal toward them (Killeen, 1989, 1991). Functional variants – ones that move the animal more quickly toward the attractor – become differentially favored in its repertoire. All learning works to minimize the path length in this multidimensional space, and continues until it is a geodesic—the shortest realizable path given the configuration of the space.

Successful movement through space needs signposts, at least to shape behavior until the trajectory is learned (Reid, Chadwick, Dunham, & Miller, 2001; Shettleworth, 2009a, 2009b). The impressive feats of trained animals, from circus beasts to the student rats visible on YouTube, typically require a scaffold of discriminative stimuli along their course. These function much the same way as the releasers in the action chains of Figure 1, opening the gate to performance of the next action. The scaffolding is effective even if there is only a brief stimulus marking the correct response (Lieberman, Davidson, & Thomas, 1985; Patterson & Boakes, 2012) or reinforcer (Fedorchak & Bolles, 1986). To the extent that reinforcers are different, differentiated performances are better learned (Trapold, 1970). These *marking* and *differential outcome* effects have been found in a variety of species (see, e.g., Miyashita, Nakajima, & Imada, 2000) for a variety of actions (see, e.g., Patterson & Boakes, 2012).

Biological Constraints

In some cases responses are modular, relatively adamant to shaping (Shettleworth, 1975; Shettleworth & Juergensen, 1980) or to appropriate control by stimuli (Hearst & Jenkins, 1974; Reid, Rapport, & Le, 2013). They are responses – movements elicited by a stimulus – but not operants – they are not under the control of the reinforcer. Like the English Pointer trapped in one link of an action chain and perpetually unable to consummate the nominal goal of that action, laboratory animals have regularly been gotten stuck at local minima. The persistent trainer then tunes his apparatus or procedure to make this bug into a feature (Breland & Breland, 1961; Timberlake, 1994), while the more cautious scientist worries over limitations and predispositions (see, e.g., Hinde, 1973). A good retrospective on the field of biological predispositions and constraints may be found in (Domjan, 2008).

Automatization

Musing over the behavior of his birds, Tinbergen (1961) observed “The most obvious thing about the gulls’ behavior is their lack of insight into the ends served by their activities, and into the way their own behavior serves these ends. ... the chick responds again and again, up to hundreds of times, to a crude dummy despite the fact the dummy never provides it with food” (p. 232). The responses, that is, are not operants, not (readily) modified by their consequences. Compare this with Weisman’s (1977) “Stubbornly we have resisted the conclusions that (a) much animal behavior is innate, and (b) animals have often not learned the correlation between their behavior and its adaptive consequence” (p. 7). Or they may have learned it and then forgot it. Weisman goes on to exemplify these points with the pro-forma behavior of cats covering scat. Hoarding behavior is equally resistant to changes in its outcomes (Luo et al., 2014).

Movements that begin as operant actions can become automatized into more reflexive responses: The correlation between responding and reinforcement may decrease to zero. “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). There are two ways in which this ritualization may happen: variation in the quality or availability of the reinforcer may have no effect on the quantity or nature of the responding (Dickinson & Balleine (1994) cite this as the *goal criterion* for calling a movement a goal-directed action); or reinforcement may not differentially select variations in the quantity or nature of responding (they call this the *instrumental*

criterion for goal directed action). In either case the correlation between them goes toward zero. In a series of studies in which the reinforcer has been devalued by satiation or a by conditioned aversion to it, animals may persist through the action chain to those reinforcers – only to recognize at the end that they were not worth the candle. This failure of changes in incentive motivation to modulate behavior occurs primarily when the response is over-trained (Adams & Dickinson, 1981; Balleine, Garner, Gonzalez, & Dickinson, 1995; Dickinson, Balleine, Watt, Gonzalez, & Boakes, 1995). In like manner addicts will often come to want their drug even if they do not particularly like it (Berridge & Robinson, 2003; Robinson & Berridge, 2013); Dickinson suggests that this may be a manifestation of the operation of habits that are no longer goal-directed actions (see, e.g., Hogarth, Dickinson, & Duka, 2010). We would suggest that they are goal-directed actions, but the goals are no longer relevant to the current motivational state. They are instrumental but not operant. Once stimuli have energized one of the modes in Figure 1, the modules and 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 original motivations that once made that a relevant goal (Dezfouli, Lingawi, & Balleine, 2014).

Stimulus Control vs. Reinforcement Control

Responses may be elicited by a stimulus but fail the operant criterion for an action if variants of it are not under the control of the reinforcer. One of the canonical laboratory behaviors, the key-pecking of pigeons, has been revealed to be easily elicited as a response to a sign stimulus – a key light signaling food – and only secondarily an operant under control of its consequences (Brown & Jenkins, 1968; Hearst & Jenkins, 1974; Sanabria, Sitomer, & Killeen, 2006). The balance between habit and action depends on the strength of the Pavlovian conditioning of the sign – the CS. It is often difficult to separate the operant and respondent support of a movement because wherever one set of operations (e.g., pairing of stimuli) exists, the other (e.g., the pairing of elicited responses with reinforcement) also exists (see the valuable discussions of these issues in Dickinson & Boakes, 1979; Hearst, 1975). In experiments designed to tease apart these causal factors, Killeen and associates (Killeen, Sanabria, & Dolgov, 2009) studied the effect of four types of trials in a probabilistic signaled food (positive auto maintenance) paradigm: trials with a peck (P) and food (F); those with no peck (\sim P) and F; trials with P and \sim F; and those with \sim P and \sim F. They predicted the number of responses on the next trial from a weighted regression of those predictors on the current trial. To measure such small trial-by-trial incremental effects with some reliability required on the order of ten thousand trials, which the pigeons obediently provided.

In most of the various conditions, the dominant weight was on the occurrence of food (F), and somewhat less on the occurrence of a response (P). The weight of the interaction term, response and food on the same trial (PF), an instance of reinforcement, increased the probability of a response on the next by an additional average of 6%, about half the weight accorded to F alone and to P alone. (The weight on PF was not significantly greater than 0 for 2 of the 6 pigeons, and averaged 8% for the remaining 4.) This weight was largest in experiments where the Pavlovian conditioning was weakest (viz., short ITIs and low probabilities of food). Thus, the Pavlovian factor of stimulus-contingent presentation of food dominated, and the momentum of responding (P) came in next. Still, the PF contingency played a role in some of the birds. Extinction trials (both CS extinction, \sim F, and response extinction, P \sim F) had weights close to 0 – reinforcing other reports of the often-negligible weighting of null cases (e.g., Huys et al., 2011; Zentall & Stagner, 2011).

The PF trials evidently served to help keep the animals in the pecking module, but had no effect on rate of responding in that module. The rate parameter in the response rate model did not change over conditions, unlike the weights for entering the pecking module, which changed systematically. A P was a strong predictor of making at least one response on the next trial, and a PF trial was in many cases better; but on those trials on which at least one peck occurred, the correlation of the response rate on a reinforced trial (viz., a PF trial) with the probability of responding on the next was essentially zero ($r = +0.035$; compare to $r = +0.081$ between response rate on an extinction trial (P \sim F) and probability of responding on the next trial). Nor was a relation found between the delay from a peck to reinforcement, and the probability of responding on the next trial. It is

clear then that a PF contingency can help keep the animal in a peck module on the next trial while having no effect on shaping the vigor of responding in that module. More responses followed by food, or followed more closely by food, was non-predictive of ensuing behavior. Pecking had become a habit.

Such automatization is consistent with Shull's (2011) observation that "Responding of various types occurs as bouts of activity separated by pauses. Deprivation and reinforcement affect the tendency to initiate and persist engaging in bouts but have little effect on performance once engaged." (p. 49). Experimental contingencies that discourage it may merely move it away from the recording instrument (F. J. Silva & Pear, 1995). This phenomenon of habit formation may be the same thing that Terrace identified as *output chunking* (Terrace, 2001).

Brembs calls the process of operant conditioning *self-learning* (in contradistinction to classical conditioning, which is *world* [or sign]-*learning*). But once well learned, responses can become automatized, as the above experiment among others show. The animal has learned what to do when, but just how to do it has become ossified. This transition from flexible to rigid learning is supported by well-defined neural (Brembs, 2009; Dezfouli et al., 2014; Everitt & Robbins, 2005; Graybiel, 2008; Killcross & Coutureau, 2003) and genetic (Brembs et al., 2011) mechanisms.

In contrast to the insensitivity of response bout length to feedback shown by Killeen and associates (2009), in a positive control condition that delivered food after 3 pecks on 10% of the trials, for the first time in the experiment the PF weight exceeded the weight on F alone. This result validates Dickinson's (1985) hypothesis that when variation in characteristics of a response is not correlated with variation in characteristics of reinforcement, that response becomes habitual and under weak or negligible control by the outcome. It is no longer an operant. Nonetheless, reinforcement may play a role in causing the animal to continue to engage that module, in concert with the often larger role played by the sign stimulus. A similar partnership between classical and neoclassical conditioning was noted by Timberlake and associates (Timberlake et al., 1982), along with many others (see, e.g., Davis & Hurwitz, 1977).

Premack (1965, p. 129) amplified his description of the little motors that constitute many action patterns, and modern research has validated those observations: 1. They occur at fixed rates (or harmonics of that rate: Killeen, Hall, Reilly, & Kettle, 2002; Kirkpatrick, 2002). 2. The duration of individual bursts of those responses is to a large extent unpredictable (they tend to be exponentially distributed: Shull, Gaynor, & Grimes, 2001). 3. The action pattern engaged depends on the actions previously engaged (Innis, Simmelhag-Grant, & Staddon, 1983; Killeen et al, 2009), and, 4. Changes in rates of measured responses are produced by affecting the pauses between successive bursts of responses – by entering the state more or less often (Brackney, Cheung, Neisewander, & Sanabria, 2011; Pear & Rector, 1979; Shull, 2011). What reinforcement does is modulate steps 3 and 4 – sign conditioning (when operative) controls the initiation of particular motors, and operant conditioning the location and timing of engagement and disengagement of those motors, and of others that may compete with them. Dezfouli and Balleine (2012) provide a recent update on the neurobiological underpinnings, sketch a hierarchical control model, and machine-learning algorithms to implement it.

Are Conditioned Stimuli Reinforcers?

Discriminative stimuli are way-signs, pointing to the next module of behavior that is lower on the action chain. For this reason, Shahan (2010) calls them signposts, and moots the issue of whether they de facto function as conditioned reinforcers. His resolution of the question is balanced, as fifty years of experiments and controversy provide evidence for and against the proposition that discriminative stimuli are conditioned reinforcers. Shahan nonetheless concludes that "A signposts or means-to-an-end account ... appears to provide a plausible alternative interpretation of [conditioned reinforcement effects]" (p. 269).

Rational animals should clearly approach such signs, which means they should work to observe them. Typically animals will, and this has been construed as evidence for the stimuli's reinforcing strength. But it is entry to the lower link of the action chain that is approached: Once the animal is there, the stimuli become supererogatory. When it is clear that stimuli point in the wrong direction, as when in an observing response experiment an S- results from an observation, those informative stimuli are avoided (Killeen, Wald, & Cheney, 1980). This is because they signal a greater distance from the goal than the status quo ante of the ambiguous stimulus. Conversely, when formerly disdained stimuli come signal access to actions that have become desirable, they are suddenly wanted (Robinson & Berridge, 2013). The early links in a chain schedule do not reinforce unless they move the animal discriminably closer to potent actions, conditioned reinforcers added to the terminal link of a concurrent chain schedule do not add value, and animals will turn their backs on an S-stimulus to return to ambiguity (Wasserman, Franklin, & Hearst, 1974). To say that because they are signposts they cannot strengthen the behavior that leads to them (Davison & Baum, 2006) is not quite correct, as there are decades of evidence showing that they can. But they do so not because they are goods in their own right, but because they point to actions that are either goods in their own right, or are closer to such actions; as was in fact the conclusion of Davison and Baum. It is not what stimuli are, they do not acquire value in themselves; it is what they let the animals do, that matters. If animals cease wanting to do those things, the signs cease to matter, or become aversive.

Stimuli that signal proximity – but not immanence – of a more potent link may come to elicit adjunctive or superstitious actions during the ensuing delay. Stimuli then associated with those delay responses may thereby become attractive, and able to act as conditioned reinforcers. In an interesting recent experiment, Thrailkill and Shahan (2014) trained rats on delayed and trace conditioning procedures. Then for one session they presented a novel stimulus after the original CS. The novel stimulus coming after the trace CS was better able to reinforce new responses than the one coming after the delay CS. During training the responses coming after the delay CS were immediate hopper approach mediated by hopper activation, and those were not available absent hopper operation in the post-conditioning session, but the trace conditioned responses were available. This provides strong evidence that it can be behavior—the arbitrary actions in the trace gap—that are potent in conditioning new stimuli, that such conditioning does not require the presence of the primary reinforcer, and that it is yet is able to reinforce instrumental actions leading to it.

Response Patterns as Reinforcers

In original experiments Neuringer and Chung (1967) presented blackouts to animals on short FI or FR schedules superimposed on a VI schedule of food. They found that if and only if the same schedule that presented blackout also presented food, then the responding took on a familiar pattern (pause and run) and overall rates more than doubled. This occurred even though the blackout was never paired with reinforcement. But if the food were delivered after a peck but independently of the blackout schedule, there was no patterning or increase in response rates. The authors attributed these effects to establishing the blackout as a “quasi-reinforcer”. In a subsequent series of similar experiments, Stubbs (1971) replicated their results, and also showed that the pairing of the brief stimuli caused no change in performance. He also showed that the patterning occurred even if the stimulus presentation were response independent, suggesting that the brief stimuli instructed the animals that a new interlude had begun, which might at the end eventuate in reinforcement. The brief stimuli acted as signposts, but were not scheduled as reinforcers and did not enhance response rates. The performance that they sustained did not look different from the pattern sustained when the brief stimuli were response dependent. Marr (1979) reviews these and more recent experiments, investigating the conditions that can generate unitary response sequences.

The powerful effects found by these investigators did not depend on Pavlovian pairing of the sign with primary reinforcement. In fact in the control condition with independent presentations, the brief stimuli were

typically closer to food than they were in the much more effective experimental procedure. This is not to say that pairing is always without effect. Stubbs reviewed some of the diverse literature where response-contingent presentations of paired stimuli augment responding. Notable is the work of Zimmerman and colleagues (e.g., Zimmerman, 1969), who demonstrated that brief hopper flashes maintain responding in the signaled absence of primary reinforcement and through extensive extinction sessions. Such stimuli elicit hopper-attending and approach responses – their location close to the site of food may make them more potent than key-light flashes. Williams (1994) and Shahan (2010) bring these considerations to and into the present century, but do not change the conclusions: Pavlovian pairing is not necessary for what appears to be conditioned reinforcement, and what appears to be conditioned reinforcement is often the signaled release of action patterns lower on the action chain.

Modules in the action chain may involve responses that the experimenter is not monitoring, and that are not necessary to move the animal to the next link. The literature is replete with examples of such adjunctive, displacement, interim, superstitious, and conditioned responses. These responses may have their origins as part of the natural repertoire of appetitive behavior (Segal, 1972), but can often be shaped by the reinforcement of progress through the chain (see, e.g., Killeen & Pellón, 2013). When evolved in these situations, those responses may acquire the ability to reinforce other instrumental acts. Falk (1966) showed that the ability to drink (excessive) amounts of water on a VI schedule of food delivery would maintain Fixed Ratio (FR) responding of 20 to 50 responses for dipper operation. Interestingly for our account, he noted that

To state that it is the water which is reinforcing in this situation [inappropriately] locates the property of reinforcement in the stimulus rather than as a relation between independent response rates (Premack, 1959)... Azrin was able to maintain FR behavior on a second key during food S- periods, the reinforcing event being the provision of another pigeon which the first bird then attacked... Phenomena [such] as polydipsia and aggression, which are produced by the environmental controls imposed by certain behavior schedules, are prepotent enough to sustain scheduled behavior in their own right. (p. 24)

Those responses may not only reinforce one another, they may compete with one another. Invoking his *appetitive structure* hypothesis, Timberlake and associates (Timberlake et al., 1982) noted that “the variation in expression of appetitive behaviors depended on the nature of the contingency and the resultant competition for expression with other behaviors. ... Misbehavior may not be a peculiar aberration but a particularly clear expression of the species-characteristic organization of stimuli and responses that probably underlies most learning.” (p. 84). That competition for expression may manifest as the roiling dynamics of interim responses (Reid, Bachá, & Morán, 1993; Reid, Vazquez, & Rico, 1985), perturbations in the stream of operant responses called behavioral contrast (Killeen, 2014), and the temporal displacement of response classes within the epoch as a function of magnitude of their memorial gradients (Killeen & Pellón, 2013) and differential rates of habituation (Balsam & Silver, 1994, p. 337).

Donahoe and colleagues have repeatedly emphasized that the control of operant behavior by outcome is mediated by the stimuli present at the time of reinforcement (Donahoe, 2006; Donahoe, Burgos, & Palmer, 1993; Donahoe, Palmer, & Burgos, 1997a, 1997b). They have also observed that classical and neoclassical procedures both select environment-behavior relations but, “because of the differences in the events that reliably occur in the vicinity of the reinforcers, the constituents of the selected relations are different” (Donahoe & Palmer, 1994, p. 65). Those events are proximities between actions and reinforcing outcomes that may associate the actions to the sign stimuli, making them responses; may through practice consolidate the responses as habits, and in that process may make them reinforcing in their own right. When this happens, then like the signs that signal available links down the action chain, those responses will be approached (Bindra, 1972). A paw above a lever is a sign of food in the hopper, and it is approached and enacted again

and again. The latency of approaching such responses depends on their own proximity to more satisfying states of affairs (Shull, 1979).

Premack's Humans

The Premackian conditioning considered above tends to be isolated to learning within subsystems and modes. There is little generalization of ability from one domain to another. Ever since Darwin humans have speculated about what makes our cognition so much more powerful from that of other animals (Boakes, 1984; Wynne, 2001). Human's ability to fabricate Premack's "little motors", to tune them and reapply them, is one of the key distinctions between man and beast. "Human faculties seem to consist not of a solution to one problem, but of the knitting together of the solutions to a number of problems. Although each of the individual components solves a specific problem, their combination provides the solution to a general problem. ... Human intelligence may consist of levels in which the primitives of one level are the combinations of a lower level" (Premack, 2010, pp. 30, 31). Language in both its metaphorical capability and its instructional role is key to this ability. I could ask you to engage in any of the actions in Figure 1, and if you obliged, you could at least "go through the motions."

Whence the sign stimuli for these actions? Conscious imagination. Through imagery, both visual and physical (Killeen & Glenberg, 2010), we release the actions, and then through practice, perfect them. Once they are habitual, we turn our minds to other employment, for "It is a general principle in Psychology that consciousness deserts all processes where it can no longer be of use. ... We grow unconscious of every feeling which is useless as a sign to lead us to our ends, and where one sign will suffice others drop out, and that one remains, to work alone" (James, 1890/1983, p. 497).

Summary

Figure 1 provides a static ethogram of an appetitive hierarchy. The Premack principle adds dynamics—transition to modules lower on that chain will reinforce actions higher on that chain. The actions in Figure 1 are natural parts of an organism's repertoire. What the conditioning process does is embellish and direct the responses in each module, and sometimes provide short cuts down the chain. If we envisage a multidimensional space of time and action, a learned sequence constitutes a trajectory through that space. One rendering of the Premack principle is that variants of an action and action-sequence that shorten the trajectory are reinforced and learned (Killeen, 1992). But the constituent actions in the chain often are, or become, attractors in their own right, leading to misbehaviors and less-than-optimal foraging. When the instrumental action becomes habitual, it may come to be preferred over the contingent behavior. This can lead to habits that are out of contact with their nominal goals, and to "running through" – instrumental responses that displace consummatory responses. It leads to the "touching the bases" of sign tracking that is unnecessary for, and often precludes the receipt of, "primary" reinforcement, or climbing over food to get to food (Stolz & Lott, 1964). This will often discomfort the experimenter more than the experimentees, who as molar behaving animals often seem to prefer sipping their wine to chugging it.

The process of conditioning may transform actions into habits, and make them reinforcing in the process: Habit is motive in the making. But habits will not motivate behavior by themselves – they need to be situated in the behavior system by stimuli or other actions that engage their modes. The ability of stimuli to do this is called incentive motivation. One role of occasion setters (Holland, 1992) is to tell animals when a subsequent sign will release action that will be effective – they are the signs of the operative modules in the behavior system. Their properties in extinction and delay of reinforcement are quite different from those of signs in the action chain (Holland, 1983, 1992).

Both conditioned stimuli and discriminative stimuli signal an epoch in which descent to a lower action module is impending. The Pavlovian CS signals an open loop epoch: Nothing an animal can do will move it through that module any faster. That does not keep the animal from responding. The Skinnerian S+ signals a closed loop; descent to the next action depends on instrumental responding. Aspects of responding that the experimenter shapes are reinforced by proximity to reinforcement, and compete against other responses either elicited in that module, or incidentally selected by the reinforcer. Classical conditioning procedures generate sign-knowledge; what stimuli will become signposts, and where they point. Operant conditioning procedures generate act-knowledge as well: What responses lead to desired outcomes. Both proceed together, with sign-learning in the early lead (Colomb & Brembs, 2010), and act-learning often facilitating sign-learning (Brembs & Heisenberg, 2000; Weiss, 2014). In all, a newly constructed action chain is forged, with Premackian conditioning signaling its links and carrying the animal to and through them.

Acknowledgments

Allen Neuringer, Björn Brembs and John Staddon provided invaluable feedback on earlier versions of this manuscript. Stan Weiss was midwife.

References

- Adams, C. D., & Dickinson, A. (1981). Instrumental responding following reinforcer devaluation. *The Quarterly Journal of Experimental Psychology*, *33*, 109-121. doi: 10.1080/14640748108400816
- Allison, J. (1983). *Behavioral economics*. New York, NY: Praeger.
- Allison, J., & Timberlake, W. (1974). Instrumental and contingent saccharin licking in rats: Response deprivation and reinforcement. *Learning and Motivation*, *5*, 231-247. doi: 10.1016/0023-9690(74)90029-0
- Balleine, B. W., Garner, C., Gonzalez, F., & Dickinson, A. (1995). Motivational control of heterogeneous instrumental chains. *Journal of Experimental Psychology: Animal Behavior Processes*, *21*, 203-217. doi: 10.1037/0097-7403.21.3.203
- Balsam, P. D., Deich, J. D., Ohyama, T., & Stokes, P. D. (1998). Origins of new behavior. In W. T. O'Donohue (Ed.), *Learning and behavior therapy* (pp. 403-420). Needham Heights, MA: Allyn & Bacon.
- Balsam, P. D., Graf, J. S., & Silver, R. (1992). Operant and Pavlovian contributions to the ontogeny of pecking in ring doves. *Developmental Psychobiology*, *25*, 389-410. doi: 10.1002/dev.420250602
- Balsam, P. D., & Silver, R. (1994). Behavioral change as a result of experience: Toward principles of learning and development. In J. A. Hogan & J. J. Bolhuis (Eds.), *Causal mechanisms of behavioural development* (pp. 327-357). Cambridge, England: Cambridge University Press.
- Baum, W. M. (2004). *Understanding behaviorism: Behavior, culture, and evolution*. New York, NY: Wiley-Blackwell.
- Baum, W. M. (2005). *Understanding behaviorism: Behavior, culture, and evolution* (2nd ed.). Malden, MA: Blackwell Publishing.
- Baum, W. M. (2012). Rethinking reinforcement: allocation, induction, and contingency. *Journal of the Experimental Analysis of Behavior*, *97*, 101-124. doi: 10.1901/jeab.2012.97-101
- Berridge, K. C., & Robinson, T. E. (2003). Parsing reward. *Trends in Neurosciences*, *26*, 507-513. doi: 10.1016/s0166-2236(03)00233-9

- Bindra, D. (1972). A unified account of classical conditioning and operant training. In A. H. Black & W. F. Prokasy (Eds.), *Classical conditioning II: Current research and theory* (pp. 453-481). New York, NY: Appleton-Century-Crofts.
- Boakes, R. A. (1977). Performance on learning to associate a stimulus with positive reinforcement. In H. David & H. M. B. Hurwitz (Eds.), *Operant-Pavlovian interactions* (pp. 67-101). Hillsdale, NJ: Erlbaum.
- Boakes, R. A. (1984). *From Darwin to Behaviourism: Psychology and the minds of animals*. Cambridge, UK: Cambridge University Press.
- Boakes, R. A., Poli, M., Lockwood, M. J., & Goodall, G. (1978). A study of misbehavior: Token reinforcement in the rat. *Journal of the Experimental Analysis of Behavior*, 29, 115-134. doi: 10.1901/jeab.1978.29-115
- Bolles, R. C. (1970). Species-specific defense reactions and avoidance learning. *Psychological Review*, 71, 32-48. doi: 10.1037/h0028589
- Bolles, R. C. (1983). The explanation of behavior. *The Psychological Record*, 33, 31-48.
- Brackney, R. J., Cheung, T. H., Neisewander, J. L., & Sanabria, F. (2011). The isolation of motivational, motoric, and schedule effects on operant performance: a modeling approach. *Journal of the Experimental Analysis of Behavior*, 96, 17-38. doi: 0.1901/jeab.2011
- Breland, K., & Breland, M. (1961). The misbehavior of organisms. *American Psychologist*, 16, 681-684. doi: 10.1037/h0040090
- Brembs, B. (2009). Mushroom Bodies Regulate Habit Formation in *Drosophila*. *Current Biology*, 19(16), 1351-1355. doi: 10.1016/j.cub.2009.06.014
- Brembs, B., & Heisenberg, M. (2000). The operant and the classical in conditioned orientation of *Drosophila melanogaster* at the flight simulator. *Learning & Memory*, 7, 104-115. doi: PMC311324
- Brembs, B., Pauly, D., Schade, R., Mendoza, E., Pflüger, H.-J., Rybak, J., Scharff, C., & Zars, T. (2011). *The Drosophila FoxP gene is necessary for operant self-learning: Implications for the evolutionary origins of language*. Paper presented at the German Neuroscience Society, Göttingen.
- Brown, P. L., & Jenkins, H.M. (1968). Auto-shaping of the pigeon's key-peck. *Journal of the Experimental Analysis of Behavior*, 11, 1-8. doi: 10.1901/jeab.1968.11-1
- Cairns, R. B., Garipey, J. L., & Hood, K. E. (1990). Development, microevolution, and social behavior. *Psychological Review*, 97, 49-65. doi: 10.1037/0033-295X.97.1.49
- Catania, A. C. (2013). *Learning* (5th ed.). Cornwall on Hudson, NY: Sloan Publishing.
- Colomb, J., & Brembs, B. (2010). The biology of psychology: 'Simple' conditioning? *Communicative & Integrative Biology*, 3, 142-145. doi: PMC2889970
- Crawford, L. L., Holloway, K. S., & Domjan, M. (1993). The nature of sexual reinforcement. *Journal of the Experimental Analysis of Behavior*, 60, 55-66. doi: 10.1901/jeab.1993.60-55
- Davis, H., & Hurwitz, H. M. B. (1977). *Operant-Pavlovian interactions*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Davison, M., & Baum, W. M. (2006). Do conditional reinforcers count? *Journal of the Experimental Analysis of Behavior*, 86, 269-283. doi: 10.1901/jeab.2006.56-05
- Denny, M., & Weisman, R. G. (1964). Avoidance behavior as a function of length of nonshock confinement. *Journal of Comparative and Physiological Psychology*, 58, 252-257. doi: 10.1037/h0048966
- Denny, M. R. (1991). Relaxation/relief: The effects of removing, postponing, or terminating aversive stimuli. In M. R. Denny (Ed.), *Fear, avoidance and phobias* (pp. 199-229). Hillsdale, NJ: Erlbaum.
- Denny, M. R., & Adelman, H. M. (1955). Elicitation theory: I. An analysis of two typical learning situations. *Psychological Review*, 62, 290-296. doi: 10.1037/h0046265

- Dezfouli, A., & Balleine, B. W. (2012). Habits, action sequences and reinforcement learning. *European Journal of Neuroscience*, 35, 1036-1051. doi: 10.1111/j.1460-9568.2012.08050.x
- Dezfouli, A., Lingawi, N. W., & Balleine, B. W. (2014). Habits as action sequences: hierarchical action control and changes in outcome value. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 369(1655). doi: 10.1098/rstb.2013.0482
- Dickinson, A. (1985). Actions and habits: the development of behavioural autonomy. *Philosophical Transactions of the Royal Society of London. B, Biological Sciences*, 308, 67-78. doi: 10.1098/rstb.1985.0010
- Dickinson, A., & Balleine, B. W. (1994). Motivational control of goal-directed action. *Animal Learning & Behavior*, 22, 1-18.
- Dickinson, A., Balleine, B. W., Watt, A., Gonzalez, F., & Boakes, R. A. (1995). Motivational control after extended instrumental training. *Animal Learning & Behavior*, 23, 197-206.
- Dickinson, A., & Boakes, R. A. (Eds.). (1979). *Mechanisms of learning and motivation: A memorial volume to Jerzy Konorski*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Domjan, M. (2005). Pavlovian conditioning: A functional perspective. *Annual Review of Psychology*, 56, 179-206. doi: 10.1146/annurev.psych.55.090902.141409
- Domjan, M. (2008). Adaptive specializations and generality of the laws of classical and instrumental conditioning. In R. Menzel (Ed.), *Learning theory and behavior* (Vol. 1, pp. 327-340). Oxford, UK: Elsevier.
- Donahoe, J. W. (2006). Contingency: Its meaning in the Experimental Analysis of Behavior. *European Journal of Behavior Analysis*, 7, 111-114.
- Donahoe, J. W. (2014). Evocation of behavioral change by the reinforcer is the critical event in both the classical and operant procedures. *International Journal of Comparative Psychology*, 27, 537-543.
- Donahoe, J. W., Burgos, J. E., & Palmer, D. C. (1993). A selectionist approach to reinforcement. *Journal of the Experimental Analysis of Behavior*, 60, 17-40. doi: 10.1901/jeab.1993.60-17
- Donahoe, J. W., & Palmer, D. C. (1994). *Learning and complex behavior*. Boston, MA: Allyn and Bacon.
- Donahoe, J. W., Palmer, D. C., & Burgos, J. E. (1997a). The S-R issue: its status in behavior analysis and in Donahoe and Palmer's learning and complex behavior. *Journal of the Experimental Analysis of Behavior*, 67, 193-211. doi: 10.1901/jeab.1997.67-193
- Donahoe, J. W., Palmer, D. C., & Burgos, J. E. (1997b). The unit of selection: What do reinforcers reinforce? *Journal of the Experimental Analysis of Behavior*, 67, 259-273. doi: 10.1901/jeab.1997.67-259
- Donahoe, J. W., & Vegas, R. (2004). Pavlovian conditioning: The CS-UR relation. *Journal of Experimental Psychology: Animal Behavior Processes*, 30, 17-33. doi: 10.1037/0097-7403.30.1.17
- Eisenberger, R., Karpman, M., & Tractner, J. (1967). What is the necessary and sufficient condition for reinforcement in the contingency situation? *Journal of Experimental Psychology*, 74, 342-350. doi: 10.1037/h0024719
- Everitt, B. J., & Robbins, T. W. (2005). Neural systems of reinforcement for drug addiction: from actions to habits to compulsion. *Nature Neuroscience*, 8, 1481-1489. doi: 10.1038/nn1579
- Falk, J. L. (1966). The motivational properties of schedule-induced polydipsia. *Journal of the Experimental Analysis of Behavior*, 9, 19-25. doi: 10.1901/jeab.1966.9-19
- 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 R. C. Bolles & M. D. Beecher (Eds.), *Evolution and learning* (pp. 185-212). Hillsdale, NJ: Erlbaum.
- Fedorchak, P. M., & Bolles, R. C. (1986). Differential outcome effect using a biologically neutral outcome difference. *Journal of Experimental Psychology: Animal Behavior Processes*, 12, 125-130. doi: 10.1037/0097-7403.12.2.125
- Gallistel, C. (1980). *The organization of action: A new synthesis*. Hillsdale, NJ: Lawrence Erlbaum Associates.

- Gallistel, C. R. (1981). Précis of Gallistel's The organization of action: A new synthesis. *Behavioral and Brain Sciences*, 4, 609-619. doi: <http://dx.doi.org/10.1017/S0140525X00000480>
- Gormezano, I., & Kehoe, E. J. (1975). Classical conditioning: Some methodological-conceptual issues. In W. K. Estes (Ed.), *Handbook of learning and cognitive processes* (Vol. 2, pp. 143-179). Hillsdale, New Jersey: Lawrence Erlbaum Associates.
- Gottlieb, G. (2007). Probabilistic epigenesis. *Developmental Science*, 10, 1-11. doi: 10.1111/j.1467-7687.2007.00556.x
- Gould, J. L., & Marler, P. (1984). Ethology and the natural history of learning. In P. Marler & H. S. Terrace (Eds.), *The biology of learning* (pp. 47-74). New York, NY: Springer-Verlag.
- Graybiel, A. M. (2008). Habits, rituals, and the evaluative brain. *Annual Review of Neuroscience*, 31, 359-387. doi: 10.1146/annurev.neuro.29.051605.112851
- Hanson, S. J., & Timberlake, W. (1983). Regulation during Challenge: A general model of learned performance under schedule constraint. *Psychological Review*, 90, 261-282. doi: 10.1037/0033-295X.90.3.261
- Hearst, E. (1975). The classical-instrumental distinction: Reflexes, voluntary behavior, and categories of associative learning. In W. K. Estes (Ed.), *Handbook of learning and cognitive processes* (Vol. 2, pp. 181-223). Mahwah, NJ: Erlbaum.
- Hearst, E., & Jenkins, H. M. (1974). *Sign-tracking: The stimulus-reinforcer relation and directed action*. Austin, TX: The Psychonomic Society.
- Herrnstein, R. J. (1988). Lost and found: One self. *Ethics*, 98, 566-578. doi: <http://www.jstor.org/stable/2380969>
- Hinde, R. A. (1973). *Constraints on learning: An introduction to the problems*. New York, NY: Academic Press.
- Hinde, R. A. (1982). *Ethology: Its nature and relation with other sciences*. New York, NY: Oxford University Press.
- Hogan, J. A. (1988). Cause and function in the development of behavior systems. In E. M. Blass (Ed.), *Developmental psychobiology and behavioral ecology* (Vol. 9, pp. 63-106). New York, NY: Plenum Press.
- Hogan, J. A. (1994). Development of behavior systems. In J. A. Hogan & J. J. Bolhuis (Eds.), *Causal mechanisms of behavioural development* (pp. 242-264). Cambridge, England: Cambridge University Press.
- Hogan, J. A. (2014). A framework for the study of behavior. *Behavioural Processes*, May 22. pii: S0376-6357(14)00123-5. doi: 10.1016/j.beproc.2014.05.003. [Epub ahead of print]
- Hogan, J. A., & Bolhuis, J. J. (2009). *Causal mechanisms of behavioural development*. Cambridge, England: Cambridge University Press.
- Hogarth, L., Dickinson, A., & Duka, T. (2010). The associative basis of cue-elicited drug taking in humans. *Psychopharmacology*, 208, 337-351. doi: 10.1007/s00213-009-1735-9
- Holland, P. C. (1977). Conditioned stimulus as a determinant of the form of the Pavlovian conditioned response. *Journal of Experimental Psychology: Animal Behavior Processes*, 3, 77-104. doi: 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. *Journal of Experimental Psychology: Animal Behavior Processes*, 6, 155-174. doi: 10.1037/0097-7403.6.2.155
- Holland, P. C. (1983). Occasion-setting in Pavlovian feature positive discriminations. In M. L. Commons, R. J. Herrnstein, & A. R. Wagner (Eds.), *Quantitative analyses of behavior: Discrimination processes* (Vol. 4, pp. 183-206). New York: Ballinger.

- Holland, P. C. (1990). Forms of memory in Pavlovian conditioning. In J. L. McGaugh, N. M. Veinberger & G. Lynch (Eds.), *Brain organization and memory: cells, systems, and circuits* (pp. 78-105). Oxford, UK: Oxford University Press.
- Holland, P. C. (1992). Occasion setting in Pavlovian conditioning. In D. L. Medin (Ed.), *The psychology of learning and motivation* (Vol. 28, pp. 69-125). San Diego, CA: Academic Press.
- Hollis, K. L. (1983). Cause and function of animal learning processes. In P. Marler & H. S. Terrace (Eds.), *The biology of learning* (pp. 357-371). Berlin, Germany: Springer-Verlag.
- Hull, C. L. (1934). The concept of the habit-family hierarchy, and maze learning. Part I. *Psychological Review*, *41*, 33-54. doi: 10.1037/h0070758
- Hull, C. L. (1943). *Principles of behavior*. New York, NY: Appleton-Century-Crofts.
- Hundt, A. G., & Premack, D. (1963). Running as both a positive and negative reinforcer. *Science*, *142*, 1087-1088. doi: 10.1126/science.142.3595.1087
- Hursh, S. R. (1984). Behavioral economics. *Journal of the Experimental Analysis of Behavior*, *42*, 435-452.
- Huys, Q. J., Cools, R., Golzer, M., Friedel, E., Heinz, A., Dolan, R. J., & Dayan, P. (2011). Disentangling the roles of approach, activation and valence in instrumental and Pavlovian responding. *PLoS Computer and Biology*, *7*, e1002028. doi: 10.1371/journal.pcbi.1002028
- Innis, N. K., Simmelhag-Grant, V. L., & Staddon, J. E. R. (1983). Behavior induced by periodic food delivery: The effects of interfood interval. *Journal of the Experimental Analysis of Behavior*, *39*, 309-322. doi: 10.1901/jeab.1983.39-309
- James, W. (1890/1983). *Principles of Psychology*. Cambridge, MA: Harvard University Press.
- Jenkins, H. M., Barrera, F., Ireland, C., & Woodside, B. (1978). Signal-centered action patterns of dogs in appetitive classical conditioning. *Learning and Motivation*, *9*, 272-296. doi: 10.1016/0023-9690(78)90010-3
- Killcross, S., & Coutureau, E. (2003). Coordination of actions and habits in the medial prefrontal cortex of rats. *Cerebral Cortex*, *13*, 400-408. doi: 10.1093/cercor/13.4.400
- Killeen, P. R. (1989). Behavior as a trajectory through a field of attractors. In J. R. Brink & C. R. Haden (Eds.), *The Computer and the brain: Perspectives on human and artificial intelligence* (pp. 53-82). Amsterdam: Elsevier.
- Killeen, P. R. (1991). Behavioral Geodesics. In D. S. Levine & J. S. Levin (Eds.), *Motivation, emotion, and goal direction in neural networks* (pp. 91-114). Hillsdale, NJ: Erlbaum.
- Killeen, P. R. (1992). Mechanics of the animate. *Journal of the Experimental Analysis of Behavior*, *57*, 429-463. doi: 10.1901/jeab.1992.57-429
- Killeen, P. R. (2014). A theory of behavioral contrast. *Journal of the Experimental Analysis of Behavior*, *102*, 363-390. doi: 10.1002/jeab.107
- Killeen, P. R., & Glenberg, A. M. (2010). Resituating cognition. *Comparative Cognition & Behavior Reviews*, *5*, 59-77. doi: 10.3819/ccbr.2010.50003
- Killeen, P. R., Hall, S. S., Reilly, M. P., & Kettle, L. C. (2002). Molecular analyses of the principal components of response strength. *Journal of the Experimental Analysis of Behavior*, *78*, 127-160. doi: 10.1901/jeab.2002.78-127
- Killeen, P. R., & Pellón, R. (2013). Adjunctive behaviors are operants. *Learning & Behavior*, *41*, 1-24. doi: 10.3758/s13420-012-0095-1
- Killeen, P. R., Sanabria, F., & Dolgov, I. (2009). The dynamics of conditioning and extinction. *Journal of Experimental Psychology: Animal Behavior Processes*, *35*, 447-472. doi: 10.1037/a0015626
- Killeen, P. R., Wald, B., & Cheney, C. D. (1980). Observing behavior and information. *The Psychological Record*, *30*, 181-190.

- Kirkpatrick, K. (2002). Packet theory of conditioning and timing. *Behavioural Processes*, *57*, 89-106. doi: 10.1016/S0376-6357(02)00007-4
- Klatt, K. P., & Morris, E. K. (2001). The Premack principle, response deprivation, and establishing operations. *The Behavior Analyst*, *24*, 173-180. doi: PMC2731497
- Konorski, J. (1967). *Integrative activity of the brain*. Chicago, IL: Chicago University Press.
- Kruijt, J. P. (1964). Ontogeny of social behaviour in Burmese red junglefowl (*Gallus gallus spadiceus*) Bonnatere. *Behaviour. Supplement*, 1-201. doi: <http://www.jstor.org/stable/30039152>
- Lieberman, D. A., Davidson, F. H., & Thomas, G. V. (1985). Marking in pigeons: The role of memory in delayed reinforcement. *Journal of Experimental Psychology: Animal Behavior Processes*, *11*, 611-624. doi: 10.1037/0097-7403.11.4.611
- Lorenz, K. (1973). Konrad Lorenz - Biographical. *Nobelprize.org*. Retrieved 5 Oct 2014, 2014
- Lubow, R. E. (1989). *Latent inhibition and conditioned attention theory*. Cambridge, England: Cambridge University Press.
- Luo, Y., Yang, Z., Steele, M. A., Zhang, Z., Stratford, J. A., & Zhang, H. (2014). Hoarding without reward: Rodent responses to repeated episodes of complete cache loss. *Behavioural Processes*, *106*, 36-43. doi: <http://dx.doi.org/10.1016/j.beproc.2014.04.005>
- Mackintosh, N. J. (1974). *The psychology of animal learning*. New York, NY: Academic Press.
- Marr, M. (1979). Second-order schedules and the generation of unitary response sequences. In M. D. Zeiler & P. Harzem (Eds.), *Advances in the analysis of behavior* (Vol. 1, pp. 223-260). New York, NY: John Wiley & Sons.
- Mason, G., Bateson, M., & Bean, D. (1999). Contra-freeloading in starlings-a test of the information hypothesis. *Behaviour*, *136*, 1267-1282. doi: <http://hdl.handle.net/10214/4700>
- Miyashita, Y., Nakajima, S., & Imada, H. (2000). Differential outcome effect in the horse. *Journal of the Experimental Analysis of Behavior*, *74*, 245-253. doi: 10.1901/jeab.2000.74-245
- Morgan, M. (1974). Resistance to satiation. *Animal Behaviour*, *22*, 449-466. doi: 10.1016/S0003-3472(74)80044-8
- Neuringer, A. J. (1970a). Many responses per food reward with free food present. *Science*, *169*, 503-504. doi: 10.1126/science.169.3944.503
- Neuringer, A. J. (1970b). Superstitious key pecking after three peck-produced reinforcements. *Journal of the Experimental Analysis of Behavior*, *13*, 127-134. doi: 10.1901/jeab.1970.13-127
- Neuringer, A. J. (2002). Operant variability: Evidence, functions, and theory. *Psychonomic Bulletin & Review*, *9*, 672-705. doi: 10.3758/BF03196324
- Neuringer, A. J., & Chung, S.-H. (1967). Quasi-reinforcement: Control of responding by a percentage-reinforcement schedule. *Journal of the Experimental Analysis of Behavior*, *10*, 45-54. doi: 10.1901/jeab.1967.10-45
- Osborne, S. R. (1978). A note on the acquisition of responding for food in the presence of free food. *Animal Learning & Behavior*, *6*, 368-369.
- Patterson, A. E., & Boakes, R. A. (2012). Interval, blocking and marking effects during the development of schedule-induced drinking in rats. *Journal of Experimental Psychology: Animal Behavior Processes*, *38*, 303-314. doi: 10.1037/a0027788
- Pavlov, I. P. (1927). *Conditioned reflexes* (G. V. Anrep, Trans.). London, UK: Oxford University Press.
- Pear, J. J., & Rector, B. L. (1979). Constituents of response rate. *Journal of the Experimental Analysis of Behavior*, *32*, 341-362. doi: 10.1901/jeab.1979.32-341
- Peden, B. F., Browne, M. P., & Hearst, E. (1977). Persistent approaches to a signal for food despite food omission for approaching. *Journal of Experimental Psychology: Animal Behavior Processes*, *3*, 377-399. doi: 10.1037/0097-7403.3.4.377

- Pitts, R. C. (2013). On multiscaled and unified. *The Behavior Analyst*, 36, 313-323.
- Premack, D. (1959). Toward empirical behavior laws: I. Positive reinforcement. *Psychological Review*, 66, 219-233. doi: 10.1037/h0040891
- Premack, D. (1963). Rate differential reinforcement in monkey manipulation. *Journal of the Experimental Analysis of Behavior*, 6, 81-89. doi: 10.1901/jeab.1963.6-81
- Premack, D. (1965). Reinforcement theory. In D. Levine (Ed.), *Nebraska Symposium on Motivation* (pp. 123-180). Lincoln, NE: University of Nebraska Press.
- Premack, D. (1971). Catching up with common sense, or two sides of a generalization: Reinforcement and punishment. In R. Blaser (Ed.), *The nature of reinforcement* (pp. 121-150). New York, NY: Academic Press.
- Premack, D. (2010). Why humans are unique: Three theories. *Perspectives on Psychological Science*, 5(1), 22-32. doi: 10.1177/1745691609356782
- Pryor, K., Haag, R., & O'reilly, J. (1969). The creative porpoise: training for novel behavior. *Journal of the Experimental Analysis of Behavior*, 12, 653-661. doi: 10.1901/jeab.1969.12-653
- Rachlin, H. R. (1985). Pain and behavior. *Behavioral and Brain Sciences*, 8, 43-83. doi: <http://dx.doi.org/10.1017/S0140525X00019488>
- Rachlin, H. R. (1992). Teleological behaviorism. *American Psychologist*, 47, 1371-1382.
- Rachlin, H. R. (2014). *The escape of the mind*. New York, NY: Oxford University Press.
- Rachlin, H. R., Battalio, R., Kagel, J., & Green, L. (1981). Maximization theory in behavioral psychology. *Behavioral and Brain Sciences*, 4, 371-417. doi: <http://dx.doi.org/10.1017/S0140525X00009407>
- Reid, A. K., Bachá, G., & Morán, C. (1993). The temporal organization of behavior on periodic food schedules. *Journal of the Experimental Analysis of Behavior*, 59, 1-27. doi: 10.1901/jeab.1993.59-1
- Reid, A. K., Chadwick, C. Z., Dunham, M., & Miller, A. (2001). The development of functional response units: The role of demarcating stimuli. *Journal of the Experimental Analysis of Behavior*, 76, 303-320. doi: 10.1901/jeab.2001.76-303
- Reid, A. K., Rapport, H. F., & Le, T.-A. (2013). Why don't guiding cues always guide in behavior chains? *Learning & Behavior*, 41, 402-413. doi: 10.3758/s13420-013-0115-9
- Reid, A. K., Vazquez, P. P., & Rico, J. A. (1985). Schedule induction and the temporal distributions of adjunctive behavior on periodic water schedules. *Learning & Behavior*, 13, 321-326. doi: 10.3758/BF03200027
- Rescorla, R. A. (1971). Summation and retardation tests of latent inhibition. *Journal of Comparative and Physiological Psychology*, 75, 77-81. doi: 10.1037/h0030694
- Reynierse, J. H., & Rizley, R. C. (1970). Relaxation and fear as determinants of maintained avoidance in rats. *Journal of Comparative and Physiological Psychology*, 72, 223-232. doi: 10.1037/h0029476
- Robinson, M. J., & Berridge, K. C. (2013). Instant transformation of learned repulsion into motivational "wanting". *Current Biology*, 23, 282-289. doi: 10.1016/j.cub.2013.01.016
- Sanabria, F., Sitomer, M. T., & Killeen, P. R. (2006). Negative automaintenance omission training is effective. *Journal of the Experimental Analysis of Behavior*, 86, 1-10. doi: 10.1901/jeab.2006.36-05
- Sawisch, L. P., & Denny, M. R. (1973). Reversing the reinforcement contingencies of eating and keypecking behaviors. *Animal Learning & Behavior*, 1, 189-192.
- Schneirla, T. C. (1959). An evolutionary and developmental theory of biphasic processes underlying approach and withdrawal. In M. Jones (Ed.), *Nebraska Symposium on Motivation* (pp. 1-42). Lincoln, NE: University of Nebraska Press.
- Segal, E. F. (1972). Induction and the provenance of operants. In R. M. Gilbert & J. R. Millenson (Eds.), *Reinforcement: Behavioral analyses* (pp. 1-34). New York, NY: Academic Press.

- Shahan, T. A. (2010). Conditioned reinforcement and response strength. *Journal of the Experimental Analysis of Behavior*, *93*, 269-289. doi: 10.1901/jeab.2010.93-269
- Shettleworth, S. J. (1975). Reinforcement and the organization of behavior in golden hamsters: Hunger, environment, and food reinforcement. *Journal of Experimental Psychology: Animal Behavior Processes*, *1*, 56-87. doi: 10.1037/0097-7403.1.1.56
- Shettleworth, S. J. (2009a). *Cognition, evolution, and behavior*: Oxford University Press.
- Shettleworth, S. J. (2009b). The evolution of comparative cognition: is the snark still a boojum? *Behavioural Processes*, *80*, 210-217. doi: 10.1016/j.beproc.2008.09.001
- Shettleworth, S. J., & Juergensen, M. R. (1980). Reinforcement and the organization of behavior in golden hamsters: Brain stimulation reinforcement for seven action patterns. *Journal of Experimental Psychology: Animal Behavior Processes*, *6*, 352-375. doi: 10.1037/0097-7403.6.4.352
- Shimp, C. P. (2013). Toward the unification of molecular and molar analyses. *The Behavior Analyst*, *36*, 295-312.
- Shull, R. L. (1979). The postreinforcement pause. In M. D. Zeiler & P. Harzem (Eds.), *Advances in the analysis of behavior* (Vol. 1, pp. 193-221). New York, NY: John Wiley & Sons.
- Shull, R. L. (2011). Bouts, changeovers, and units of operant behavior. *European Journal of Behavior Analysis*, *12*, 49-72.
- Shull, R. L., Gaynor, S. T., & Grimes, J. A. (2001). Response rate viewed as engagement bouts: Effects of relative reinforcement and schedule type. *Journal of the Experimental Analysis of Behavior*, *75*, 247-274. doi: 10.1901/jeab.2001.75-247
- Silva, F. J., & Pear, J. J. (1995). Stereotypy of spatial movements during noncontingent and contingent reinforcement. *Animal Learning & Behavior*, *23*, 245-255.
- Silva, F. J., Timberlake, W., & Koehler, T. L. (1996). A behavior systems approach to bidirectional excitatory serial conditioning. *Learning and Motivation*, *27*, 130-150. doi: 10.1006/lmot.1996.0008
- Silva, K. M., & Timberlake, W. (1997). A behavior systems view of conditioned states during long and short CS-US intervals. *Learning and Motivation*, *28*, 465-490. doi: 10.1006/lmot.1997.0986
- Skinner, B. F. (1938). *The behavior of organisms*. New York: Appleton-Century-Crofts.
- Staddon, J. E. R. (1979). Operant behavior as adaptation to constraint. *Journal of Experimental Psychology: General*, *108*, 48-67. doi: 10.1037/0096-3445.108.1.48
- Stokes, P. D., & Balsam, P. D. (1991). Effects of reinforcing preselected approximations on the topography of the rat's bar press. *Journal of the Experimental Analysis of Behavior*, *55*, 213-231. doi: 10.1901/jeab.1991.55-213
- Stolz, S. B., & Lott, D. F. (1964). Establishment in rats of a persistent response producing a net loss of reinforcement. *Journal of Comparative and Physiological Psychology*, *57*, 147-149. doi: <http://dx.doi.org/10.1037/h0042991>
- Stubbs, D. A. (1971). Second-order schedules and the problem of conditioned reinforcement. *Journal of the Experimental Analysis of Behavior*, *16*, 289-313. doi: 10.1901/jeab.1971.16-289
- Terhune, J. G., & Premack, D. (1970). On the proportionality between the probability of not-running and the punishment effect of being forced to run. *Learning and Motivation*, *1*, 141-149. doi: 10.1016/0023-9690(70)90080-9
- Terrace, H. S. (2001). Chunking and serially organized behavior in pigeons, monkeys and humans. In R. G. Cook (Ed.), *Avian visual cognition* [On-line]. Available: <http://www.pigeon.psy.tufts.edu/avc/terrace/>.
- Thrailkill, E. A., & Shahan, T. A. (2014). Temporal integration and instrumental conditioned reinforcement. *Learning & Behavior*, *42*, 1-8. doi: 10.3758/s13420-014-0138-x
- Timberlake, W. (1983). The functional organization of appetitive behavior: Behavior systems and learning. In M. D. Zeiler & P. Harzem (Eds.), *Advances in Analysis of Behavior* (Vol. 3. Biological factors in learning, pp. 177-221). New York, NY: Wiley.

- Timberlake, W. (1993). Behavior systems and reinforcement: An integrative approach. *Journal of the Experimental Analysis of Behavior*, *60*, 105-128. doi: 10.1901/jeab.1993.60-105
- Timberlake, W. (1994). Behavior systems, associationism, and Pavlovian conditioning. *Psychonomic Bulletin & Review*, *1*, 405-420. doi: 10.3758/BF03210945
- Timberlake, W. (2001). Motivational modes in behavior systems. In R. R. Mowrer & S. B. Klein (Eds.), *Handbook of Contemporary Learning Theories* (pp. 155-209). Mahwah, NJ: Erlbaum.
- Timberlake, W., & Allison, J. (1974). Response deprivation: An empirical approach to instrumental performance. *Psychological Review*, *81*, 146-164. doi: 10.1037/h0036101
- Timberlake, W., & Lucas, G. A. (1989). Behavior systems and learning: From misbehavior to general principles. In S. B. Klein & R. R. Mowrer (Eds.), *Contemporary Learning Theories: Instrumental Conditioning Theory and the Impact of Constraints on Learning* (pp. 237-275). Hillsdale, NJ: Erlbaum.
- Timberlake, W., Wahl, G., & King, D. A. (1982). Stimulus and response contingencies in the misbehavior of rats. *Journal of Experimental Psychology: Animal Behavior Processes*, *8*, 62-85. doi: 10.1037/0097-7403.8.1.62
- Timberlake, W., & White, W. (1990). Winning isn't everything: Rats need only food deprivation and not food reward to efficiently traverse a radial arm maze. *Learning and Motivation*, *21*, 153-163. doi: 10.1016/0023-9690(90)90017-I
- Timberlake, W., & Wozny, M. (1979). Reversibility of reinforcement between eating and running by schedule changes: A comparison of hypotheses and models. *Animal Learning & Behavior*, *7*, 461-469. doi: 10.3758/BF03209702
- Tinbergen, N. (1961). *The Herring Gull's World*. New York, NY: Basic Books.
- Trapold, M. A. (1970). Are expectancies based upon different positive reinforcing events discriminably different? *Learning and Motivation*, *1*, 129-140. doi: 10.1016/0023-9690(70)90079-2
- Turkkan, J. S. (1989). Classical conditioning beyond the reflex: The new hegemony. *Behavioral and Brain Sciences*, *12*, 121-137. doi: <http://dx.doi.org/10.1017/S0140525X00024572>
- Wasserman, E. A., Franklin, S. R., & Hearst, E. (1974). Pavlovian contingencies and approach versus withdrawal to conditioned stimuli in pigeons. *Journal of Comparative and Physiological Psychology*, *86*, 616-627. doi: 10.1037/h0036171
- Weisman, R. G. (1977). On the role of the reinforcer in associative learning. In H. Davis & H. M. B. Hurwitz (Eds.), *Operant-Pavlovian interactions*. Hillsdale, NJ: Erlbaum.
- Weiss, S. J. (2014). Instrumental and classical conditioning. In F. K. McSweeney & E. S. Murphy (Eds.), *The Wiley Blackwell Handbook of Operant and Classical Conditioning* (pp. 417-451). John Wiley & Sons, Ltd, Oxford, UK. doi: 10.1002/9781118468135.ch17.
- Williams, B. A. (1994). Conditioned reinforcement: Neglected or outmoded explanatory construct? *Psychonomic Bulletin & Review*, *1*, 457-475. doi: 10.3758/BF03210950
- Wynne, C. D. (2001). *Animal cognition: The mental lives of animals*. New York, NY: Macmillan.
- 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). *Proceedings of the Royal Society, B: Biological Sciences*, *278*, 1203-1208. doi: 10.1098/rspb.2010.1607
- Zimmerman, J. (1969). Meanwhile... back at the key: Maintenance of behavior by conditioned reinforcement and response-independent primary reinforcement. In D. P. Hendry (Ed.), *Conditioned reinforcement* (pp. 91-124). Homewood, IL: The Dorsey Press.

Financial Support: The author received support from his wife.

Conflict of Interest: The author declares no conflict of interest.

Submitted: July 14th, 2014

Resubmitted: October 11th, 2014

Accepted: October 30st, 2014

Glossary

Movement: A change of physical location or position.

Action: The process of doing something, typically to achieve an aim; a goal-directed movement.

Response: A movement in reaction to a stimulus. An instance of behavior.

Behavior: One of many ways in which an animal might respond to particular situations (contexts and stimuli). It is a movement caused by the organism; responding.

Habit: A regular, repeated behavior, especially one that is hard to extinguish.

Reinforcement: 1. A change in characteristics (e.g., probability, rate, resistance to extinction, etc.) of actions due to their proximity to a biologically potent event. 2. Maintenance of that changed characteristic.

Classical (Pavlovian, or respondent, or sign) conditioning: 1. The operation: The conditional probability of a biologically potent action (the UR) is greater soon after the sign (the CS) than at other times. 2. The effect: After conditioning, the organism exhibits conditional responses (CRs) to the CS.

Operant (Thorndikian or Skinnerian or instrumental or neoclassical or act) conditioning: 1. The operation: The conditional probability of a biologically potent action (such as the consummation of a reinforcer) is greater soon after the response than at other times. 2. The effect: After conditioning, the characteristics of the response change. Operant conditioning is what reinforcement does, and the conditional behavior is called an *operant*.

Premackian conditioning: 1. The operation: The conditional availability of a more potent (*contingent*) behavior is greater soon after the emission of a less potent (*instrumental*) behavior than at other times. 2. The effect: a) After conditioning, the organism exhibits a change in characteristics of the instrumental response (act-learning). b) Premackian conditioning may also establish stimuli as signals of that transition (sign-learning).

Behaviorally potent: An action capable of sustaining conditioning. 1. In behavioral parlance, a UR, consummatory response or escape response, and actions that lead to them. 2. In biological parlance, a biologically significant event.

Response characteristics: The form, location, latency, rate and probability of a response.

Sign-tracking: The tendency of animals to approach and contact signs of the availability of more potent positive actions, and avoid and remove signs of impending less positive responses.

These definitions start with the most general thing that we study, *movement*. Anything can have its position changed, and most animate things can change it themselves. Sunflowers follow the sun as tides follow the moon. It is a useful neutral term, one that can refer to the motion of a dog's paw without prejudging whether it was moved by its master's hand or its master's command. To add that the dog moved it makes it either an action or a response. If we cannot tell which, all we can say is "the paw moved".

"Actions [are] responses defined by their relation to the environment" (Catania, 2013, p. 9). In common parlance actions assume agency, and the actions of agents are typically *goal-directed*. They are *instrumental* actions, in that they "serve as an instrument or means in pursuing an aim or policy". Sign-tracking, the tendency to approach signs of more biologically potent events, is a ubiquitous instrumental response. If in addition actions are *modifiable* by the consequences that are obtained, then they are *operants*. Sometimes actions are done to achieve an aim, but are not modified by their consequences. They are instrumental actions, but not operants. We all know individuals who repeat the same counterproductive actions over and over. Those actions may be goal directed, but they are not under the control of the outcomes. If they are routine we call them *habits*.

Habits are not necessarily aimless. Runners may start running for health, but the movement of running may become its own goal, and might continue despite negative health consequences. Such habits remain instrumental, but the goal has shifted, sometimes imperceptibly. Conversely, movements do not have to be goal-directed to be operants; they need only be modified by their consequences. The judgment as to whether an action is instrumental is Olympian, as it arrogates to the observer what the goal should be. Auto-shaped key pecking achieves the goal of making contact with a sign of reinforcement, even though the act of doing so may have been arranged to decrease the eventual probability of reinforcement (Peden, Browne, & Hearst, 1977). We shake our heads, while in blissful ignorance of what he should be doing the pigeon keeps nodding his.

A *response* is by definition a reaction to something - the something that we call a stimulus. "Neither [stimulus or response] may be defined as to its essential properties without the other" (Skinner, 1938, p. 9). A spontaneous movement is a response whose stimulus we cannot discern (Bindra, 1972). Catania noted that "spond" or "sponse" might be more appropriate for actions that are not a response to an obvious stimulus (Catania, 2013).

A loud noise causes a startle response, also called a *reflex*. Unconditioned reflexes, as their name implies, do not require a particular history of experience, although they may be modified by such history, as a warning stimulus can modify the strength of the response to the noise stimulus. Classically conditioned responses are responses to classically conditioned stimuli—sounds or lights that predict the occurrence of a biologically-potent event (Baum's (2005) *phylogenetically important event*).