

## Recursive Self-Prediction as a Proximate Cause of Impulsivity The Value of a Bottom-Up Model

George Ainslie  
Veterans Affairs Medical Center, Coatesville, PA<sup>1</sup>  
University of Cape Town, South Africa  
[George.Ainslie@va.gov](mailto:George.Ainslie@va.gov)

This material is the result of work supported with resources and the use of facilities at the Department of Veterans Affairs Medical Center, Coatesville, PA, USA. The opinions expressed are not those of the Department of Veterans Affairs or of the US Government.

Published in Gregory Madden and Warren Bickel, eds., *Impulsivity: The Behavioral and Neurological Science of Discounting*. Washington, D.C., APA Books, 2010, pp. 389-410.

There seems to be a corollary of Murphy's Law that applies to human choice: Whenever there is way to get a poorer deal to pay off faster than a better one, some people will fall for it, and some of those will become addicted to it. The grim awareness that "we have met the enemy, and he is us" has required motivational science to take a closer look than has been customary at the basic mechanisms of choice. This has been an opportunity to re-think the elementary process of behavior selection, and in particular to examine whether the reductionist approach, somewhat disused in the human literature since the cognitive revolution of the 1970s, might lead to explanations that more intuitively acceptable approaches have failed to find. That is, a mechanistic or *bottom-up* model of elementary reward-seeking processes that combine by simple principles may describe the complexity of human behavior more parsimoniously than the extensions that holistic or *top-down* theories have had to make to accommodate addictive behavior. Meanwhile brain imaging techniques are beginning to show a functional neuroanatomy of impulsiveness in general and addictions in particular that promises to make the basic units of choice visible (assembled recently in Redish *et.al.* in press), but so far these studies have not revealed how the many processes that have been located at various sites in the brain combine to determine choice. I will not try to anticipate those answers here, but rather will examine how much can be predicted from existing knowledge of the basic properties of choice.

I will first discuss the differences in top-down and bottom-up approaches and the two leading theories of impulsiveness that have been associated with them. Then as an example I will compare these approaches in the case of a common experience that often precedes impulsive choice, sudden craving occasioned by non-informative reminders of consumption. I will argue that the top-down solution, adding classical conditioning as an externality, is inadequate, while the seeming failure of the bottom-up approach in this example is repaired by the same interaction that builds higher-order mental processes ("ego functions"), *recursive self-prediction*. Finally I

will put this example in context by briefly summarizing implications of the basic hyperbolic evaluation function that I have described elsewhere, in particular intertemporal bargaining—another area of recursive self-prediction-- and the reward-dependence of involuntary mental processes, which frees motivational theory from any need to invoke classical conditioning beyond an informational role.

### Top-down and Bottom-up Theories

There are two ways to model human choice systematically. You can start with familiar phenomena—addiction, financial investment, consumption, emotion—and look for simplifying regularities; or you can start with the simplest elements of choice—binary alternatives, a selective factor, a discount function—and look for combining properties by which you can fit the phenomena. Most of recent motivational science has favored the former, *top down* approach, which has the advantage of allowing the theorist to stay close to her target. It also permits her to avoid the question of reductionism—whether human choice can be accounted for entirely by the interaction of simpler mechanisms, a notion that many people find vaguely offensive.

The oldest theories have had to be top down ones, for the obvious reason that these are closest to the observational level. Philosophers such as Aristotle, Aquinas, Spinoza, Kant, and Davidson, psychologists such as James, Lewin, Allport, and Baumeister, and economists generally, take the whole person as a given and study her characteristics. In the area of choice, top-down theories start with norms for rational choice-making, for instance in social settings and competitive markets, and make hypotheses about what factors disturb the execution of these norms in practice. Each might be called a modified rational choice theory (RCT). RCT depicts the self as an autonomous decision-maker within—or above—whatever motivational mechanisms have been discovered. This self assigns value to alternative goals without being bound by these mechanisms. Valuation is said to be a matter of interpreting information according to principles that keep value both internally consistent—comprehensive and transitive—and consistent over time in the absence of new information (Boudon, 1996). The governing self assigns value to external events according to its irreducible judgment, but to be rational these assignments must be transitive and consistent.

Thus rationality demands that the self discount the value of future events according to the only function that will keep their relative values from shifting with changes in delay:

$$V_d = V_0 \delta^d \quad (1)$$

where  $V_d$  is the discounted value of the future event,  $V_0$  is the value of the event if immediate,  $\delta = (1 - \text{discount rate})$ , and  $d$  is delay. Any other function, if it does not generate a straight line, will generate curves from a given amount that sometimes cross the curves from some other amounts at other moments, simply because of the passage of time; that is, it will describe inconsistent preference. Inconsistent preference is clearly irrational because it leaves a person susceptible to being a money pump, that is, it opens her to exploitation by a competitor who repeatedly buys from her when her valuations fall below their exponentially discounted value and sells back to her when they rise (Arrow, 1959; Conlisk, 1996). Inconsistent preference implies in general that a person can expect to make future choices that she does not currently want. In order to deal with instances where people's choices have repeatedly deviated from rationality, and where

explanations such as naivety or a simple difference in discount rate do not apply, RCT has been modified to add an unmotivated principle that overrides rational choice: classical conditioning. Modified RCT is a top-down theory, in that it starts with a coherent executive faculty and seeks to explain its malfunctions.

Aside from errors in information processing (*e.g.* Kahneman *et.al.*, 1982), the principal malfunction discussed has been impulsivity. Impulses are most usefully defined as temporary preferences for options that an individual usually values less than their alternatives. The problems raised by other senses of the word are trivial—spontaneous or whimsical choice, in one usage, or faulty motor control in another (Parker *et.al.*, 1993)—but the phenomenon of temporary preference is the central puzzle in drug addictions and the growing number of behaviors that are seen to have addictive patterns: overeating, credit abuse, habitually self-destructive relationships, absorption in electronic entertainments to the exclusion of relationships, and, archetypically, pathological gambling, to name just a few (Offer, 2006; Ross *et.al.*, 2008). It has been calculated that substance abuse alone is the greatest cause of preventable death in young adults (Robins & Regier, 1991). Impulses are clearly motives. The addict engages in her behavior wholeheartedly and even shrewdly, despite being motivated at other times to limit or avoid it. The critical question is, then, how an ordinarily inferior motive is amplified enough to dominate the motives that had dominated it, in the absence of new information about it. In particular, why does this dominance last for comparatively short periods of time, to be followed repeatedly by either repudiation of or resignation to a consciously disliked behavior pattern?

The answer that has been most intuitively appealing is that there is an additional motivational factor, controlled by an unmotivated process such as association. This seeming invasion by a foreign process leaves RCT itself undisturbed. An association theory is especially credible because of the familiar experience of sudden craving for some specific pleasurable activity, occasioned by only a reminder of this activity. Appetites and emotions seem particularly susceptible to this kind of pattern, so it has been suggested that they form a special class of *visceral* rewards which, when triggered by small reminders, can produce the temporary amplification of reward that a theory of impulsivity needs (Loewenstein, 1996; Laibson, 2001).

Visceral factors include drive states such as hunger, thirst, and sexual desire, moods and emotions, physical pain, and most importantly for addiction, craving for a drug... At intermediate levels, most visceral factors, including drug craving, produce similar patterns of impulsivity, remorse, and self-binding. At high levels, drug craving and other visceral factors overwhelm decision making altogether, superseding volitional control of behavior (Loewenstein, 1999, p. 235).

Ever since Plato made passion a distinct part of the soul, attribution of impulses to a separate motivational force has seemed necessary to preserve RCT as a descriptive—as opposed to merely normative—theory of choice-making. Triggering by reminders is the mechanism described by classical conditioning, a venerable theory of unwanted behavior, of which visceral reward theory is just the latest restatement.

A bottom-up approach can do without this dualism. From the time Descartes realized that human physiology obeyed physical laws theorists have speculated about how to build it from

parts. The associationism of the empiricist philosophers and La Mettrie's first attempt to model choice-making as mechanical (in *Man a Machine*, 1748/1999) have been followed by many reductionist theories, most famously those of Freud, who developed an "economic" theory of what he imagined to be neural energy (Freud, 1895/1956), later called "libido" (Freud, 1923/1956), and Skinner, whose notoriety approached that of La Mettrie after he asserted that all choice can be traced to environmental reinforcement (*e.g.* 1948). A bottom up approach such as behaviorism takes the reductionist bull by the horns and tests how parsimoniously a higher-order process can be predicted by the interaction of simpler elements. However, the behaviorists' models have been limited by a methodological norm that forbids the modeling of mental processes, leaving all complexity to be modeled in the external contingencies of reinforcement that the subject faces (Alston, 1974); this approach will be blind to any contingencies that lie within her. Ironically, only the methods developed by the behaviorists themselves have produced data precise enough to take the modeling of mental processes beyond the common sense level.

I will argue that the extension of behavioral concepts to the conflict of motives within a subject can permit a model that fits familiar experience better than the behaviorists' own environmental contingency model, and that the inadequacy of an RCT model modified by conditioned appetite makes the new model necessary. This model starts with a well-established observation (Green & Myerson, 2004; Kirby, 1997), that all reward-seeking organisms show a robust tendency to devalue expected reward according to a hyperbolic function (Mazur, 1987):

$$V_d = \frac{V_0}{1 + kd} \quad (2)$$

where  $k$  is the discounting rate and the other variables are as defined above. Where smaller rewards precede larger alternatives subjects regularly prefer the larger reward when both are distant, but change to preferring the smaller reward as it becomes imminently available. Left uncompensated, spontaneous preference will make an individual impulsive by the elementary operation of this discount function, without the contribution of any other factor. Compensation for this phenomenon is a major task for the ego functions—perhaps *the* major task, but the same hyperbolic discount function that created the problem can be expected to motivate the learning of solutions. I have developed this argument elsewhere (Ainslie, 2001, pp. 73-140) and will summarize relevant parts at the end of this chapter. In this chapter, however, my main object is to examine a serious objection to the hyperbolic discount curve as the mechanism of impulsiveness: The hyperbolic discount curve itself does not predict the sudden craving elicited by stimuli associated with reward consumption, when these stimuli do not predict increased availability or proximity of this consumption.

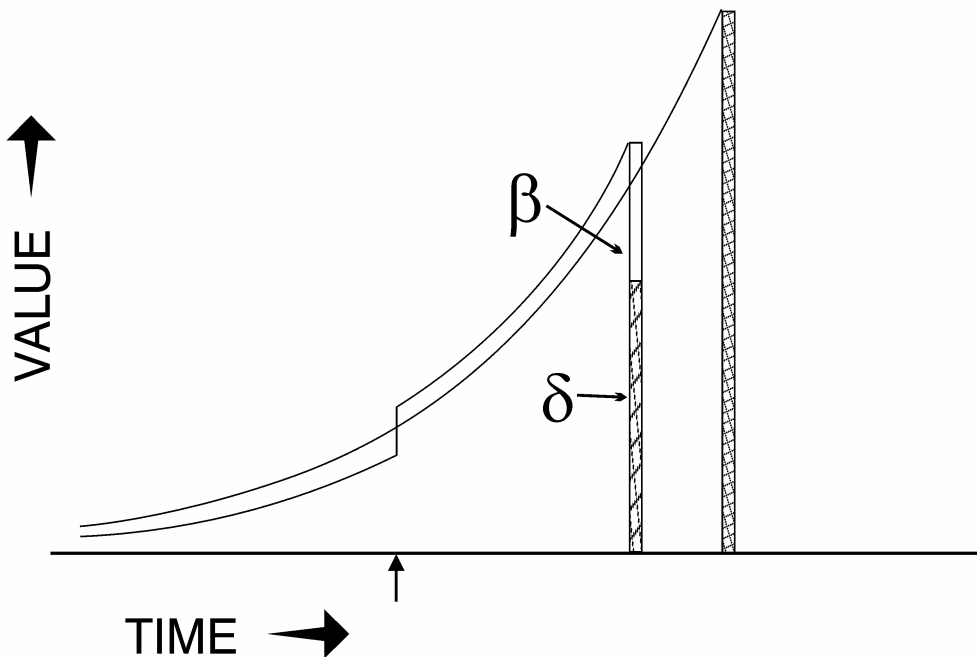
In craving induced by these uninformative cues, some amplifying factor beyond immediacy is clearly operating. Conditioned appetite is an obvious alternative: Perhaps even uninformative cues can amplify a subject's valuation of a good from having been associated with ("conditioned to") its consumption in the past. The conditioned appetite model meshes with the experience of sudden temptation that often precedes impulses. Many authors have continued to rely on conditioned appetite/emotion as an explanation for temporarily amplified valuations (Drummond *et.al.*, 1995; O'Brien, 1997). Problematic impulses usually involve visceral rewards—indeed the occurrence of impulses has been used as a defining property of viscosity, as in the passage

above. An addict may suddenly get intense cravings while watching a show about drugs, for instance. Since such sudden cravings are sometimes implicated in relapses (e.g. Tiffany, 1995), the question naturally arises of whether conditioned appetite is responsible.

The discount function that became associated with visceral reward theory originally described the sudden amplification of value only in situations where reward is imminently available:

$$V_d = V_0 \beta \delta^d \quad (3)$$

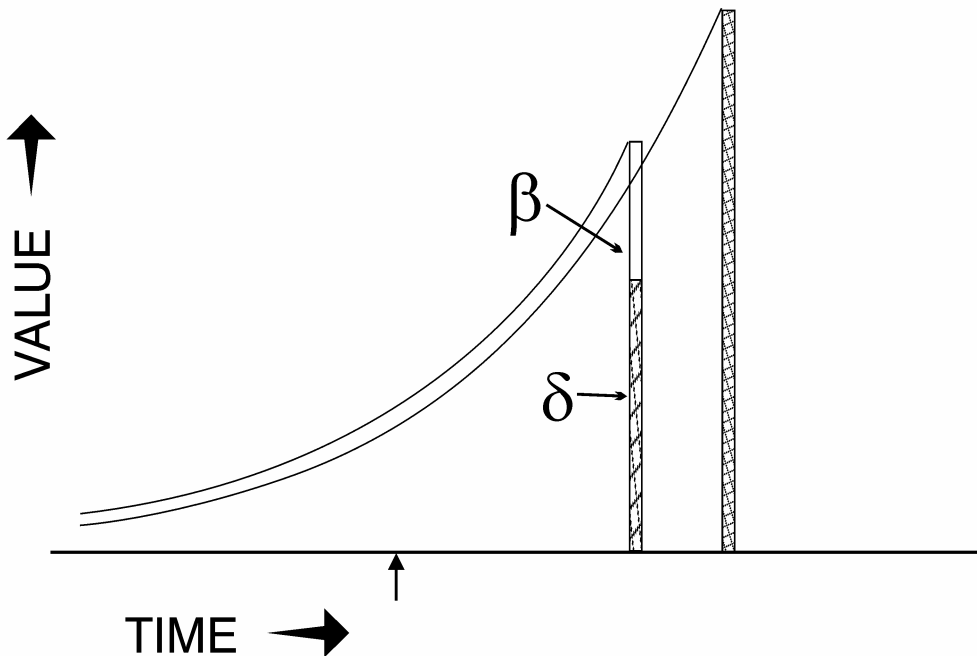
where  $\beta$  (visceral excitatory factor) has one of only two values: if reward is not immediate,  $0 < \beta < 1$ ; if reward is immediate,  $\beta = 1$  (Laibson, 2001; McClure *et.al.*, 2004). The visceral effect was said to come from the immediacy of reward itself. This step function produces a curve that somewhat resembles a hyperbola, an effect that was intentional: Economist David Laibson originally adopted its dual, “hyperboloid,” curve from an article on intergenerational transfers of wealth (Phelps & Pollack, 1968) because “the discount structure [of the curve] mimics the qualitative property of the hyperbolic discount function, while maintaining most of the analytical tractability of the exponential discount function” (Laibson, 1997, p. 450). However, this formula has the same limitation as the hyperbolic function that it approximates (Equation 2). Spikes of appetite often occur without any cue that predicts greater availability or proximity of the reward. Thus it has been proposed that a sudden evocation of viscerality has the same effect as immediacy in sending  $\beta$  to 1.0 (McClure *et. al.*, 2004). This hypothesis converts Laibson’s original proposal from a straightforward discounting theory to a two-factor, conditioning-and-exponential-discounting theory (Figure 1A).



**Figure 1A.**  $\beta$ - $\delta$  (exponential) curves drawn according to Equation 3, except that  $0 < \beta < 1$  before a CS occurs at arrow on x-axis, and  $\beta = 1$  afterwards. The CS elicits a CR (appetite, the unfilled  $\beta$  part of the bar) that raises the effective height of the SS reward.

The mechanism that evokes this viscerality obviously has to be classical conditioning. Research on conditioning has evolved considerably since Pavlov's first experiments (Pavlov, 1927). The initial observation was that some events (unconditioned stimuli, UCSs) elicit reflexive responses (unconditioned responses, UCRs). Initially UCSs seemed to select for the transfer of UCRs to arbitrarily designated stimuli (conditioned stimuli, CSs) that predicted their occurrence, regardless of whether the CSs predicted that the transferred responses (now called conditioned responses, CRs) would be rewarded. However, experimenters' initial conclusion that UCSs were sufficient to select for the transfer of UCRs to new stimuli did not stand up. The site of selection by UCSs is on stimuli, not behaviors; that is, the pairing of novel stimuli with UCSs produces only CSs, not true CRs (Mackintosh, 1983; Rescorla, 1988). Only information is learned by association; the occurrence and timing of "CRs" depends on what incentives are created by this information. Even the selection of the CR was recognized very early not to depend on what response is elicited by the UCS-- Only a few CRs happen to be the same, in detail or even in approximate kind, as those elicited by the UCS (Upton, 1929; Zener, 1937). However, many emotions and appetites lack external signs and are apt to be innately connected to particular kinds of expectation, making it unclear to what extent the disproof of conditioning as a mechanism of response selection applies to them. CRs might still exist in the form of appetites or emotions. We will have to leave this possibility open.

The assumption that appetites and thus amplification of reward can be transferred by association is consistent with the seemingly unmotivated appetites of visceral learning theory, but experiments on the timing of CRs—or operant responses to CSs—have been less supportive. These experiments have regularly shown that responses based on associated stimuli reflect the exact timing of when the original objects have been available. In parametric experiments CRs anticipate the occurrence of UCSs with great accuracy. If a CS occurs or begins well before a UCS is due, subjects learn to estimate the delay and emit the CR just before the UCS (Gallistel & Gibbon, 2000; Kehoe *et al.*, 1989; Savastano *et al.*, 1998). Appetites specifically-- cue-induced craving for cigarettes, and skin conductance and salivation related to craving-- are strongly dependent on whether puffing is available within the next minute (Carter & Tiffany, 2001; Field & Duka, 2001). Thus, in the laboratory involuntary responses closely track the prospect of reward. Furthermore, if the general process of learning behavioral contingencies in daily life counts as conditioning—the usual assumption—then non-anticipatory appetites are an anomalous variant here as well. Where the consumption of an addictive substance never happens in a given circumstance, humans do not crave it: Opiate addicts and alcoholics in programs that allow consumption on only certain days report absence of craving on other days (Meyer, 1988), and observant orthodox Jews, who never smoke on the Sabbath, are reported not to crave cigarettes then (Dar *et al.*, 2005; Schachter *et al.*, 1977). Classical conditioning, which is just associative learning, does not explain appetites that are disproportionate to the prospect of consuming and that change without changes in this prospect. It is true that some cue-induced appetites/emotions have been reported to grow without further contact with UCSs (Eysenck, 1967); but such examples, if they are valid (Malloy & Levis, 1990) go beyond passive association and thus need explanation themselves. Furthermore, visceral reward theory does not make it clear why the associative process should not lead amplified visceral rewards to be anticipated and discounted like any other reward, so that they become preferred consistently, rather than temporarily, where the amplification is great enough to ever make them preferred (Figure 1B).



**Figure 1B.**  $\beta$ - $\delta$  curves at equilibrium drawn according to Equation 3, except that  $0 < \beta < 1$  before a CS occurs at arrow on x-axis, and  $\beta = 1$  afterwards. Repeated elicitation of appetite by a CS should make the rise in SS height expectable, eliminating the change of preference.

After all, rewards are usually remembered as consumed in the presence of appetite; adjusting their value for current hunger or satiety is a distinct process (Balleine & Dickinson, 1998). Conditioning should not reduce the efficiency of the reward process, but simply assign a true value to the prospect of a reward that has been consumed at the usual level of appetite.

Given the strictly informational role of conditioning, the premature occurrence of appetite is still a puzzle for visceral reward theory. Why does the addict develop craving when merely reminded of consumption? A straightforward application of hyperbolic discounting also fails to predict spikes of appetite in response to cues that do not convey new information. An extension of either conditioning or hyperbolic discounting theory can let these theories handle the occurrence of prematurely spiking appetites/emotions. However, I will argue that this extension works well only for the hyperbolic theory.

### **Recursive Self-prediction of CRs**

In laboratory experiments consumption goods are necessarily outside of the subject's control. Appetite is studied as a function of when the experimenter signals their availability. In daily life, by contrast, goods that might be consumed impulsively are available much of the time, and their

consumption is limited by a person's decisions. CSs in life situations will be very different than they are in the laboratory.

Conditioning theory has long departed from the notion that CSs have to be concrete stimuli. They can be just temporal patterns, interpreted stochastically by the subject (Gallistel, 2002), a finding that can be summarized by saying that the expectation of a UCS, in whatever form, functions as a CS. In humans, and possibly other organisms to a limited extent, expectation includes predictions of the individual's own behavior. If a person's conscious intentions entirely committed her future behavior such prediction would be superfluous, of course; she could predict her behavior directly by examining these intentions. But behavior in even the near future is increasingly recognized as beyond the scope of such examination (Wegner, 2002, especially pp. 63-144), and may depend on the dynamics of a population of motivated processes (Ainslie, 2001, pp.39-44). This means that expectation is apt to be recursive, with an expectation of a UCS, say, taking cocaine, functioning as a CS and inducing the CR of appetite. But where availability is not a limiting factor, an increase in appetite will itself increase the likelihood of taking the cocaine. If this likelihood increases, the CS of expecting cocaine should increase, and in turn the CR of appetite again.

Wherever a person's consumption is limited mainly by her own choice, appetite can be a positive feedback system of the kind first described by Darwin, James, and Lange:

The free expression by outward signs of an emotion intensifies it. On the other hand, the repression, as far as this is possible, of all outward signs softens our emotions. He who gives way to violent gestures will increase his rage; he who does not control the signs of fear will experience fear in greater degree (Darwin, 1872/1979, p. 366).

As a mechanism for finding a response, Darwin's suggestion has been disproven by a number of experiments (concisely summarized by Rolls, 2005, pp. 26-28), but none of these findings touches on its possible role in the modulation of a given response once an individual has focused on it. Whenever an arbitrary stimulus has been associated with consumption in the past, the appearance of that stimulus might accurately predict an increased current likelihood of consumption, and accordingly function as a conventional CS. A sudden spike of appetite could thus come from the existence of positive feedback conditions. These conditions may obtain whenever the person's consumption is determined mainly by her choice about a readily available consumption good, but are apt to have the strongest effect when she has weak-to-moderate resolve not to consume: Where a person is not trying to restrain consumption she will keep appetite relatively satisfied; where she is confident of not consuming regardless of appetite (as in cases of opiates in scheduled addicts and smoking in orthodox Jews) she will not expect appetite to lead to consumption. In both of these cases, a stimulus associated with consumption should be only a trivial CS, and thus not lead to an exceptional CR. In a recovering addict or restrained eater, by contrast, cues predicting that she might lapse could elicit significant CRs. Even in the numerous laboratory experiments where drug cues induce craving in addicts, the subjects know that they might be able to actually obtain drugs as soon as they leave.

However, conditioning theory itself limits the viability of this recursive model. CRs are supposed to be UCRs that have been passively transferred to CSs because CSs predict UCSs. If CRs are only such anticipatory responses, their amplitude should be limited to no more than that

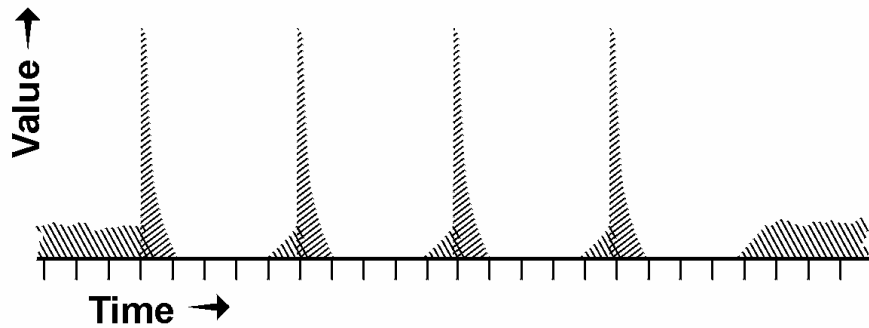


of their UCRs. If a person's expectation of consumption increases by  $x\%$ , her appetite (CR) should increase by no more than  $x\%$  of the UCR, or at least of what the CR would be when certainty is 100% and delay is zero. Since we are discussing delays that are significantly greater than zero—the cases where hyperbolic curves *per se* do not explain the spiking—the increase in appetite should be markedly less than  $x\%$ . Conversely, if a person's appetite increases by  $x\%$ , the increase in estimated probability of consumption that this causes should also be fractional, reflecting the proportion of times when that much increase in appetite has been followed by actual consumption. If a recovering addict, for instance, has moderate resolve not to relapse, an initial confrontation with a drug stimulus should increase her likelihood of relapsing by only a marginal amount. This increase should in turn have only a small effect on her conditioned craving (CR), which would be expected to increase her expectation of relapse by an even smaller amount again. The positive feedback effect should be damped down unless the percentage of each increase is perfectly preserved, and even in that case the CR will be capped at the level of the UCR and discounted for whatever delay is unavoidable. The qualitative elements for explosive craving are there, but quantitatively the amplification that results from the recursive process will be limited.

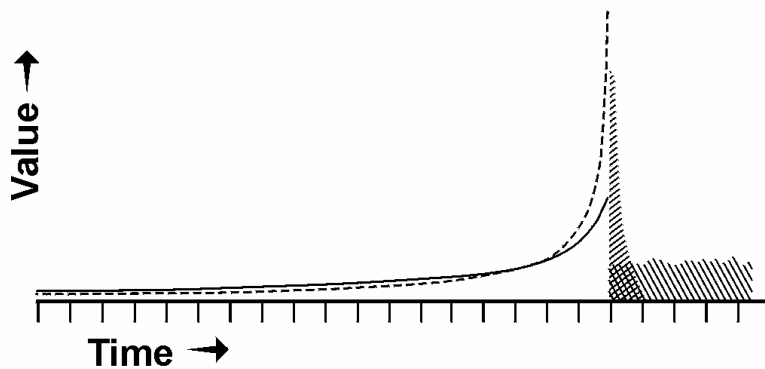
### **Recursive Self-prediction of Motivated Appetite**

If appetites themselves are operants—that is, if appetites are selected not by the simple transfer of a UCS but by reward for their activity, as I have proposed elsewhere (Ainslie, 2001, pp. 48-70, 161-174), this limitation disappears. However, this is not a trivial change in the conventional view of appetite. I need to briefly review my argument that hyperbolic discount curves permit involuntary and even aversive processes to be incorporated into a unified motivational marketplace.

First we must strip the selective factor, “reward,” of its connotations of pleasure and reduce it to its defining function: that which selects for a process that it follows.<sup>2</sup> Then we can predict what pattern will be produced by recurring periods of reward followed, if chosen, by obligatory periods of nonreward (or reduced reward). By changing the durations of the reward periods we can variously describe the binge-followed-by-hangover pattern seen in many addictions (duration-before-negative-consequences = maybe an hour for bulimics, up to several days before a binge drinker gets sick), the repeated urges reported in psychogenic itches, tics, and unwanted mannerisms (duration-before-negative-consequences = seconds), and, by extension, the attraction of attention that is fused experientially with behavioral aversion in negative emotions such as panic, the emotion-like aversive component of pain itself (“protopathic” pain-- Melzack & Casey, 1970), and, variably, rage (duration-before-negative-consequences = fractions of a second). Figure 2A depicts a level of background reward being replaced by a period of cyclic reward spikes that are each followed by reward inhibition. The figure can represent binges, tics, or negative emotions depending on cycle length.



**Figure 2A.** Motivated processes that are avoided at a distance range from behavioral impulses to painful emotions, depending on whether the time scale represents days or fractions of a second. Any such process can be accounted for as a cycle of brief, intense reward (rightward hatching) that interrupts an ongoing baseline reward (leftward hatching) for a relatively longer time (from Ainslie, 1992).



**Figure 2B.** Hyperbolic discount curves drawn from a single spike of reward (or urge) versus a lower but longer-lasting alternative reward. Each curve is the sum of the curves from each moment of reward. The spike has less area than the baseline reward to which it is an alternative; but because it is taller it will be preferred just before it is available (from Ainslie, 1992).

Figure 2B shows hyperbolic discount curves drawn to the alternatives of a single cycle vs. the same length of background reward. The spike that begins the cycle can be expected to make it preferred temporarily when it is close, even when the net effect on reward is strongly negative. At the short end of the spectrum of durations the initial spikes characterize any option that is experienced as vivid; the subsequent troughs may vary in depth, with low ones creating aversive experiences that individuals try to escape, while shallow or absent troughs create positive

experiences (thrill, elation, some examples of rage) that are limited by satiation. In this model conditions that make panic or rage possible can be viewed as enabling appetites for them just as nonsatiety enables appetites for the consumption of substances—offers of reward, even if they are frequently “offers you can’t refuse.” The value of this model is that both positive and negative appetites can be seen as luring organisms into participating in them, rather than springing up automatically like reflexes, outside of the marketplace of choice. As long as it is discounted hyperbolically, reward can then be the selective factor not only for enduring pleasures, but for temporary, regretted pleasures and for urges that do not feel pleasurable at all.

Soon after the description of classical conditioning, experimenters noted that all stimuli that could induce conditioned responses had motivating power as well (Hull, 1943; Miller, 1969). The possibility that reward and the selective principle in UCSs are identical has been proposed before (reviewed in Pear & Eldridge, 1984; Donahoe *et.al.* 1993), but has not been pursued extensively, perhaps because a separate selective principle has seemed necessary to explain why aversive experiences can compete for attention, and because no important hypotheses have hinged on this identity. But with hyperbolic discounting, seduction over various time courses accounts for attention to aversive experiences without any need for a separate selective principle, as I have just described. Of course, it may turn out that this attention can be better explained by other hypotheses that integrate reward and nonreward in close temporal proximity, or even simultaneity, as reported qualities such as “motivational salience” are analyzed (*e.g.* Berridge & Robinson, 1998, pp. 348-349), but so far there has been no other hypothesis at the behavioral level about how an incentive can attract attention while repelling behavior in the marketplace of choice. As to whether any behavioral hypotheses depend on the question of one vs. two basic selective factors, I am about to argue that the sudden occurrence of appetite when its object is no nearer requires appetite to be reward-dependent—that is, selected by the same factor as motor behavior.

If appetite is an operant, a recursive reward-seeking model is possible which can predict the same observations as a classical conditioning model but without the damping effect: A cue predicting availability will be a cue for generating appetite if the potential for consumption to be rewarding (nonsatiety) exists. When availability and nonsatiety exist, and when consumption is limited by self-control—probably the case only in humans—appetite itself has the potential to obtain fast-paying reward by motivating the abandonment of (slow-paying) self-control. The most rewarding amount of appetite then may not be that which optimizes the experience of consumption, given its probability, but rather that which makes consumption most probable. The most productive timing of such appetite will take the form of concentrated attempts on discrete occasions; if appetite does not succeed in inducing consumption on a particular occasion it would waste effort by prolonged activity. Occasions for appetites could be arbitrary, especially at higher levels of deprivation, but the occasions that are the most apt to promise successful attempts are limited ones-- an external reminder, or a circumstances where they have succeeded in the past. In this view, the force of symbols and other reminders in relapse comes not from their effect as CSs but from their providing reward-seeking appetites with focal occasions to try to overturn self-control.

This theory of appetite is part of a broader bottom-up model that depicts reward-dependent processes as competing for acceptance on the basis of the current, hyperbolically-discounted value of the prospective reward for these processes. The fact that preferability among a set of

processes can shift as a function of time alone puts these processes in a strategic competition with each other (Ainslie, 1992, pp. 154-179; 2001, pp. 90-100). They operate as somewhat independent agents that have some but not all interests in common, on the basis of what are common contingencies of reward but differently discounted valuations of them. An appetite in this model arises when an individual perceives the opportunity for consumption that can be made either more rewarding or more likely by this appetite. The contingencies that determine whether appetite as a quasi-independent agent asserts itself in a given situation will be roughly the same as those determining whether a pet begs its owner for food. Begging is a low-cost behavior and is apt to occur whenever the pet encounters food-related cues; again, continuous begging will not be worthwhile. But if the pet is never fed in a particular circumstance the begging gradually extinguishes. Likewise, if food is so available that significant deprivation does not occur, begging adds no value. By analogy, the restrained eater or recovering addict has the potential to experience intense reward by indulging in immediate consumption. Insofar as appetite makes consumption look even a little more likely it will pay for itself, and any signs of weakening will serve as cues that still more appetite may succeed in motivating consumption. The low cost of appetite may explain why it must go unrewarded consistently and over many trials in a given circumstance to extinguish—for instance, why the orthodox Jews who do not get cravings on the Sabbath do get cravings at work, even though they know they cannot smoke (Dar *et al.*, 2005).

According to this theory a person does not come to anticipate the higher value of the smaller-sooner reward, as in the progression from Figure 2A to 2B, because the eruption of the positive feedback cycle is not a reliable occurrence. It may even be that people who are trying to prevent an impulse try not to anticipate the cycle for fear of triggering it.<sup>3</sup>

Thus appetite as an operant can be subject to the same positive feedback mechanism as conditioned appetite; but there are two important differences between this and the recursive conditioning model. One is that the degree of appetite will not be limited to mere anticipation, but can be whatever increases the prospect of reward. The other is that appetite can occur without a stimulus, or can be occasioned by a stimulus that has not been associated with consumption. Addicts commonly experience craving spontaneously and even invite it; and in a well-studied example of negative appetites, the phobias, conditioning events are not usually found (Lazarus, 1972; Wolpe, 1981). There will still be constraints on the generation of appetites—in modalities where unsatisfied appetite brings hunger pangs or withdrawal symptoms these will be deterrents; and appetite without a limited occasion will extinguish (see Ainslie, 2001, pp. 166-171)—but the explosive appetite that so often ends people's efforts at controlled consumption can be understood as a motivated process that has sought to do exactly that.

For completeness I should mention a factor that is likely to increase the role of recursive self-prediction in appetite: the apparently intrinsic rewarding effect of some appetites (Herrnstein, 1977). Some people get pleasure from reading cookbooks without expecting to eat what they read about, and people who are prone to panic may have panic episodes without perceiving any danger. Some appetites at least must be intrinsically rewarding. This rewardingness will be limited not only by the aforementioned pangs or withdrawal symptoms that may follow it, but also by its own tendency to habituate, as is the case with emotions (Ainslie, 2001, pp. 164-174).

In summary, the model of appetite as an operant cued by recursive self-prediction is uniquely able to account for explosive appetite—that is, for why a cue associated with consumption but

not predictive of increased availability of a consumption good should lead to a great increase in appetite. In this model the cue is needed only to give occasion, that is, to select one moment over another for a focused attempt at reversing the dominant preference. The model depends on the hyperbolic shape of the discount curve, since an individual with consistent preferences over time would have no short range motive to undermine her own resolutions, or indeed any long range motive to make resolutions in the first place. Similar predictions are made by the “hyperboloid” ( $\beta$ - $\delta$ ) step function of visceral reward theory (Equations 1 or 3); but this shape can be generated only by the explosive appetite that, according to the damping argument presented above, the theory’s classical-conditioning mechanism would be inadequate to produce. That is, the only viable mechanism for the  $\beta$ - $\delta$  discount curve is for discount curves to have an elementary hyperbolic or other hyperconconcave shape to begin with.

### **A Self-prediction Model Describes Will (and its Failure) as Well as Sudden Appetite**

Both modified RCT and hyperbolic discounting theory have come to include means of explaining most of the ways that people make choices. I have selected a case where the difference in their predictiveness is apparent: the ways in which the models deal with the failure of one basic ego function, maintenance of consistent choice over time. This example should be seen in the context of larger models, which predict differences in how this ego function operates to begin with. These differences in turn are part of the fundamental difference in how these models depict the action and scope of motivation, which I have already described: RCT accounts for experientially unmotivated processes with the transferred reflexes of conditioning theory; hyperbolic discounting theory encompasses all processes that are not truly reflexes in a comprehensive marketplace of motivation. RCT assumes the maintenance of consistency to be implicit in the accurate gathering of information about goals, since rationality itself dictates consistent long range behavior; a faculty of will is sometimes recognized, but explicit hypotheses about willpower essentially restate common intuitions about it—for instance that it resembles a muscle that can be exhausted by overuse but strengthened by exercise (Baumeister *et.al.*, 1994). In the hyperbolic model will is essential to consistent choice over time. Because of the innate tendency to form temporary preferences, hyperbolic discounting predicts not only inconsistency over time but recursive motivational processes, from which emerge both higher-order mental functions and impulses that do not depend on temporal proximity.

The mechanism of willpower as *intertemporal bargaining* based on recursive self-prediction has been often presented over the past thirty years (Ainslie, 1975, 1992 pp.144-173, 2001, pp. 78-116) and will only be summarized here: Hyperbolically discounted reward will create what is in effect a population of reward-seeking processes that can be grouped loosely into *interests* on the basis of common goals, just as economic interests can be identified in market economies. The choice-making self will have many of the properties of an economic marketplace, with a scarce resource—access to the individual’s limited channel of behavior—bid for with a common currency—reward. Maintenance and change of choice will be governed by intertemporal bargaining, the activity in which reward-seeking processes that share some goals (*e.g.* long term sobriety) but not others (when to have drinks) maximize their individual expected rewards, discounted hyperbolically to the current moment. This *limited warfare* relationship is familiar in interpersonal situations, where it often gives rise to “self-enforcing contracts” such as nations’

avoidance of using a nuclear weapon lest nuclear warfare become general. In interpersonal bargaining, stability is achieved in the absence of an overarching government by the parties' recognition of repeated prisoner's dilemma incentives. In intertemporal bargaining, *personal rules* arise through a similar recognition among the successive motivational states of an individual, with the difference that a future state is not motivated to retaliate, as it were, against past states that have defected. The risk of future states' loss of confidence in the success of the personal rule, and consequent defection in their own short term interests, will present the same threat as the risk of actual retaliation. The reason that a recovering alcoholic avoids taking a single drink is not that it will make her drunk but that it will impair the credibility of her sobriety, without which she does not have much current reason not to get drunk.

As with interpersonal negotiations, intertemporal cooperation is threatened by the availability of alternative truce lines. Under the pressure of current temptation the alcoholic may reason that drinking on New Year's Eve would not reduce her expectation of staying sober the rest of the year. But then that might be true of her birthday too, or *your* birthday, or Saturdays... At least with alcohol there does exist a *bright line* between some drinking and none, whereas an overeater or spendthrift has much less defense against rationalizations—She has to eat *some* food and spend *some* money, and it is hard to see one diet or budget at irreplaceable. This is a large topic, but the point for the present discussion is that the intertemporal bargaining situation that hyperbolic discount curves create for a self-aware person does not require any kind of overseer to reach a stable equilibrium. Repeated prisoners dilemma contingencies can create a will without an organ, serving a self without a seat, just as the “will” of nations not to use nuclear weapons seems to be guided by an invisible hand.

Thus the model that hyperbolic discounting makes possible does not postulate any preformed or overarching faculty that makes choice coherent. Its starting place is a sequence of preferences based on the shifting dominance of incentives, which strictly determine all choice, but are not strictly controlled by external events. Higher-order mental processes, which look beyond immediate advantage and make choice somewhat coherent, are formed from simpler reward-seeking processes that must compete on the basis of the leverage afforded by superior foresight. They therefore cannot stand outside of the reward process in order to evaluate it. This is a bottom-up theory, in that it starts with a population of processes that have been shaped by differential reward and seeks to explain the ways that this population creates an executive faculty.

The differences between this model and the muscle analogy are not only that the will's direction is implicit in the mechanism itself and that the will can be simultaneously strong and weak in different areas, but also that this model predicts familiar pathologies of will—compulsiveness, denial, boredom, and circumscribed dyscontrol symptoms where a person abandons attempts at will (Ainslie, 2001, pp. 143-197): The awareness of your choices as test cases for larger plans makes you lawyerly, perhaps at the expense of being able to live fully in the here-and-now; the loss of prospective reward that comes from recognizing a lapse moves you not to notice key aspects of your choice-making; overly successful self-control reduces surprise, which is necessary to prevent the attenuation of emotional experience; and lapses in a particular circumstance often lead people to stop trying to use their will there, lest its general credibility be reduced.

The positive feedback system that I have just described for the sudden explosion of appetite follows the same pattern as the sudden wilting of a resolution following a lapse. The sudden appetite may indeed initiate the crash and form part of its motivational feedback loop, although it will probably not itself be experienced as a failure of will, but as the temptation that led to the failure. The difference between these two recursive processes—the collapse of will and the explosion of appetite-- may be in the time it takes to review alternatives versus only registering new self-predictions about a single alternative. The process of reviewing predictions about the impact of alternative choices on future bargaining (sometimes called vicarious trial-and-error or VTE—Tolman, 1939; Johnson & Redish, 2007) must obviously take time, if only some fractions of a second; so “deliberate” choice-making will govern only behaviors that do not kindle too rapidly once thought of. A shift of attention, for instance, occurs too rapidly to be tested against a personal rule, and the components of VTE itself cannot be made contingent on a process that depends on VTE-- hence the famous impossibility of using will to not think about white bears (Wegner, 1989); once you have tested whether you are thinking about white bears, you have already thought about them. People are strongly motivated to discriminate those recursive self-prediction processes that may support personal rules from the rapid ones that cannot support them, even when the difference is better described as a temporal continuum than as a dichotomy; you risk blaming yourself, and weakening your will, where you perceive personal rules as relevant to your action, but not where the action is “involuntary”. This discrimination may have contributed to the intuition that appetites are reflexive, unmotivated processes.

## Conclusions

People often deviate from their own plans after seeing or even thinking about poorer but faster paying alternatives. Since this experience often resembles laboratory observations of classical conditioning, both top-down (holistic) and bottom-up (mechanistic) theories of impulsiveness have borrowed this concept to explain irrational or seemingly less rewarded choices. However, increasingly analytic experiments have shown that conditioning is just the learning of information, and conditioned responses, including appetites, are timed precisely to make the most effective use of that information. Such learning does not account for the apparent elicitation of appetites by uninformative cues. The bottom-up approach offers a solution: Sudden appetite in this circumstance is best attributed to foraging by appetites as reward-seeking processes, becoming explosive (positively fed back) where consumption or not of the appetite’s object is a matter of self-control. The theory that appetites and other involuntary mental processes are operants requires—and follows readily from—the hyperbolic discounting of reward, and the definition of reward as whatever increases the frequency of a process that it follows. Amplification of prospective reward by recursive self-prediction, rather than by arbitrary association, lets the concept of conditioned responses finally return to the role that Pavlov originally described, responses simply “conditional” on new information (Dinsmoor, 2004).

## References

- Ainslie, G. (1974) Impulse control in pigeons. *Journal of the Experimental Analysis of Behavior* 21, 485-489.
- Ainslie, G. (1975) Specious reward: A behavioral theory of impulsiveness & impulse control. *Psychological Bulletin* 82, 463-496.
- Ainslie, G. (1991) Derivation of "rational" economic behavior from hyperbolic discount curves. *American Economic Review*, 81, 334-340.
- Ainslie, G. (1992) *Picoeconomics: The Strategic Interaction of Successive Motivational States within the Person*. Cambridge: Cambridge U.
- Ainslie, G. (2001) *Breakdown of Will*. New York: Cambridge U.
- Ainslie, G. (2005) Précis of *Breakdown of Will*. *Behavioral & Brain Sciences* 28(5). 635-673.
- Alston, W.P. (1974) Can psychology do without private data? *Behaviorism*, 1, 71-102.
- Arrow, K. J. (1959) Rational choice functions & orderings. *Economica*, 26, 121-127.
- Balleine, B. & Dickinson, A. (1998) Goal-directed action: contingency and incentive learning and their cortical substrates. *Neuropharmacology*, 37, 407-419.
- Becker, G. & Murphy, K. (1988) A theory of rational addiction. *Journal of Political Economy*, 96, 675-700.
- Bennett, A. (1918) *Self and Self-Management*, New York: George H. Doran.
- Berridge, K. C. (2003) Pleasures of the brain. *Brain & Cognition*, 52, 106-128.
- Berridge, K. C. & Robinson, T. (1998) What is the role of dopamine in reward: Hedonic impact, reward learning, or incentive salience. *Brain Research Reviews*, 28, 309-369.
- Carter, B. L. & Tiffany, S. T. (2001) The cue-availability paradigm: The effects of cigarette availability on cue reactivity in smokers. *Experimental & Clinical Psychopharmacology*, 9, 183-190.
- Conlisk, J. (1996) Why bounded rationality? *Journal of Economic Literature*, 34, 669-700.
- Cubitt, R. P. & Sugden, R. (2001) On money pumps. *Games & Economic Behavior*, 37, 121-160.
- Dar, R., Stronguin, F., Marouani, R., Krupsky, M., & Frenk, H. (2005) Craving to smoke in orthodox Jewish smokers who abstain on the Sabbath: A comparison to a baseline and a forced abstinence workday. *Psychopharmacology*, 183, 294-299.
- Darwin, Charles (1872/1979) *The Expressions of Emotions in Man and Animals*. London: Julian Friedman.
- 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.
- Drummond, D. C., Tiffany, S. T., Glautier, S., & Remington, B. (1995) *Addictive Behavior: Cue Exposure Theory & Practice*. Wiley.
- Eysenck, H. J. (1967) Single trial conditioning, neurosis & the Napalkov phenomenon. *Behavior Research & Therapy*, 5, 63-65.
- Field, M. & Duka, T. (2001) Smoking expectancy mediates the conditioned responses to arbitrary smoking cues. *Behavioural Pharmacology*, 12, 183-194.
- Flavell, J. (1976) Metacognitive aspects of problem solving. In B. Resnick (ed.), *The Nature of Intelligence*. Erlbaum.



- Gallistel, C. R. (2002) Frequency, contingency and the information processing theory of conditioning. In P. Sedlmeier & T. Betsch (Eds.). *Frequency Processing and Cognition*. Oxford.
- Gallistel, C. R. and Gibbon, J. (2000) Time, rate, and conditioning. *Psychological Review* 107, 289-344.
- Green, L. & Myerson, J. (2004). A discounting framework for choice with delayed and probabilistic rewards. *Psychological Bulletin*, 130, 769-792.
- Hollander, E. & Stein, D. J. (2006) *Clinical Manual of Impulse-Control Disorders*. American Psychiatric Publishing.
- Hull, C.L. (1943) *Principles of Behavior*. New York: Appleton-Century-Crofts.
- Johnson, A. and Redish, A. D. (2007) Neural ensembles in CA3 transiently encode paths forward of the animal at a decision point. *Journal of Neuroscience*, 12, 483-488.
- Kehoe, E. J., Graham-Clark, P., & Schreurs, B. G. (1989) Temporal patterns of the rabbit's nictitating membrane response to compound and component stimuli under mixed CS-US intervals. *Behavioral Neuroscience*, 103, 283-295.
- Kirby, K. N. (1997) Bidding on the future: Evidence against normative discounting of delayed rewards. *Journal of Experimental Psychology: General*, 126, 54-70.
- Laibson, D. (1997) Golden eggs and hyperbolic discounting. *Quarterly Journal of Economics*, 62, 443-479.
- Laibson, D. (2001) A cue-theory of consumption *Quarterly Journal of Economics*, 66, 81-120.
- La Mettrie, J. (1748/1999) *Man a Machine*. Open Court.
- Lazarus, A. (1972) Phobias: broad-spectrum behavioral views. *Seminars in Psychiatry* 4, 85-90.
- Loewenstein, G. (1996) Out of control: Visceral influences on behavior. *Organizational Behavior and Human Decision Processes*, 35, 272-292.
- Loewenstein, G. (1999) A visceral account of addiction. In J. Elser & O.-J. Skog (Eds.) *Getting Hooked: Rationality and Addiction*. Cambridge U.
- MacKintosh, N. J. (1983) *Conditioning and Associative Learning*. New York, Clarendon.
- Malloy, P.F. and Levis, D.J. (1990) A human laboratory test of Eysenck's theory of incubation: A search for the resolution of the neurotic paradox. *Journal of Psychopathology and Behavioral Assessment*. 12, 309-327.
- Mazur, J.E. (1987) An adjusting procedure for studying delayed reinforcement. In M.L. Commons, J.E. Mazur, J.A. Nevin, & H. Rachlin (Eds.), *Quantitative Analyses of Behavior V: The Effect of Delay and of Intervening Events on Reinforcement Value*. Hillsdale, NJ: Erlbaum.
- McClure, S. M., Laibson, D. I., Loewenstein, G., & Cohen, J. D. (2004) The grasshopper and the ant: Separate neural systems value immediate and delayed monetary rewards. *Science*, 306, 503-507.
- McClure, S. M., Ericson, K. M., Laibson, D. I., Loewenstein, G., & Cohen, J. D. (2007) Time discounting for primary rewards. *The Journal of Neuroscience*, 27, 5796-5804.
- Melzack, R. & Casey, K.L. (1970) The affective dimension of pain. In M.B. Arnold (Ed.), *Feelings and Emotions*, pp. 55-68. New York: Academic.
- Meyer, R. E. (1988) Conditioning phenomena and the problem of relapse in opioid addicts and alcoholics. In Ray, B. (Ed.), *Learning Factors in Substance Abuse. NIDA Research Monograph series 84*, 161-179. NIDA.
- Miller, N. (1969) Learning of visceral and glandular responses. *Science*, 163, 434-445.

- O'Brien, C. (1997) A range of research-based pharmacotherapies for addiction. *Science*, 278, 66-70.
- Offer, Avner (2006) *The Challenge of Affluence: Self-control and Well-Being in the United States and Britain since 1950*. Oxford.
- Parker, J. D. A., Bagby, R. M., and Webster, C. D. (1993) Domains of the impulsivity construct: A factor analytic investigation. *Personality and Individual Differences* 15, 267-274.
- Pavlov, I. P. (1927) *Conditioned Reflexes: An Investigation of the Physiological Activity of the Cerebral Cortex* (G. V. Anrep, Trans.). Oxford.
- Pear, J., J. & Eldridge, G. D. (1984) The operant-respondent distinction: Future directions. *Journal of the Experimental Analysis of Behavior*, 42, 453-467.
- Phelps, E. S. & Pollack, R. A. (1968) On second-best national saving and game-equilibrium growth. *Review of Economic Studies*, 35, 185-199.
- Piaget, J. (1937/1954) *Construction of Reality in the Child*. M. Cook, Trans. New York: Basic.
- Redish, A. David, Jensen, Steve, and Johnson, Adam (in press) A unified framework for addiction: Vulnerabilities in the decision process. *Behavioral and Brain Sciences*.
- Rescorla, R. A. (1988) Pavlovian conditioning: It's not what you think it is. *American Psychologist*, 43, 151-160.
- Robins, L.N. and Regier, D.A. (1990) *Psychiatric Disorders in America*. New York: Free Press.
- Rolls, E. T. (2005) *Emotion Explained*. Oxford.
- Ross, Don, Sharp, Carla, Vuchinich, Rudy and Spurrett, David (2008) *Midbrain Mutiny: The Picoeconomics and Neuroeconomics of Disordered Gambling*. MIT.
- Savastano, H. I., Hua, U., Barnet, R. c. & Miller, R. R. (1998) Temporal coding in Pavlovian conditioning: Hall-Pearce negative transfer. *Quarterly Journal of Experimental Psychology*, 51, 139-153.
- Schachter, S., Silverstein, B. & Perlick, D. (1977) Psychological and pharmacological explanations of smoking under stress. *Journal of Experimental Psychology: General*, 106, 31-40.
- Skinner, B.F. (1948) Superstition in the pigeon. *Journal of Experimental Psychology* 38, 168-172.
- Tiffany, S. T. (1995) Potential functions of classical conditioning in drug addiction. In D. C. Drummond, S. T. Tiffany, S. Glautier & B. Remington (Eds.), *Addictive Behavior: Cue Exposure Theory and Practice*. Wiley.
- Tolman, E. C. (1939) Prediction of vicarious trial and error by means of the schematic sowbug. *Psychological Review*, 46, 318-336.
- Upton, M. (1929) The auditory sensitivity of guinea pigs. *American Journal of Psychology*, 41, 412-421.
- Watson, J. B. (1924) *Behaviorism*. New York: The Peoples Institute.
- Wegner, D. M. (1989) *White Bears and Other Unwanted Thoughts: Suppression, Obsession, and the Psychology of Mental Control*. Penguin.
- Wegner, Daniel M. (2002) *The Illusion of Conscious Will*. MIT.
- Wolpe, J. (1981) The dichotomy between classically conditioned and cognitively learned anxiety. *Journal of Behavior Therapy and Experimental Psychiatry*, 12, 35-42.
- Zener, K. (1937) The significance of behavior accompanying conditioned salivary secretion for theories of the conditioned response. *American Journal of Psychology*, 50, 384-403.

## Notes

---

<sup>1</sup> I thank R. Phoebe Dunkle for extensive help researching and preparing this chapter.

<sup>2</sup> Berridge (2003) has reported behavioral and neurophysiological evidence that seems to require a kind of reward that is not pleasurable, which he calls “non-hedonic”.

<sup>3</sup> Edward Tolman’s original concept of vicarious trial and error, in which a subject estimates the reward for alternative choices by serially initiating them without committing to them (1939) has recently been validated with single hippocampal neurons (Johnson & Redish, 2007).