

De Gustibus Disputare:
Hyperbolic delay discounting integrates five approaches to impulsive choice

George Ainslie
Veteran Affairs Medical Center
Coatesville, PA, USA 19320
and
School of Economics,
University of Cape Town
Rondebosch 7710, South Africa

George.Ainslie@va.gov
ga@picoeconomics.org
1 610 384-7711 ext. 4260

From corrected proof, *Journal of Economic Methodology* 24(2), 166-189, 2017
<http://dx.doi.org/10.1080/1350178X.2017.1309748>

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, and is thus not subject to copyright. The opinions expressed are not those of the Department of Veterans Affairs or of the US Government.

Disclosure statement

I have no potential conflicts of interest.

Abstract

Impulsiveness is the tendency to shift preference from a better, later option to a poorer, earlier option as the two get closer. Explanations have extrapolated from five piecemeal elements of psychology—failure of cognitive rationality, Pavlovian conditioning, force of habit, incentive salience, and long chains of secondary reward—but in doing so have created models that stretch the properties of these elements as observed in the laboratory. The models are hard to integrate with each other, much less with the utility-maximizing structure of microeconomics. The increasing prominence of addictions, particularly addictions that do not involve a substance, have confronted behavioral science with these shortcomings. I suggest that the roles of all five phenomena follow from the hyperbolic discounting of expected reward:

Rationality can be seen as maximizing expected discounted reward as seen at a distance, which can be approximated by making choices in bundles; Pavlovian conditioning reduces to reward learning when a very short term rewarding effect of conditioning stimuli is factored in; “habit” subsumes various kinds of reward-seeking, especially that which has been narrowed by failures of self-control; incentive salience is a kind of short term reward; self-reward is possible—and probably common—without being secondary to inborn turnkeys, limited rather by not being occasioned so often that the appetite(s) for it fail to build up.

Keywords

Impulsiveness, addiction, hyperbolic delay discounting, intertemporal bargaining, endogenous reward, artificial intelligence, Pavlovian conditioning, habit, incentive salience, idleness aversion

Text

With good reason economists have regarded human tastes as imponderable. Of course, the population of internal components that interact to form preferences have to be analyzed on a different scale than literal populations, the stuff of microeconomics (Ross, 2014). However, the problem is not just one of scale, but of a seeming incompatibility of concepts. As recently as half a century ago this incompatibility was not recognized. The sciences of choice all assumed that people's decisions naturally followed something like expected utility theory (EUT).. To describe even normal decision-making-- much less deviations from it-- without hard data on internal factors, the field of psychology has had to extrapolate from the findings of diverse methodologies. The resulting picture is not a black box, perhaps, but a semi-opaque box full of levers and gears. The edges of the mechanisms that are visible by experiment have inspired models of human tastes, with increased urgency as we realize the scope of maladaptive tastes-- most glaringly for addictions. People are innately impulsive—that is, we have a stubborn tendency to form temporary preferences for options we would not choose on overview, and even which we expect to regret. The good news is that several models have suggested what goes wrong with choice. The bad news is that these models have been stretched well beyond the properties seen in the laboratory, and are hard to relate to each other.

Five principles have been proposed as bases of impulsive choices. A person might have shifting cognitions about the outcome contingences; she might have reflexively created, “conditioned” urges; she might be locked in habit; her choice might be hijacked by the nonhedonic “incentive salience” of an option; or she might be misled by secondary rewards that do not predict the outcomes accurately. I will show how these principles may be understood as products of a single economic mechanism—selection by hyperbolically discounted reward—and suggest how this understanding might affect treatments for impulsiveness. The reward in this approach has its roots in the operant reinforcement of B. F. Skinner and R. J. Herrnstein (Herrnstein, 1997)), but the hyperbolic discounting first demonstrated by their methods reveals a more dynamic scope than was first imagined, so the effect of reward no longer stops at the borders of the five other models—or, in the case of the fifth model, no longer depends on external turnkeys..

Limitations of the five models

Failure of cognitive rationality

It was behavioral economics that first began to use psychological findings to test the validity of EUT, which the sciences of choice had assumed to be basic. It had two approaches, cognitive and motivational. The cognitive half of behavioral economics was founded on subjects' responses in thought experiments where they shifted valuations, sometimes radically, if a constant set of alternatives was framed differently— for instance with different reference points or as losses versus gains.(reviewed in Ainslie, 2015). Many examples have been described by prospect theory (Kahneman & Tversky, 1979). These shifts have often been attributed to cognitive errors, aberrations in a naturally rational valuation process. A large proportion disappear “at equilibrium,” when subjects have thought them over (Slovic & Tversky, 1974) or been debriefed (see Gigerenzer, Fiedler, & Olsson, 2012). That finding suggested that a person's susceptibility to such shifts would respond to teaching. However, there seemed to be no general principle underlying the shifts, leading one commentator to characterize the field as just “economics minus the assumption that people are rational maximizers of their satisfactions“ (Posner, 1998, p. 1552). “It would not be surprising if many of these phenomena turned out to be

unrelated to each other, just as the set of things that are not edible by man include stones, toadstools, thunderclaps, and the Pythagorean theorem” (*ibid.*, p. 1560).

The motivational half of behavioral economics came from experiments showing inconsistent preferences over time. From Socrates until the nineteenth century the standard of rational choice was to go by the outcomes’ objective values, regardless of delay (Plato, 1892; Jevons, 1871/1911). Subsequent authors had mostly accepted as rational the discounting of delayed outcomes, as long as it did not lead preference to shift as delays shorten—that is, in an exponential discount function, as Samuelson almost casually described it (1937; Grune-Yanoff, 2015). However, since the 1970s subjects have commonly been found to report temporary preferences for smaller, sooner (SS) outcomes over larger, later (LL) ones as delays to the pairs shortened (Kirby, 1997); These findings were the beginnings of the approach I propose; but it is enough to note here that they are not the product of cognitive error. There have been no studies of whether such subjects change their reports after being confronted with consequences such as inconsistency and the potential for being money-pumped. However, in daily life complaints about having low willpower are common, implying that the limiting factor in intertemporal consistency may not be cognitive awareness. The temporary preference effect seems to have the staying power of a Gestalt illusion that cannot be overcome by insight, as can, for instance, the initial impression that distant objects are smaller. In both human and nonhuman subjects the temporary preference for SS rewards has been a robust finding, in contradiction to EUT.

Conditioned responses

An obvious explanation for why someone does not learn to maximize reward is that her behavior is not governed by reward—that is, it is pushed by preceding stimuli rather than pulled by prospective outcomes. The two common examples of this model are Pavlovian conditioning and habit. Conditioning is often seen as imposing impulsive choices on people by reflexively creating appetites or emotions (Tiffany, 1995, O’Brien, Ehrman, & Ternes, 1986). It was once thought to govern behavior directly—so the clinking of ice in a glass might elicit drinking behavior, for instance—but has since been shown to govern only information (Dolan & Dayan, 2013; Huys, Guitart-Masip, Dolan, & Dayan, 2015; Rescorla, 1988): What is pushed by preceding stimuli is not behavior but expectation, the common experience of associative memory. However, psychology has never worked out whether visceral responses such as appetites and emotions are intrinsic parts of the expectations that induce them, or are separable behaviors motivated by these expectations. Responses that cannot be emitted deliberately have usually been assumed to form part of expectation, even though there is reason to think they are behaviors just like other kinds of response: Conditioned responses are plastic, rather than being like Windows images dragged from one site to another; they differ in detail from the originals, a change that must have been taught afresh (Rescorla, 1988; reviewed in Ainslie, 2001, pp. 20-22). They sometimes become simply voluntary, as when a bulimic learns to vomit without gagging herself or even imagining gagging herself. They compete with, and can be overridden by, motivated responses (Ainslie & Engel, 1974). They sometimes even vary between opposite responses, such as a high versus the start of withdrawal (O’Brien, Ehrman, & Ternes, 1986). Furthermore, appetites/emotions are often observed to respond to a stimulus in the absence of further pairings-- even to get stronger, the Napalkov effect (Eysenck, 1967). They often arise in the absence of a stimulus. Changes to an emotion-evoking memory with rehearsal are not conscious and are thus usually assumed to be passive—“reconsolidation” (Lane, Ryan, Nadel, &

Greenberg, 2015)—but the urge to rehearse them is obviously a motive. The point is that something more than learning of information must be happening when *responses* are conditioned. Craving cannot be based on mere association—Something must be pulling it. That is, the appetite/emotion must be *rewarding* or *rewarded*. Pavlovian conditioning is the name of a pairing that elicits visceral responses, but it only begins to explain how such responses are selected.

Students of addiction continue to accept appetite as passively driven by stimuli, without worrying about these dynamic properties. However, theorists have sometimes suggested that conditioning stimuli actually function as rewards. They have pointed to the longstanding observation that conditioning can be observed only when the conditioning stimulus—pain, food, drug—has motivational value (Miller, 1969), and they have built computer models in which the same selective factor operates both in obviously rewarded learning and in conditioning (Donahoe, Burgos, & Palmer, 1993). However, conditioning has seemed necessary to govern appetites and emotions, because they are often negative and thus in need of a mechanism for intruding on attention. The problem for an all-operand—that is, motivational-- model in such cases is how to mix negative and positive components without simply averaging them, which would cause them to cancel each other out rather than combining attraction and aversion. This, I suggest, is the crux of why psychological theory has retained two otherwise similar selective mechanisms: one that is only attractive and one that forces you to entertain painful experiences.

Force of habit

The persistence of addictions despite deteriorating reward has suggested the relevance of behavioral experiments on habit—the persistence of a choice after it is no longer rewarded. Overlearning of a routinely rewarded response is accompanied by a shift of neural activity in midbrain striatal areas from “planning” or “voluntary” to “habitual” systems (Everitt & Robbins, 2013). A similar shift has been described from “goal-directed” or “model-based” to “model-free” systems (Voon et al., 2015). Such shortcuts in response to regular reward could be called *routine habits* (Ainslie, 2016a). Some kinds of brain lesions make rats unable to change overlearned responses when reward contingencies change, and addiction to stimulants has been accompanied by similar changes in human brain activity, leading to the suggestion that these midbrain changes are also responsible for *bad habits*, that is, repeated impulsive choices (Everitt & Robbins, 2005). However, although addicts have been observed in choice-making tasks to respond less well than nonaddicts to changed contingencies of reward, this difference has usually been moderate, as it has been even in patients with gross brain lesions (Fellows & Farah 2005). Furthermore, it is of questionable relevance to their choices of whether or not to consume substances; addictive “habits” have very little to do with mindless repetition, but on the contrary require a high degree of flexible, goal-directed behavior to evade a hostile society.

Incentive salience

Another functional dichotomy besides reward-responsive/habitual has been reported, also based on observations of the midbrain striatum. Dopaminergic agents such as cocaine can become “wanted” on the basis of their “incentive salience,” an effect that may persist despite not being followed by a supposedly true reward, the “liked” activation of the brain opiate system (Berridge & Kringelbach 2008). This effect has been produced in both nonhumans and human patients (Heath, 1972) by stimulating the dopaminergic areas of their ventral striata. Once stimulation

has started the subjects work vigorously to repeat it, but do not show facial signs (rats) or give reports (patients) of pleasure. Moreover the subjects often fail to start self-stimulation again once it has been stopped for a period. The areas of the midbrain that are involved are roughly the same as those that govern the habit (model-free) system. Evidence from local ablations in non-humans and dopaminergic neuron loss in Parkinsonian patients suggests that reward indeed has two separable components, wanting and liking, such that wanting leads to action as just described, but liking does not lead to action unless wanting also occurs. The incentive salience that leads to wanting has been regarded as “nonhedonic” (Berridge 2003)--unrelated to reward. But the obvious question, then, is how do wanted options compete with others for adoption? Is incentive salience a third selective factor, after operant reward and conditioning stimuli?

Long chains of secondary reward

The closest analog of economists’ and philosophers’ EUT is the maximization of reward/minimization of punishment, which behavioral psychology has studied in detail. In the laboratory choice is a highly regular function of consequent events, *primary rewards*, the expectation of which is necessary to give value to all events that do not have the power to reward in their own right-- *secondary rewards*. (This is Pavlovian conditioning in the modern, narrow sense of associative learning.) This model is assumed to describe motivation in more complex human situations as well, an assumption that has rarely been reexamined. Even for humans, “A primary reinforcer derives its strength from an innate drive, such as hunger or sexual appetite; a secondary reinforcer derives its strength from learning” (Wilson & Herrnstein, 1985, p. 45; see also Baum 2005, pp. 77-86). However, in the laboratory secondary rewards that stop predicting the occurrence of a primary lose their rewarding power—*extinguish*. Since most non-addictive rewards in modern life are abstract and chosen arbitrarily by individuals or societies, this theory requires long chains of association or deduction, leading to primary rewards at such long delays or such low probabilities that their power ought to have been largely lost.¹

The list of obvious primary rewards-- outcomes that are hardwired to reward-- is rather short: food, drink, sexual stimulation, bodily comfort, sleep, pain relief, perhaps some kinds of exercise—and of course drug effects. Direct access to these is readily understood as inviting impulses; but we are increasingly aware of impulses that themselves apparently lack primary rewards (Zhang et.al., 2012). It is hard to model non-concrete rewards that are not clearly secondary without least an hypothesis about what determines their value. There has been little movement in this direction.

The paucity of primary rewards has been used as an argument against reward-based models of human behavior (e.g. Deci & Ryan, 1985, pp. 179-189). Recently roboticists have taken up this challenge to model “intrinsic” primary rewards that do not depend on external input-- but they still conceive of these rewards as inborn, “inherently interesting or enjoyable” (Oudeyer & Kaplan, 2007; Ryan & Deci, 2000; Singh, Lewis, & Barto, 2009). Even this extended reward theory has to stretch to account for the vast range of human tastes, especially when these tastes

¹ Social psychologists sometimes group motivations to get higher-order outcomes into categories such as “needs” (e.g. Atkinson & Raynor, 1975), but without thought to their motivational mechanisms.

intensify during their pursuit, sometimes to the point of addiction. The variety of human rewards that do not lead to apparent primary rewards suggests that some rewards are *endogenous*—not only mental but also assignable, that is, independent of any inborn turnkeys of reward. This indeed implies that people can coin their own reward, perhaps leading to a positive feedback system—a worry that has probably deterred theorizing about this possibility. The internal reward process clearly has to be limited in some way, but not necessarily by the prospect of primary rewards.

Some recent reports illustrate this problem. There have been a number of experiments that elicit activity in brain reward centers (ventral striatum and medial prefrontal cortex) using small amounts of money, a supposedly secondary reward. In one, subjects chose between amounts of expected money ranging from \$0.01 to \$0.24 at delays ranging from 2 sec to 64 sec. The value of delayed amounts declined as a hyperbolic function of the delay, even though the money itself would not be delivered until the end of the experiment, and obviously could not be spent until even later (Wittman, Lovero, Lane, & Paulus 2010). As the authors pointed out, the substantial behavioral impact of these meaningless delays indicated that the reward announcements were valued for their own sake, like points in a video game. In a similar experiment, activity in the ventral striata of about half the subjects discounted real monetary winnings signaled by pictures; but the brains of the others did not show discounting at all, realistically reflecting the uniform delay to actual delivery (Gregorios-Pippas, Tobler & Schultz, 2009). The observation that half of subjects did not show discounting also suggests that this valuation did not occur through a passive process such as secondary reward. If an important amount of reward is not limited by predicting a primary reward (or another secondary reward), what does limit it?

A puzzle for econometrics

These five models accord with common assumptions about motivation, which are a similar potpourri: It is natural enough to equate reward with utility. Conventionally we have pictured this reward as something a person cannot coin by herself, but must get from outside. It often connotes only one of various influences on her ego, which is assumed to be a higher authority that might make choices on some other basis besides reward, such as logic, morality or even arbitrary preference—“free will.” We are also apt to see reward as the symmetrical opposite of undesirable states—pain, grief, anger-- the capacity of which to attract attention we must then ascribe to Pavlovian conditioning. Research has recently shown this attractive capacity to share properties with reward in some cases, leading the researchers to propose the hybrid concept of incentive salience (Berridge & Kringelbach, 2008). And subjectively, habit is a process without motivational properties. This picture must be a nightmare for the econometrician: a market dominated by option futures with tenuous bases in fact, not fully governed by price, invaded by demands that are either price-insensitive or based on a foreign, non-convertible currency, and subject to arbitrary restrictions that maintain the status quo. No wonder personal motivation is an economic no man’s land, not to be disputed about.

Basic properties of reward

A change in the way we understand reward could repair these ambiguities and anomalies. It is possible to draw a more parsimonious picture without contradicting any established facts of motivation. However, this picture will contradict most of the assumptions in the foregoing paragraph.

Reward is the ultimate determinant of choice

First we should recall that we are evolved organisms, and that our reward process evolved as a proxy for Darwinian adaptiveness—a genetic program of the incentives that lured our ancestors into leaving the most surviving children. Where that program has turned out not to work under modern conditions, for instance in motivating the use of opiates-- or of birth control!-- it may be corrected, but only in future generations. We are not motivated to maximize adaptiveness, only reward. Nor do logical, moral, or other cognitive processes *per se* have any necessary call on our choice-making. We are creatures whose only job is to forage for reward.

Reward is the process that selects among replaceable behaviors

Reward can be defined as that which makes the behaviors it follows more likely to recur. A behavior is subject to reward—*motivated*-- to the extent that it can replace other behaviors on the basis of its expected consequences, rather than depending on hardwired triggers as in a reflex. Of course, being expected implies being learned. The set of motivated behaviors is larger than might at first appear, including behaviors while sleeping (Granda & Hammack, 1961) and autonomic processes that can be modified by biofeedback (Schwartz & Andrasik, 2016). Brain imaging is now confirming the long-held suspicion that even perception is somewhat a motivated behavior (Pessoa, in press), so even chains of expectations laid down by Pavlovian conditioning may be modified purposively.

The boundary between motivated and hardwired behavior is sometimes difficult to discern, and may be diffuse—For instance, changes of heart rate that can be made voluntarily with biofeedback ride on top of autonomous excitation pathways in the heart, just as the voluntary steps of walking are superimposed on a hardwired crossed-extension motor reflex. Breathing, which might be thought of as occurring autonomously unless inhibited, actually depends on recurring motivation, as shown by conditions where this motivation fails, such as Ondine’s curse (Kuhn, Lütolf, & Reinhart, 1999). However diffuse the boundary may be, the set of motivated behaviors is clearly larger than the set of deliberate behaviors, so lack of voluntary controllability is not an argument against their being motivated. Importantly, the property of being motivated does not imply the absence of a neurological substrate, which obviously has to underlie all behaviors.

Rewards must be commensurable

Several authors have pointed out that there must be a common basis on which motivated behaviors compete for selection, and such a basis must exist between any two processes of which one can be chosen instead of the other (Cabanac, 1992; Montague & Berns, 2002; Shizgal & Conover, 1996).² Brain imaging is now beginning to visualize such a process, both in nonhumans (Chen, Lakshminarayanan, & Santos 2006; Glimcher, 2009; Johnson & Redish, 2007) and humans (2010; Levy & Glimcher, 2012). Various aspects of value—probability, modality, emotional salience, and others besides delay itself— may be computed in different

² Strictly this is what David Spurrett has called a “proximal currency,” part of the actual mechanism of choice, as opposed to “ultimate currencies” such as adaptive fitness (Spurrett, 2014). I would argue that all the proximal currencies that he reviews are consistent with the model proposed here.

sites, but they are observed to feed into a single final path (Carter et.al., 2010; Platt & Padoa-Schiappa, 2009). In such experiments we can see the competition of discrete external rewards taking place according to the principles laid down by the behaviorists of the last century. In a person who can make an arbitrary choice among a million options, each of the million must be ratable within a choice on this common basis. Although most possible choices never confront each other—such as betting on this poker hand versus voting for this political candidate—they are parts of a web of choices in which you can get from any one to any other, for instance by choosing whether to go to the game or the polling place. On the way, this web must potentially include the urge to get angry at another driver or pay attention to a sore toe. Nor is there reason to believe that this web excludes complex intellectual activities, such as searching for mathematical proofs or deducing moral certainties. As Spinoza said, "No affect can be restrained by the true knowledge of good and evil insofar as it [the knowledge] is true, but only insofar as it is considered as an affect."—quoted in Hirschman, 1977, p. 23.). The implication is that all choices, including commitments to protect certainties and “higher” values, must compete to survive in an internal marketplace based on reward.

Delayed prospective rewards are discounted hyperbolically

In interpersonal markets the value of delayed goods is normally discounted in an exponential curve.

$$\text{Present value} = \text{Value}_0 \times \delta^{\text{Delay}} \quad \text{Formula 1}$$

where Value_0 = value if immediate and $\delta = (1 - \text{discount rate})$.

That is the only mathematical formula that will keep an agent from changing her preferences as goals separated by a fixed interval get closer, which would make her vulnerable to money-pumping (Conlisk, 1996). In a person’s internal marketplace, however, brain imaging has corroborated behavioral evidence that delayed outcomes tend to be evaluated like other psychophysical qualities, in a hyperbolic function of the stimulus in question, in this case delay (Kable & Glimcher, 2007).

$$\text{Present value} = \text{Value}_0 / [1 + (k \times \text{Delay})] \quad \text{Formula 2}$$

where Value_0 = value if immediate and k is degree of impatience (Mazur, 1987).

This evaluation pattern creates a limited warfare relationship among an individual’s successive motivational states, leading to inconsistent preferences over time and incentives to make commitments against this inconsistency. A person’s skill at such commitment determines her self-control and consequent rationality, including her sometime exponential pattern of discounting prospective rewards (Ainslie, 2001, pp. 100-104, 2005).

In hyperbolic discounting we can discern the elementary process of reward that makes human behavior patterns diverge from consistent utility-maximizing. This shape is especially important in environments where hardwired rewards have become cheap, as in modern civilization. The most verifiable effects have been described extensively elsewhere (Ainslie, 2001, 2005, 2012), and I summarize them only briefly. I then describe how the implications of hyperbolic

discounting suggest a coherent rationale for the five kinds of phenomena described above, in a single marketplace of behavior with a single currency of reward. This marketplace in turn determines the demand for external goods seen in microeconomics, including substances of abuse and addictive activities that do not involve a substance. Where lack of hard data limits this proposal to a conceptual framework, I assert that the parsimony of this framework makes it the most likely hypothesis about what cannot yet be measured.

A review of hyperbolic delay discounting.

The mental processes that a person learns in order to get a reward can be called its interest, by analogy to economic interests in society. Hyperbolic discounting allows various longer- and shorter-term interests to survive in durable competition with each other, as long as each sometimes gets the reward on which it is based. People and other vertebrates will temporarily prefer an SS reward to an LL one as the SS reward draws closer (*dynamic inconsistency*), unless they have learned to forestall this change. The means of forestalling might be chosen before the SS reward becomes dominant, and thus involve adopting a physical or social commitment, diverting attention, or exciting a contrary emotion. However, such means are often awkward, costly, or unavailable before the SS prevails.

Alternatively, if a person arranges to make a whole series (or *bundle*) of SS/LL choices at once, the relatively flat tails of hyperbolic curves should add up to increase her differential incentive to choose LL options, as in Figure 1. This mathematical prediction has been borne out experimentally (Kirby & Guastello, 2001; Ainslie & Monterosso, 2003; Hofmeyr, Ainslie, Charlton, & Ross, 2010). The necessary bundle may be formed by discerning—or defining—a variant of repeated prisoner’s dilemma among these expected choice-makers, as in limited warfare: Defection in the present case makes defection in future cases more probable, not from a motive of retaliation (which would be present in a classical repeated prisoner’s dilemma) but by making present cooperation seem likely to be wasted. The stake of LL reward may be aggregated from the evident consequences of each choice, as in binges with hangovers, or the stake may be a more distant anticipated condition such as good health or adequate savings. The necessary element for the self-prediction process is that the person’s expectation of getting a category of LL reward as a whole is put at stake when each opportunity within a definable set of SS rewards occurs.³ This self-prediction process is *recursive*, in that each estimate of future self-control is fed back into the estimating process.

³ This mechanism provides a rationale for the motivational force of goals in the “goal-setting” literature (e.g. Koch & Nafziger, 2011), which is otherwise unaccountable.

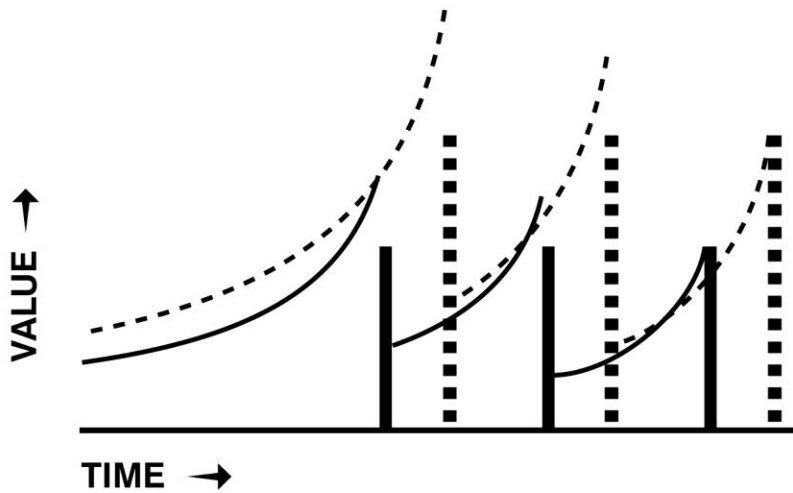


Figure 1. Summation of the expected values of a bundle of three choices between an SS and LL reward, made all at once: LL will be preferred at all choice points in advance of the first pair. If the choice were only between the last pair, the SS option would come to be preferred.

The *intertemporal bargaining* that creates bundles of rewards gives long range interests their greatest weapon, one that matches common descriptions of willpower. Such bargaining also converts the battle between long and short term interests from sequential dominance to a lawyer-like competition of rationales, to define whether a current gratification represents a defection. The winning bargains function as explicit or implicit *personal rules*. These rules are most stable when their criteria stand out from other potential criteria, separated by *bright lines* (Ainslie, 2001, pp. 94-100).

Personal rules are compromises to stabilize the bargaining of long and short term interests, and as such they are apt to be brittle. In normal self-control they are added to social commitments—leaving yourself open to the influence of the people around you, and selecting those people partly with an eye to what that influence will be. Over-reliance on personal rules creates an incentive structure that is sometimes perverse:

- Lapses cause bigger losses of expected reward, so a person may stop attempting self-control in areas where they have happened, in order to protect her personal rules in other areas. This abandonment leads to circumscribed areas of dyscontrol: “I can never talk in front of a group, give up smoking, etc.”
- The potential damage from lapses creates an incentive not to see them, giving rise to a motivated unconscious: suppression, repression, denial.
- Choices may become more important as test cases than for their own sakes, making it hard to live in the here-and-now. Awareness that damaging a personal rule may destabilize a larger network of intertemporal bargains may lead to compulsiveness—giving too wide a berth to risks of lapses.

A way to unite the five models

A hyperbolically discounted reward process with these properties offers a unified framework for the five mechanisms that have been derived from different research traditions. I will argue for:

- defining rationality in light of hyperbolic delay discounting of reward
- deriving supposedly conditioned visceral responses from reward
- rejecting the notion that habits become significantly independent of reward
- interpreting incentive salience as a form of short term reward
- recognizing that people often coin reward endogenously, without expecting hardwired events.

Cognitive rationality requires consistent choice

Hyperbolic discounting makes impulsiveness the state of nature. Rewards are seen most in proportion to their objective sizes when evaluated from a distance. Thus a practical definition of rationality might be whether an option is preferred from the viewpoint of distance. This viewpoint might be approximated by bundling rewards through intertemporal bargaining,⁴ except that rigid bundling begins to depart from reward maximization again—as just described. The robust “anomalies” reported in prospect theory can mostly be traced to the strategic management of temporary preferences made necessary by the hyperbolic shape of the discount curve; however, a weighting of losses more than gains may be innate (Ainslie, 2015).

Learning the cognitions that evaluate rationality itself must be subject to reward, and the resulting routines must be hired or not in the service of one interest or another. Examples include arguments to claim exemption from personal rules by short term interests; testing of what plans are consistent, realistic, or moral by long term interests; and building a narrative that will stand up to questions—and questioning it—by either kind of interest. Cognitions that are testably true but less desirable are left by the wayside, as when people learn *ad hoc* heuristics to solve problems in preference to logically solid means such as Bayes’ theorem (Gigerenzer, Fiedler, & Olsson, 2012).

Conditioned responses depend on reward

The shape of hyperbolic discounting offers a solution to the problem of how negative appetites and emotions can bargain for a person’s involvement.. It permits cycles of strong, brief reward and relatively longer, consequent periods of reward inhibition, creating an experience that is vivid but aversive. That is, aversive mental processes such as painful emotions and intrusive memories may involve the same sequence of attraction and cost that characterize addictions and itches, but alternating too rapidly to be discriminated (Ainslie, 2001, pp. 51-61; 2009).⁵ Some other combination that is not literally sequential is conceivable, but the fact remains that negative experiences have to attract our attention. That they sometimes fail, and can be resisted

⁴ Options in the far future are foreseen more in imagination than from experience, but still tend to be discounted hyperbolically (Ainslie, 2016a; Green, Fry, & Myerson, 1994). Thus we might expect a cash payment delayed for six years to become temporarily preferred to a much larger payment delayed eight, and so need a personal rule—for instance to discount money exponentially—for it to be valued consistently over time.

⁵ As a model it has been possible to get pigeons to actively avoid schedules of tiny, immediate rewards (Appel, 1963), but no neurophysiological example of very brief reward alternating with nonreward has yet been created.

deliberately as in some obstetric and dental programs (Melzack, Weisz, & Sprague, 1963) implies a market competition. A painful stimulus may make us an offer we can almost never refuse, but it is still an offer and sometimes we learn to refuse it. The observation that responses to a stimulus can grow over time is also evidence against their being reflexive (see the ensuing endogenous reward section). Furthermore, the incentives are often not so one-sided. Subjectively, if another driver cuts me off in traffic I consciously weigh the urge to get angry against the attraction of whatever I had been thinking about. This kind of experience is well depicted in the recent animated film *Inside Out*. Inside the heroine's head a personified Sorrow and other negative emotions do not coerce the ego-like Joy, but pester for her attention. These emotions are not reflexive, but entreat for acceptance on the same basis as positive ones, implying a rewarding component.⁶

The important point is that hyperbolic discounting makes the conditioning account of appetites/emotions unnecessary to explain the urge to attend to them. The implication most relevant to temptations is that motivated appetites can take part in recursive self-prediction, accounting for such phenomena as explosive craving after mere reminders of consumption (discussed in Ainslie, 2010 and 2016b). Thus if you are confident of not giving in to an appetite it will stop arising; whereas if you see signs of wavering, the appetite will strengthen. Such confidence usually requires a long period of consistent abstinence, but can be created by an iron-clad personal rule (Dar, Stronguin, Marouani, Krupsky, & Frenk, 2005) or radical reinterpretation of options (Miller & C'de Baca, 2001). By contrast, counter-conditioning does not keep addictive appetites from arising (Conklin & Tiffany, 2002).

Habit is not a selective factor

At least four phenomena may lead someone to continue an addictive behavior after the pleasure it gives would no longer seem competitive. The most obvious is that alternatives may have deteriorated as well, so the addictive choice is still the most rewarding. The second phenomenon is the relative preservation of reward at the beginning of addictive activities as they deteriorate, keeping them competitive for a hyperbolic discounter despite low average reward over time. Part of this front-loading may be what has been described as incentive salience, as I discuss in the next section. Third, the challenge of getting illegal substances may have created game-like activities that are rewarding in their own right, the same pattern as seen in shoplifting and other seemingly unrewarded addictions. I discuss this possibility in the next section but one. Perhaps most importantly, after repeated failures of intertemporal bargains, a person may have learned not to attempt self-control in this circumscribed area, as described above. This is a failure of intertemporal bargaining, but is apt to be described as habit:

Urges may feel negative but hard to resist (to panic, to attend to an obsession), or they may be consciously tempting (to use drugs or get into destructive relationships), but for all of them we face the choice of giving in or trying to control them. We monitor our attempts to control urges with recursive self-prediction, and in doing so create a history of successful and failed

⁶ I would have suggested a Cookie Monster-like character as well, to represent a consumption interest. Unlike *Inside Out*, most films that imagine the inside of the head portray it as a set of faculties manipulated by an ego-type figure, like an engineer running a complex robot from a control room-- Daniel Dennett's "Cartesian theater" (2003, pp. 122-126).

commitments that constrains subsequent intertemporal bargains. Our personal rules are often implicit— we sense the extra significance of the choice but cannot articulate what is at stake. Lapses damage our confidence in our intertemporal cooperation, engendering guilt and leading us to abandon attempts at self-control in areas where it has failed. Thereafter we avoid the kind of situation where we were overwhelmed, concluding that we “can’t face embarrassment,” “can’t resist chocolate,” or “can’t stand heights,” thus establishing a circumscribed symptom or “bad habit.” To restore intertemporal cooperation we redefine our rules with rationalization, and we develop repression and denial to avoid recognizing lapses. Over time we accumulate commitments and failures of commitments that make us rigid in much the way old economies or bureaucracies become rigid (Olson 1982).

“Habit” describes a repeated activity for which many motives may be responsible, not a selective mechanism in itself. The narrowing of choice that comes from failed intertemporal bargaining, in particular, has been the major concern of dynamic psychotherapies, which target misguided and overgrown attempts at self-control: “cognitive maps” (Gestalt), “conditions of worth” (client-centered), “musturbation” (rational emotive), “overgeneralization” (cognitive behavioral), and of course the punitive superego (summarized in Corsini & Wedding 2011).

Incentive salience is very short term rewardingness

The system(s) associated with attention, initiation, and the apparently short term reward of behavior (“wanting”) can be experimentally dissociated from the basis of more stable preference (“liking”—Berridge, 2003; Berridge & Kringelbach, 2008). However, to the extent that wanted processes compete for adoption with liked alternatives, they must obviously have a reward value themselves (see Ainslie, 2009). The sharp decline in the attractiveness of midbrain striatal stimulation with distance suggests that the rewarding effect of salience is very short term. As described in the foregoing “conditioning” section, hyperbolic delay discounting can produce short term reward by an option that is poorly rewarding on the average, including, hypothetically, pains or other emotions where a cyclical reward is so brief that it cannot be discriminated within an urgent, negative experience.

Such a cycle may also play out with longer, consciously discriminable periods, as tolerance to an addictive activity develops. The activity does not decline into frank aversiveness, but delivers a decreasing duration of pleasure or relief—the smoker who lights cigarettes and then stubs them out, the opiate addict whose high turns stale rapidly, the overeater who must nibble-- in what may look like robotic repetition. The mean level of reward is low, but it is still front-loaded, causing a hyperbolic discounter to seek it. Furthermore, if you resist the urge, it recurs repeatedly, like an itch. If you don’t believe that you can abstain enough times to make the urge extinguish, the effort may seem as though it will be wasted. There may be many cases where a “wanting” component endures while a “liking” component fades, and these may be related to the properties of different neurotransmitter systems. It has been suggested that changes in midbrain dopaminergic neurons that accompany some addictions make them unresponsive to differential reward (e.g. Kalivas & Volkow, 2005). However, as with “habits” that are supposedly nonhedonic, wanting must still compete via the final common selective factor of reward.

Chains of secondary reward are often chimerical

Not only are people rewarded by outcomes with no obvious connection to primary rewards, but our activities typically build on themselves to become elaborate projects to which we attach great significance—feats of personal mastery, sports championships, religious revelation, hobbies, fandom. A soft currency rewarding game-like activities might seem to need backing by a hard currency that is outside a person's control, lest she short-circuit the reward process and reward herself *ad lib*. However, the hyperbolic discount function permits a model of primary reward in imagination that depends not on realism but on other factors that also narrow the person's field of arbitrary choice.

The properties of reward outlined above were discovered largely with the methods of the behaviorists. To use these properties in modeling the behavior of imagination we will have to abandon their discipline of avoiding mental constructs, but otherwise their description of the reward process serves us well.

When you can consume rewards *ad lib*, the limiting factor is appetite. Motivational science has been poor with superabundant goods. Economic goods must be in limited supply relative to demand, which is what creates value. As Adam Smith originally observed, that is the rationale by which air has less market value than diamonds, although air is more necessary. A factor best called *appetite* not only limits the rewardingness of freely available goods, but it also produces reward most when it is allowed to build up. Think of hunger in the case of food, refractory time in the case of sex, or deprivation in the case of sleep or cigarettes. These are physical appetites, but they can serve as models of subtler ones—for puzzles, fiction, or mastery of an unnecessary task. When we seek reward from an appetite, we have an urge to harvest its potential prematurely, just as hyperbolic delay discounting predicts. Building appetite requires restraint or opposition. We pace ourselves by social conventions or by personal rules. We can snack *ad lib*, but mostly we learn to wait until meal times or even “work up” an appetite. Teenagers learn to pace sexual gratification, first with fiction, then by having a partner. Nordics have learned to create appetite for temperature change with saunas. Heroin addicts learn to withdraw voluntarily to restore their appetites. Means to build appetite become economic goods: appetizers, anaesthetics to delay orgasm, saunas. Conversely, where an appetite such as hunger distracts us from a more valuable goal, we learn to subdue it by eating whenever a little bit appears—“grazing.”

The same contingencies operate with reward in imagination—*endogenous reward*-- which is also in unlimited supply. However, it is harder to define what the appetites for endogenous reward are. Whereas physical appetites can be distinguished from each other by independent satiability, the appetites for endogenous rewards are less distinguishable. They are somewhat apparent in the case of the emotions, but there is no useful way to assess satiety for humor versus tragedy, for instance, or wish fulfillment versus rehearsing past satisfactions.⁷

Where rewarding events are limited by their availability, our realized reward is proportional to the efficiency with which we get them. But to the extent that a reward is endogenous or

⁷ However, it is important to distinguish between *fatigue*—loss of appetite for a particular modality from recent consumption—and *habituation*—permanent devaluation of a goal because it is fully anticipated, such as the punch line of an old joke.

otherwise available at will, learning to be efficient reduces appetite. This can be a factor in limiting the life satisfactions of advanced societies. As S. S. Tomkins put it: “The paradox is that it is just those achievements which are most solid, which work best, and which continue to work that excite and reward us least. The price of skill is the loss of the experience of value-- and of the zest for living” (Tomkins, 1978, p. 212).

The marketplace for endogenous rewards: One model

When we seek purely endogenous rewards, a smaller, sooner harvesting of appetite competes with a larger, later one in the familiar SS/LL format. The properties of mental appetites are largely unknown. The most effective factors in building appetites are surprises and challenges—risk, pursuit, opposition—but these are many and varied. Rather than attempt to model these, I assume for simplicity that appetite both builds and is consumed linearly, after which it builds again (Figure 2a). The reward from expected consumption (stippled triangles, height x duration) is discounted hyperbolically for delay. Expected discounted reward competes with alternatives (not shown), and prevails when it rises above the prevailing market level set by those alternatives (horizontal dashed line). In this example, the summed areas under the descending consumption curves total 6.75 arbitrary units ($1.5 \times 2.8 \times 0.5 \times 3$).

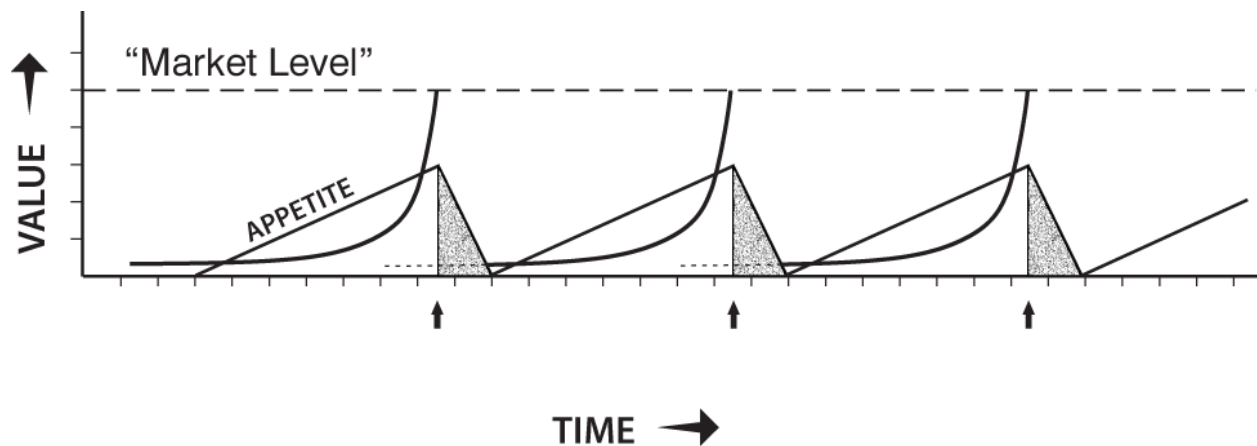


Figure 2a. Discounted reward value of the consumption from an appetite (stippled area under descending line), to start when the aggregate value of this consumption reaches a competitive (market) level. Appetite rises again as soon as it is consumed. The episodes are independent.

If some factor enforces a delay from the moment of choice (when discounted value reaches market level, figure 2b) until consumption begins, appetite builds higher, and the total consumption area is 9.45 ($2.25 \times 4.2 \times 0.5 \times 2$).

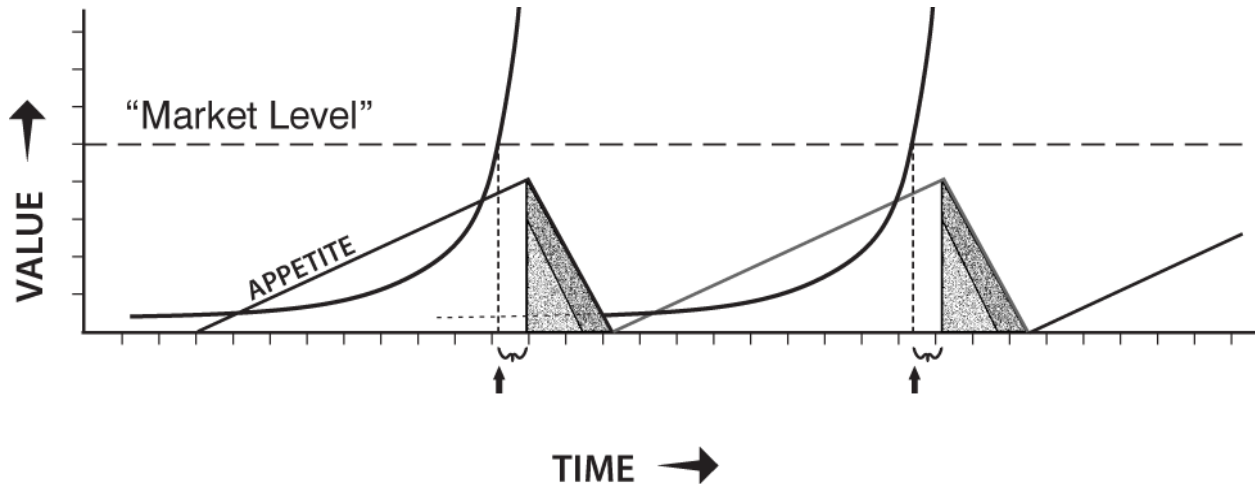


Figure 2b. Additional discounted reward when consumption is delayed from the moments of choice (darker stippled area under descending line). Consumption starts after the bracketed amount of time from the moments of choice, which occur when the discounted value of the consumption reaches competitive (market) level.

Different times of consumption may compete like any other SS/LL pair (figure 7c).

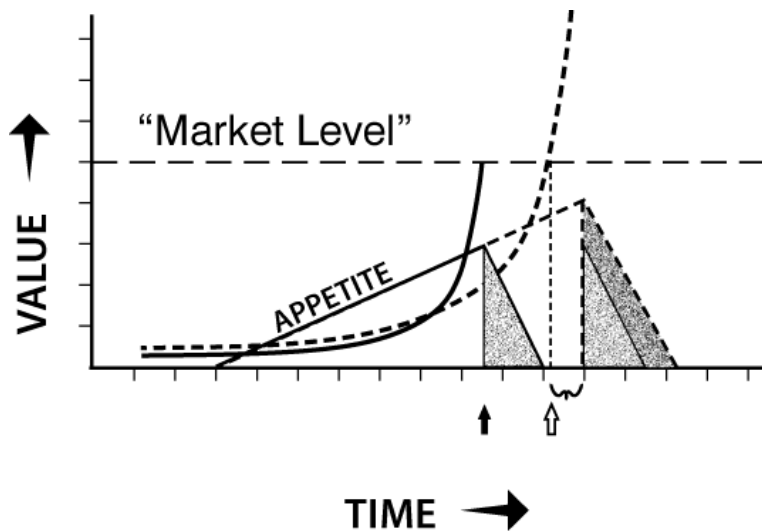


Figure 2c. Competition of the discounted values of the consumption patterns in figures 2a and 2b. An obstacle that imposes a delay after choice (bracket) makes the expected reward grow higher before its discounted value reaches market level (at the open arrow); such an obstacle would be chosen at any time before the discount curves cross.

Figure 2d depicts a still less rewarding pattern: letting attention wander in the absence of good sources of appetite, perhaps checking your watch or odometer when bored-- the equivalent of grazing. Small bits of curiosity capture attention and are trivially satisfied. Assuming a constant

linear rate of increase in appetite, the area under the triangles sums to only 0.54 ($0.225 \times 0.42 \times 0.5 \times 17$).

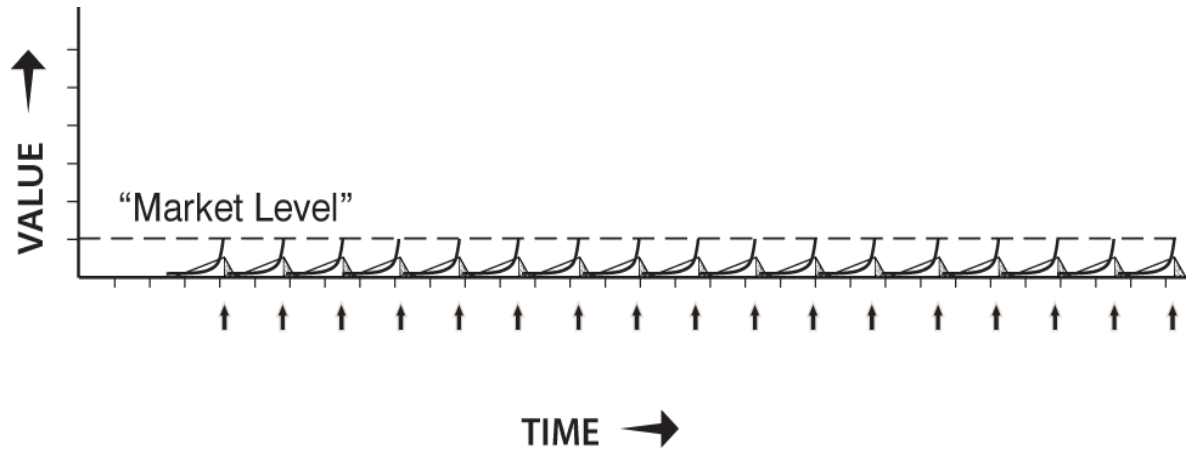


Figure 2d. Value of reward *ad lib* when there are no valuable alternatives (low market level). Discounted reward value of the consumption from an appetite (stippled area under descending line), starting when the aggregate value of this consumption reaches a competitive (market) level.

In this model, maximizing reward depends on finding challenge or being committed to delay. As you grow up, the poor returns for high frequencies of harvesting appetite probably shape attention skills that avoid the shortest range harvesting patterns, even in infancy or toddlerhood. Fantasy and play become complex, but nevertheless often compete decreasingly well with more challenging outcomes. Some temptations to premature consumption respond to deliberate self-control, such as by putting your watch aside when bored to keep from looking at it repeatedly, or making personal rules not to cheat at solitaire or read ahead in a story. Where the harvesting depends on physical objects such as watches, books or cards, personal rules against shortcuts are easy to enforce. Where the harvesting is entirely mental, ways to pace it are more conjectural. Rules for such harvesting necessarily begin in discerning criteria for payoffs—*occasions*-- that are either external or hard to create at will. Adoption of such criteria might best be called *betting*: on occasions in a book or card game, on whether a sports team will win, or on whether you can master a task—which can be a mental task, as long as benchmarks for mastering it are clear. It could be whether one stick will float under a bridge before another (“Pooh-sticks”), or whether you can predict which will be first, or whether you can throw it in a way to make it be first. It could be whether you can make the team, or whether the team wins, or whether you can help the team win.

You set up endogenous rewards in imagination by the same process as in daydreaming, but letting them grow by betting them on infrequent and unpredictable outcomes. At some level of involvement you care about these outcomes. Caring is something you are said to *do*, but it is also something that you discover yourself *having done* without deciding or perhaps even wanting to do. Your caring could be measured by the extent to which reward is at stake; but the amount of reward at stake also depends on how much you care. Thus you could be said to *give* the activity *hedonic importance*, but also to *find* it to be more or less important. The process of

giving/finding importance seems to be what Freud meant by “cathexis” (*besetzung* or sitting on in the original; Strachey, 1933/1956), but mostly we have lacked a word that captures its dual implication.⁸ You can grow hedonic importance by cultivating a skill or an alliance, “rooting for” one team or another, or dedicating yourself to one faith or another. The steps that lead to caring can be deliberate—when you move to a new city and decide whether you root for the new team or the old one. But the degree of caring that brings you the greatest hedonic gains, or commits you to major losses—when mighty Casey strikes out or the Atlanta Braves unexpectedly lose the pennant (Sternbergh 2011)—must grow over time.

The experience of hedonic betting is familiar, but the way the bets are enforced is not obvious. Seemingly they are enforced by recursive self-prediction as in personal rules-- to maintain the appetite until the right occasions occur, rather than redefining the occasions or withdrawing your attention to alternative bets. A mental action that spoils a bet is less conspicuous than turning a page or card. If you withdraw your investment in a movie when it gets too scary, you may or may not be conscious of “saying to yourself” that it’s only a story; but doing so reduces the potential of the movie to give you occasions for reward subsequently. “Saying to yourself” redefines the set of rewards at stake, from those that depend on the suspense that has built up to a matter of ordinary curiosity—an immediate relief at the expense of a long term reduction in importance, which may deter you from saying it. Likewise you may or may not be conscious of rooting less for a sports team, or valuing a friendship less, but you are apt to notice a change in your relevant payoffs. As with personal rules, seeing yourself pull investment out of a movie will make you less likely to resist the urge to do so later, and not just in this movie but in other movies and perhaps other projects. This investment is the stake of your bet(s), which you experience when you “give importance” to a project and “find it important” in turn. Noticing such changes is as close as we’re apt to get to discerning a deliberate action that modulates hedonic importance.

Hedonic importance and personal rules are both recursive

Self-prediction in the case of personal rules may be deliberate or not, but with the importance of a bet it is usually not. The two kinds of self-prediction may interact, in that the stake of a personal rule is apt to vary with the hedonic importance of the outcomes. But personal rules use self-prediction specifically to increase the success of long term interests against short term ones, whereas the assignment of importance does not necessarily serve either faction. A love for Rosaline may fade and that for Juliet grow; the old team may lose importance to the new; and a world view may be outgrown or renounced for a different one; none with any necessary implication for long term reward. Unlike the stake of a personal rule, hedonic importance can strengthen addictive and frankly aversive options too: An addiction to personal risk, increasingly risky shoplifting, self-destructive mountain climbing (Leamer 1999). a pet hatred, a phobic

⁸ Another definition of endogenous reward is that reward which you can make contingent on bets, without its being secondary to a species-specific or hardwired outcome. It is thus separate from the *intrinsic* reward for following internal but inborn patterns, as modeled by Singh et.al. (2010) and others—although their argument for going beyond the conventional model of secondary reward is excellent. Likewise *hedonic importance* resembles Renninger’s recursive *interest* (1992), but implies recruitment of motivational force. See my earlier discussion, without recursion (Ainslie, 1992, pp. 243-273), and a mechanical model (*ibid.*, pp. 274-300).

anxiety, or preoccupation with a jealous suspicion may be cultivated into overwhelming size despite the opposition of personal rules. For want of a motivational explanation the role of hedonic importance has often been dismissed as habit—the third case in the foregoing section on habit.

Occasions for endogenous reward must be singular

The great weakness of hedonic bets is their arbitrariness, the fact that they can as well be based on any number of scenarios leading to any number of outcomes, as in daydreams. Of course, tastes for endogenous reward cannot be purely arbitrary. They are undoubtedly constrained by kinds of evolved readiness—to hoard, to avoid contamination, to experience others' rewards vicariously—and resistance to satiety, varying among individuals on a continuum from a high of fantasy-proneness (Rhue & Lynn, 1987) to a low in sensation-seeking and reward-deficient people, who have an excessive need for stimulation. Remembered deprivation or trauma may make appetites in some modalities satiate less. But subject to these influences a wide range of possible scenarios remain, and tend to undermine each others' importance. This weakness can be overcome by betting on *singular* occasions—those that stand out from other possible occasions by being infrequent and easily distinguishable from those that are more frequent. Trivially, singularity is the reason we admire a difficult rhyme, commemorate a round anniversary, or value the rare provenance of a work of art. But singularity also governs the significance we attribute to places close to home, relatives close to us, or times close to the present. Situations that offer singular occasions optimally, and especially those that unfold new occasions as familiar ones habituate, can be said to have good *texture* (Ainslie, 2013a).

As with the bright lines that anchor personal rules, singular features are apt to anchor recurring bets. A history of having bet on an activity or faction creates additional singularity besides whatever formal singularity it has—the one contender that has been mine, the one team that I have long cared about, the unique kind of collection or quest that has a personal meaning, the religious belief that I grew up with or for which I find a unique rationale—or the paranoid delusion that explains many things at once—such foci are landmarks in the open field of endogenous reward, attractors of the repeat betting that multiplies hedonic importance. The bets still have to create appetite that is neither gratified too soon nor utterly frustrated, and which attach to new scenarios after old ones habituate; but to the extent that I have found a source of such occasions to bet on, the resulting importance can accumulate as “consumption capital” (Becker & Murphy, 1988). The recursive self-prediction of hedonic importance can sometimes result in extreme quests—striving to spot a white tiger in nature, make a hole-in-one, climb Mount Everest, die a martyr. The growth of such monumental importance can certainly not be explained as a product of secondary reward.

Some ostensibly secondary rewards may be largely endogenous

Goals may become singular also by being instrumentally important, so their benchmarks can serve as occasions for hedonic reward in addition to being secondary external rewards. It is difficult to distinguish the two processes by observation, although some studies have demonstrated their separate roles in the case of money (Wittman et.al., 2010 as discussed above; Lea & Webley 2006). Arguably the long chains of association or deduction that should cause many a benchmark to extinguish as a secondary reward still preserve its singularity, so it can remain a good occasion for endogenous reward. The role of accomplishments as occasions is

conspicuous when ostensibly productive activities that are good pacers of endogenous reward become process addictions, sometimes in disguise, as in “skilled” gambling, day trading of stocks, or dealing in collectibles. In effect, interests in endogenous rewards can parasitize instrumental rewards and create incentives to misperceive the most effective ways to get them. The same use of instrumental tasks to occasion endogenous reward is probably a factor in workers’ resistance to new technologies (see Navarro & Osiurak, 2015). Experimental exploration of the value of instrumentality *per se* has begun under the name “idleness aversion,” as in a report of subjects’ preferring prizes that required relatively more effort to those requiring less, if and only if they had a pretext to seek them (Hsee, Yang, & Wang, 2010). These authors put their fingers exactly on its seeming paradox: “People dread idleness, yet they need a reason to be busy... many purported goals that people pursue may be merely justifications to keep themselves busy” (ibid., p. 926).

Examples abound in ordinary life. People notoriously waste effort putting the exact postage on letters, squeezing the last toothpaste from the tube, and looking for tiny differences in the price of gasoline (at least in the USA). People search for a parking spot with the least walking to a gym, where they will labor on a treadmill. Rituals of recycling look like good fodder for a similar criticism. Accrued hedonic importance may elevate objectively trivial transactions into highly motivating ones: Grocery savings stamps and similar coupon schemes are notorious. Significantly, all these pastimes are distinguished by having a rationale as useful. They would lose importance without the singularity of “good reason.” Arguably the crossed purposes of increasing productivity versus cultivating appetite underlie much economic irrationality in general, but that is another topic (Ainslie, 2013b).

Summary

Hyperbolic delay discounting makes the recursive process of self-prediction an important determinant of expected reward. Two kinds of recursiveness help determine human tastes: intertemporal bargaining and hedonic importance based on endogenous reward.⁹ These patterns of reward can be discerned under five disparate choice mechanisms that have been proposed to underlie human behavior—and, hence, misbehavior:

1. (1) *Rational choice* is consistent with maximizing hyperbolically discounted reward if the goal is either committed to at a distance or chosen in a way that predicts choice of a bundle of similar rewards in the future.
2. (2) “*Conditioned*” *appetite/emotion*, including negative cases, is a reward-seeking process. That is, it not only gives rise to incentives, but is itself incentivized by opportunities for reward—including the very short term reward that must be the seductive factor in aversive urges.
3. (3) *Habit* is not a selective factor for behavior, but just a description of repetition that might be the result of several motivational factors. The greatest such factor may be the narrowed options that result from failed intertemporal bargains.

⁹ A third kind of recursiveness, positively fed back appetite or emotion, is made possible by their reward-dependence, but was mentioned only briefly above (under conditioning; see Ainslie, 2010).

4. (4) There is experimental evidence that *incentive salience* has a neurophysiological basis distinct from longer term rewards, but this is consistent with its functioning as a short term reward in the same marketplace.
5. (5) Motivational analysis should leave room for *endogenous reward*—reward that does not depend on predicting a hardwired primary reward. Hyperbolic discounting suggests how hedonic importance is cultivated by *betting* on singular occasions; but practical observation is complicated by two dualities in many kinds of accomplishment:
 - Events may be both instrumental secondary rewards and occasions for endogenous reward.
 - Recursive self-prediction is the mechanism of both the hedonic importance of bets and the enforcement of the bets themselves.

The need to trace all reward to hardwired events-- pre-programmed, or external in the sense of not originating in a person's choices—might be the biggest failure of behaviorist theory. It is, of course, a consequence of the anti-theoretical philosophy shared by the founders (e.g. Skinner, 1969, pp. 236-240). However, hyperbolic delay discounting—itsself based on behaviorist methods—suggests that endogenous reward need not be secondary to hardwired rewards when it is occasioned by singular events. This reward is a personal fiat currency, but, like all fiat currencies, dependent on factors that prevent inflation.

The econometric takeaway is that because of the recursive modulation of many motives, individuals' choices in microeconomic markets cannot be predicted by knowledge of her prior incentives (Ainslie, 2011). Still, understanding impulses as market phenomena based on a single currency may suggest better interventions than do the notions of illogic, control by stimuli, automatic responding, neural damage, and poor predictive ability that crop up in the literature of addiction and other self-defeating behaviors.

References

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.

Ainslie, G. (2009) Pleasure and aversion: Challenging the conventional dichotomy. *Inquiry* 52 (4), 357-377.

Ainslie, G. (2010) The core process in addictions and other impulses: Hyperbolic discounting versus conditioning and cognitive framing. In D. Ross, H. Kincaid, D. Spurrett, & P. Collins (Eds.), *What Is Addiction?* MIT, pp. 211-245.

Ainslie, G. (2011) Free will as recursive self-prediction: Does a deterministic mechanism reduce responsibility? In J. Poland & G. Graham (Eds.) *Addiction and Responsibility*. MIT Press, pp. 55 - 87.

Ainslie, G. (2012) Pure hyperbolic discount curves predict “eyes open” self-control. *Theory and Decision* 73, 3-34. 10.1007/s11238-011-9272-5

Ainslie, G. (2013a) Grasping the impalpable: The role of endogenous reward in choices, including process addictions. *Inquiry* 56, 446-469. DOI: 10.1080/0020174X.2013.806129. <http://www.tandfonline.com/eprint/8fGTuFsnfFunYJKJ7aA7/full>

Ainslie, G. (2013b) Money as MacGuffin: A factor in gambling and other process addictions. In Neil Levy, ed., *Addiction and Self-Control: Perspectives from Philosophy, Psychology, and Neuroscience*. Oxford University Press, pp. 16-37

Ainslie, G. (2015) The cardinal anomalies that led to behavioral economics: Cognitive or motivational? *Managerial and Decision Economics*. DOI:10.1002/mde.2715.

Ainslie, G. (2016a) Intertemporal bargaining in habit. *Neuroethics*. DOI 10.1007/s12152-0169294-3.

Ainslie, G. (2016b) Palpating the elephant: Current theories of addiction in the light of hyperbolic delay discounting. In Heather, Nick, & Segal, Gabriel (Eds.) *Addiction and Choice: Rethinking the Relationship*. Oxford U. pp-. 227-244.

Ainslie, G. & Engel, B. T. (1974) Alteration of classically conditioned heart rate by operant reinforcement in monkeys. *Journal of Comparative and Physiological Psychology* 87, 373-383.

Ainslie, G. & Monterosso, J. (2003) Building blocks of self-control: Increased tolerance for delay with bundled rewards. *Journal of the Experimental Analysis of Behavior* 79, 83-94.

Atkinson, J. W. & Raynor, J. O. (1975) *Motivation and Achievement*. Winston & Sons.

Baum, W. M. (2005) *Understanding Behaviorism* 2d Edition. Blackwell.

Becker, G. & Murphy, K. (1988) A theory of rational addiction. *Journal of Political Economy* 96, 675-700.

Berridge, K. C. (2003) Pleasures of the brain. *Brain and Cognition* 52, 106-128.

Berridge, K. C., & Kringelbach, Morten L. (2008). Affective neuroscience of pleasure: reward in humans and animals. *Psychopharmacology*, 199(3), 457-480.

Cabanac, M. (1992) Pleasure: The common currency. *Journal of Theoretical Biology* 155, 173-200.

Carter, R. M., Meyer, J. R., & Huettel, S. A. (2010) Functional neuroimaging of intertemporal choice models: A review. *Journal of Neuroscience, Psychology, and Economics* 27-48.

Chen, M. K., Lakshminarayanan, V., & Santos, L. R. (2006) How basic are behavioral biases? Evidence from Capuchin monkey trading behavior. *Journal of Political Economy* 114, 517-537.

- Conklin, C. A., & Tiffany, S. T. (2002). Applying extinction research and theory to cue-exposure addiction treatments. *Addiction*, 97(2), 155-167.
- Conlisk, J. (1996). Why bounded rationality? *Journal of Economic Literature*, 34, 669-700.
- Corsini, R. J. & Wedding, D. (2011) *Current psychotherapies, ninth edition*. Brooks/Cole.
- 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.
- Deci, E. L. & Ryan, R. M. (1985) *Intrinsic Motivation and Self-Determination in Human Behavior*. Plenum.
- Dennett, D. C. (2003) *Freedom Evolves*. Viking.
- Dolan, R. J., & Dayan, P. (2013). Goals and habits in the brain. *Neuron*, 80(2), 312-325.
- 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.
- Everitt, B. J. & Robbins, T. W. (2005) Neural systems of reinforcement for drug addiction: From actions to habits to compulsion. *Nature Neurosciences* 22, 3312-3320.
- Everitt, B. J., & Robbins, T. W. (2013). From the ventral to the dorsal striatum: devolving views of their roles in drug addiction. *Neuroscience & Biobehavioral Reviews*, 37(9), 1946-1954.
- Eysenck, H.J. (1967) Single trial conditioning, neurosis and the Napalkov phenomenon. *Behaviour Research and Therapy* 5, 63-65.
- Fellows, L. K. & Farah, M. J. (2005) Different underlying impairments in decision-making following ventromedial and dorsolateral frontal lobe damage in humans. *Cerebral Cortex* 15, 58-63.
- Gigerenzer, G., Fiedler, K., & Olsson, H. (2012). Rethinking cognitive biases as environmental consequences. In P. M. Todd, G. Gigerenzer, & the ABC Research Group (Eds.). *Ecological Rationality: Intelligence in the World* (pp. 80–110). New York: Oxford University Press.
- Glimcher, P. W. (2009) Choice: towards a standard back-pocket model. In P. W. Glimcher, C. Camerer, Russell A. P., & Ernst Fehr (Eds.) *Neuroeconomics: Decision Making . Brain*. Elsevier, pp. 503 – 521.
- Granda, A.M. & Hammack, J.T. (1961) Operant behavior during sleep. *Science* 133, 1485-1486.
- Green, L., Fry, A., & Myerson, J. (1994) Discounting of delayed rewards: A life-span comparison. *Psychological Science* 5, 33-36.

- Gregorios-Pippas, L., Tobler, P. N., & Schultz, W. (2009) Short-term temporal discounting of reward value in human ventral striatum. *Journal of Neurophysiology* 101, 1507-1523.
- Grüne-Yanoff, T. (2015). Models of Temporal Discounting 1937–2000: An Interdisciplinary Exchange between Economics and Psychology. *Science in Context*, 28, 675-713
doi:10.1017/S0269889715000307
- Heath, R. G. (1972). Pleasure and brain activity in man: Deep and surface electroencephalograms during orgasm. *Journal of Nervous and Mental Disease*, 154, 3-18..
- Herrnstein, R. J. (1997) *The Matching Law: Papers in Psychology and Economics*. Edited by H. Rachlin and D. I. Laibson. New York: Sage.
- Hirschman, A. (1977) *The Passions and the Interests*. Princeton, N.J. : Princeton University Press.
- Hofmeyr, A., Ainslie, G., Charlton, R. & Ross, D. (2010) The relationship between addiction and reward bundling: An experiment comparing smokers and non-smokers. *Addiction* 106, 402-409.
- Hsee, C. K., Yang, A. X., & W., Liangyan (2010) Idleness aversion and the need for justifiable busyness. *Psychological Science* 21, 926-930.
- Huys, Q. J., Guitart-Masip, M., Dolan, R. J., & Dayan, P. (2015). Decision-theoretic psychiatry. *Clinical Psychological Science*, 2167702614562040.
- Jevons, W.S. (1871/1911) *The Theory of Political Economy*. London, Macmillan.
- Johnson, A. & 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.
- Kable, J. W. & Glimcher, P. W. (2007) The neural correlates of subjective value during intertemporal choice. *Nature Neuroscience* 10, 1625-1633.
- Kahneman, D. & Tversky, A. (1979) Prospect theory: an analysis of decision under risk. *Econometrica* 47, 263-291.
- Kalivas, P. W., & Volkow, N. D. (2005). The neural basis of addiction: a pathology of motivation and choice. *American Journal of Psychiatry*, 162(8), 1403-1413.
- Kirby, K. N. (1997) Bidding on the future: Evidence against normative discounting of delayed rewards. *Journal of Experimental Psychology: General* 126, 54-70.
- Kirby, K. N., and Guastello, B. (2001) Making choices in anticipation of similar future choices can increase self-control. *Journal of Experimental Psychology: Applied* 7, 154-164.

Koch, A. K., & Nafziger, J. (2011). Self-regulation through Goal Setting. *Scandinavian Journal of Economics*, 113(1), 212-227.

Kuhn, M., Lütolf, M., & Reinhart, W. H. (1999). Ondine's curse. *Respiration*, 66(3), 265-265.

Lane, R. D., Ryan, L., Nadel, L., & Greenberg, L. (2015). Memory reconsolidation, emotional arousal, and the process of change in psychotherapy: New insights from brain science. *Behavioral and Brain Sciences*, 38, 1-64.

Lea, S. E.G. & Webley, P. (2006) Money as tool, money as drug: The biological psychology of a strong incentive. *Behavioral and Brain Sciences* 29, 161-209.

Leamer, L. (1999) *Ascent: The Spiritual and Physical Quest of Legendary Mountaineer Willi Unsoeld*. Minot, ND, Quill.

Levy, D. J., & Glimcher, P. W. (2012). The root of all value: a neural common currency for choice. *Current opinion in neurobiology*, 22(6), 1027-1038.

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*. Erlbaum.

Melzack, R., Weisz, A.Z. & Sprague, L.T. (1963) Stratagems for controlling pain: contributions of auditory stimulation and suggestion. *Experimental Neurology* 8,239-247.

Miller, N. (1969) Learning of visceral and glandular responses. *Science* 163, 434-445.

Miller, W. R. & C'de Baca, J. (2001) *Quantum Change: When Epiphanies and Sudden Insights Transform Ordinary Lives*. Guilford.

Montague, P. R. & Berns, G. S. (2002). Neural economics and the biological substrates of valuation. *Neuron*, 36, 265-284.

Navarro, J., & Osiurak, F. (2015). When Do We Use Automatic Tools Rather Than Doing a Task Manually? Influence of Automatic Tool Speed. *The American journal of psychology*, 128(1), 77-88.

O'Brien, C. P., Ehrman, R. N., & Ternes, J.W. (1986) Classical conditioning in human dependence. In S. R.Goldberg & I. P. Stolerman (Eds), *Behavioral Analyses of Drug Dependence*. Orlando, FL: Academic, pp. 329-356.

Olson, M. (1982) *The rise and decline of nations*. Yale University Press.

Oudeyer, P. Y., Kaplan, F., & Hafner, V. V. (2007). Intrinsic motivation systems for autonomous mental development. *IEEE Transactions on Evolutionary Computation*, , 11(2), 265-286.

Pessoa, L. (in press) Précis of *The Cognitive-Emotional Brain*. *Behavioral and Brain Sciences*.

- Plato (1892) Protagoras. *The Dialogues of Plato* (B. Jowett, trans) section 356, p. 181. Macmillan.
- Platt, M. L. & Padoa-Schioppa, C. (2009) Neuronal representations of value. In Paul W. Glimcher, Colin Camerer, Russell Alan Poldrack, and Ernst Fehr, eds., *Neuroeconomics: Decision making and the brain*. Elsevier, pp. 441 – 462.
- Posner, R. (1998) Rational Choice, Behavioral Economics, and the Law. *50 Stanford Law Review* 1551-1575),
- Renninger, K. A. (1992) Individual interest and development: Implications for theory and practice. In K. A. Renninger, S. Hidi, & A. Krapp, (Eds.), *The Role of Interest in Learning and Development*. Erlbaum.
- Rescorla, R. A. (1988) Pavlovian conditioning: It's not what you think it is. *American Psychologist* 43, 151-160.
- Rhue, J. W. & Lynn, S. J. (1987) Fantasy proneness: The ability to hallucinate "as real as real" *British Journal of Experimental and Clinical Hypnosis* 4, 173-180. NB: Is this orig? one journal has it in v. 6, 1988.
- Ross, D. (2014) *Philosophy of Economics*. Palgrave.
- Ryan, R. M. & Deci, E. L. (2000c) Intrinsic and extrinsic motivations: Classic definitions and new directions. *Contemporary Educational Psychology* 25, 54-67.
- Samuelson, P.A. (1937) A note on measurement of utility. *Review of Economic Studies* 4, 155-161.
- Schwartz, M. S. & Andrasik, F. (Eds.) (2016) *Biofeedback: A Practitioner's Guide*. Guilford.
- Shizgal, P., & Conover, K. (1996) On the neural computation of utility. *Current Directions in Psychological Science* 5, 37-43.
- Singh, S., Lewis, R. L., & Barto, A. G. (2009). Where do rewards come from. In *Proceedings of the Annual Conference of the Cognitive Science Society*, pp. 2601-2606.
- Skinner, B.F. (1969) *Contingencies of Reinforcement: A Theoretical Analysis*. Appleton-Century-Crofts.
- Slovic, P., & Tversky, A. (1974) Who accepts Savage's axiom? *Behavioral Science* 19, 368-373.
- Spurrett, D. (2014) Philosophers should be interested in 'common currency' claims in the cognitive and behavioural sciences. *South African Journal of Philosophy* 33, 211-221.

Sternbergh, A. (2011) The thrill of defeat for sports fans. *New York Times Magazine* October 23, 2011, pp. 18-20.

Strachey, J. (1933/1956) The emergence of Freud's fundamental hypotheses. In J. Strachey & A. Freud (Eds.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*. vol. 3. 63-68. Hogarth.

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.

Tomkins, S. S. (1978) Script theory: Differential magnification of affects. *Nebraska Symposium on Motivation* 26, 201-236.

Voon, V., Derbyshire, K., Rück, C., Irvine, M. A., Worbe, Y., Enander, J.,... & Robbins, T. W. (2015). Disorders of compulsivity: a common bias towards learning habits. *Molecular psychiatry*, 20(3), 345-352.

Wilson, J. Q. & Herrnstein, R. J. (1985) *Crime and Human Nature*, New York: Simon & Schuster.

Wittmann, M., Lovero, K. L., Lane, S. D., & Paulus, M. P. (2010) Now or later? Striatum and insula activation to immediate versus delayed rewards. *Journal of Neuroscience, Psychology and Economics* 3, 15-26. Doi: 10.1037/a0017252.

Zhang, Y., Tian, J., von Deneen, K. M., Liu, Y., & Gold, M. S. (2012). Process addictions in 2012: food, internet and gambling. *Neuropsychiatry*, 2(2), 155-161.