

Emotion: The Gaping Hole in Economic Theory

George Ainslie

Veterans Affairs Medical Center, Coatesville, PA 19320

www.picoeconomics.com

Presented at the 31st Annual conference of the Eastern Economic Association, New York City, March 6, 2005; published in In B. Montero and M. D. White, Eds., *Economics and the Mind*. Routledge. 2006.

Economics is the study of choice among limited resources. This kind of choice is easy to conceive and, often, to study when the limiting factor is the physical availability of external goods. Choice among internal processes that are available at will is harder to conceive and almost impossible to study, which is undoubtedly why economists have thought very little about them. And yet the most prominent of these internal processes, the emotions, are also limited in regular ways, and could be said to be far bigger creators of value in modern societies than goods which are physically limited. I will suggest a conception of how emotions generate value that is based on what are now widely accepted research findings, and point out some likely implications for more conventional economics.

The best sources of economic data have been marketplaces, which have revealed a basic constraint on competitive choice-making: Successful marketplace behavior depends on assigning each good a value that declines exponentially with delay. People who do not buy and sell according to exponential discounting become what are called money pumps, since a more rational agent can buy from them when they undervalue a good and then simply wait until they overvalue the good to sell it back to them at a profit. Thus economic theories of choice have converged on the model that assigns this value according to exponential discount curves, rational choice theory (RCT). Formalized almost as an afterthought by Samuelson (1937), RCT has become the established theory of how people value goods in the absence of pathology. It has even been argued that some forms of pathology, that is, of self-destructive choice-making, can arise within the framework of RCT via an extremely steep discount function (Becker & Murphy, 1988). Insofar as other fields assume that choice will stay constant over time in the absence of new information they can be said to follow RCT as well, since concave discount curves based on any other function predict change of preference between smaller, sooner (SS) and larger, later (LL) goods as a function of elapsing time.

However, an increasing number of cases have been brought to light where people regularly depart from RCT. I have argued that RCT describes a special case of human motivation that arises as people adapt to competitive marketplaces. Elementary motivation in humans and nonhumans alike is described by a non-exponential discount curve, one that creates conflict among successive evaluations. The resulting patterns of

choice must be studied by analysis of this conflict, a micro-microeconomics or *picoeconomics* (Ainslie, 1986, 1992).

A recent article catalogued cases anomalous for RCT under the rubrics of bounded (limited) rationality, bounded willpower, and bounded self-interest (Jolls et.al., 1998). I will use these three categories to describe the potential contributions of picoeconomics to macro/microeconomic theory. In particular, I will expand Jolls et.al.'s "bounded self-interest" category to begin an examination of how occasions for emotion acquire value alongside other limited resources, a topic which by the extent of motivation involved represents a gaping hole in RCT and economic theory generally.

Of the three, bounded rationality has been the subject of the most research, but represents the least problem for conventional theory. It describes what look mostly like simple perceptual or computational errors in estimating value—framing effects, endowment effects, and many of the inconsistencies of choice described by Kahneman, Tversky, and their collaborators (1982). Some of these may actually represent evasions of self-control, such as the defense of sunk costs in order to postpone the realization of loss, or attempts to achieve self-control, such as choosing illiquid investments at the cost of poorer returns (Harris & Laibson, 2001); but many seem to be innocent errors, that is, they do not seem to be motivated. Rational allowance for limited cognitive capacity has been well described in the literature on satisficing (starting with March & Simon, 1958).

Bounded willpower

According to RCT bounded willpower should not even be observed, since there is the conventional approach predicts no need for will in the first place. A rational agent simply maximizes her exponentially discounted prospects at all times, and would have no incentive to restrict her future range of choice. It was the failure of choice to retain its predicted consistency over time in the absence of new information that first made the need for a radical re-evaluation of RCT evident. I have ascribed this failing to the shape of the basic discount curve (Ainslie, 1975, 2001). A large number of experiments have now confirmed that both humans and nonhuman animals tend to discount prospective events in hyperbolic rather than exponential curves (Green & Myerson, 2004; Kirby, 1997). Humans seem to achieve exponential discounting only in special situations, particularly where competition for quantifiable goods threatens to make them money pumps to the extent that they fail. The highly bowed shape of hyperbolic curves predicts that SS goods will often be preferred temporarily over LL goods, in the period just before the SS goods become available (*impulsiveness*; figure 1). The relatively high tails of hyperbolic curves predict also that making sequential choices in *bundles* should increase the incentive to pick LL options (figure 2). That is, series of LL options will be more apt to be chosen over the series of their SS alternatives than is the single LL option of figure 1 to be chosen over the single SS option. Furthermore, these high tails should motivate a person (or nonhuman, if given a simple method) to commit her future behavior so as to forestall temporary preferences for SS options. These effects have all been observed experimentally: Both human and nonhuman subjects switch preferences from LL to SS alternatives as the SS alternatives draw near, increase their tendency to pick LL

alternatives when choosing whole series of rewards at once, and learn behaviors the only effect of which is to commit them to make LL choices (reviewed in Ainslie, 2001 and 2005).

Figure 1

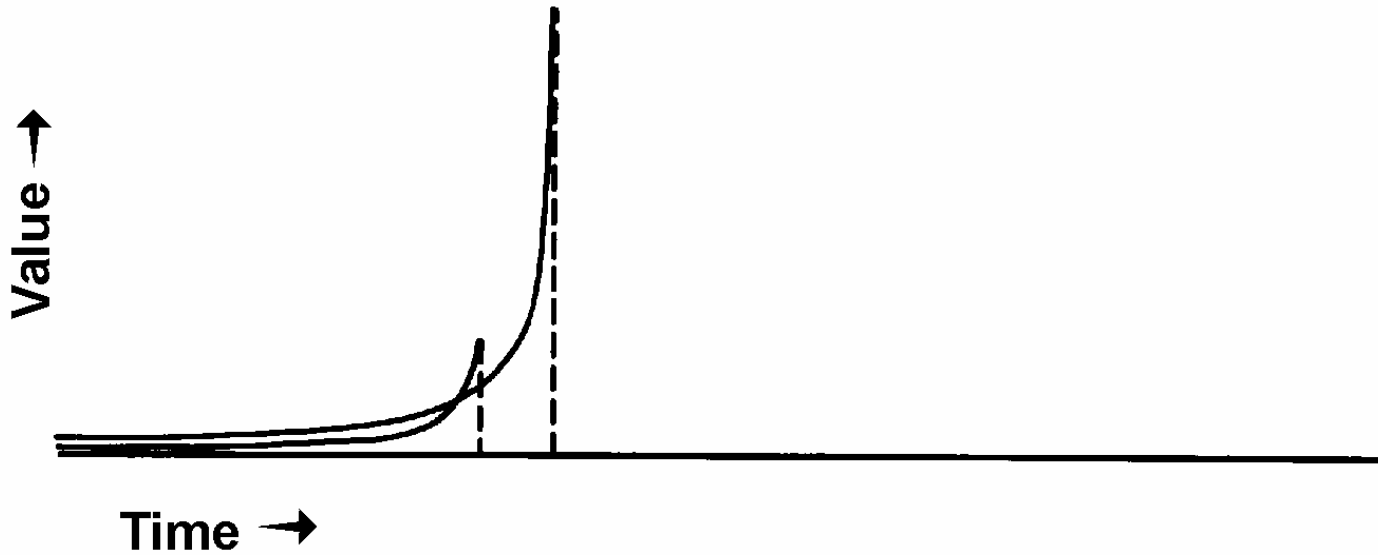


Figure 1. Hyperbolic discount curves from two rewards of different sizes available at different times. The vertical bars represent the value of the reward when immediate, and each curve represents the discounted value of that alternative as a function of the time before it will be available. The smaller, sooner (SS) reward is temporarily preferred for a period just before it is available, as shown by the portion of its curve that projects above the curve from the larger, later (LL) reward.

Figure 2

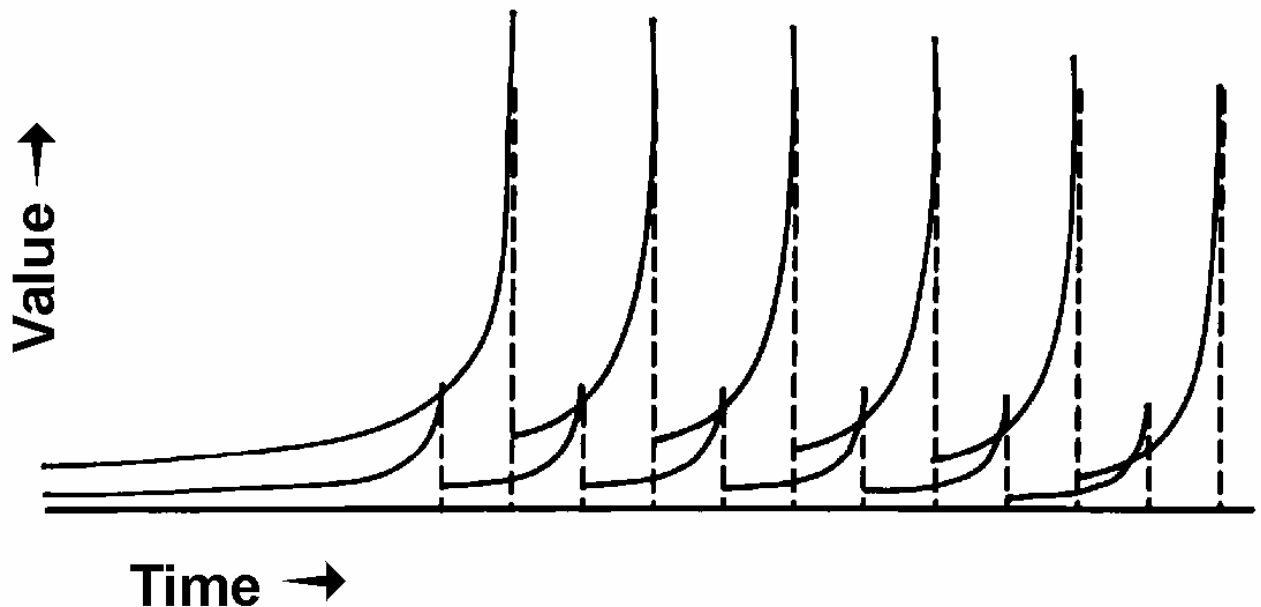


Figure 2. Summed *hyperbolic* curves from a series of LL rewards and an alternative series of SS rewards. The curve from each reward represents the discounted value of that reward when summed with all the other rewards of the same size occurring later in time (to the right). The period of temporary preference for the series of all six SS rewards is about zero, but as time passes and there are fewer choices in the series, periods of temporary preference for the series of SS rewards get longer. The curves from the final pair of rewards show the same period of temporary preference that is depicted in figure 1.

The greatest objections to hyperbolic discounting have been (1) that it lacks an adaptive purpose and therefore should have been selected against in evolution, and (2) that as often as not people avoid the impulsiveness it predicts. (1) It is not possible to state positively why hyperbolic discounting should have evolved as the basic principle of evaluating future events. The idea that it motivates organisms to attend to immediate opportunities and threats is not adequate, since exponential curves would provide a more objective weighing of immediate versus delayed events. Hyperbolic discounting might have been adaptive by making individuals follow instincts that lead them to sacrifice their own long range interests for the sake of their offspring—to lure women into having babies and men into defending kin—but the hyperbolic shape was in place at least as early as rats and pigeons evolved, before there was much likelihood that an animal would recognize its long range interest in self-preservation. However, traits often survive not because they increase fitness but because they are not-too-costly side effects of traits that do. Hyperbolic discounting of prospective rewards is likely to have been a consequence of the more general hyperbolic principle of estimating sensory magnitudes, the Weber-Fechner law (Gibbon, 1977), a consequence that was harmless until significant foresight

evolved. And evolution requires time to fine-tune fitness. Organisms sufficiently foresighted to get in trouble from hyperbolic discounting have existed for no more than a hundred thousand years or so. Whatever the case may be, hyperbolic discounting is an empirical given, and any analysis of choice motivation needs to include the motivational consequences of this discounting.

(2) When people avoid choosing SS over LL rewards, it is usually by self-control rather than spontaneous preference. The hyperbolic shape of spontaneous discount curves themselves suggests several means of self-control, such as finding external precommitments, diverting attention from likely sources of SS reward, and cultivating incompatible emotions that have some momentum: but the workhorse of self-control is willpower. The decrease in impulsiveness that results when choices are made in bundles suggests a mechanism for willpower, including a mechanism by which we can sometimes get our choices to approximate RCT. It also provides rationales for both the familiar experience of freedom of will and maladaptive overcontrol (compulsiveness). Again, these phenomena are described in Ainslie, 2001 and 2005, and are only summarized here. Basically, an adequately self-observant organism—probably only a human—will come to see frequent temptation by SS rewards as creating an intertemporal variant of the bargaining game, repeated prisoner's dilemma. Of course, she will usually recognize this property under some other name or have no name for it, just as *interpersonal* repeated prisoner's dilemmas that arise naturally have usually been recognized only tacitly, even by effective players.¹ Once you have become aware of an *intertemporal* prisoner's dilemma relationship, even indirectly, giving in to a current temptation will make you expect to give in to similar temptations in the future, and thus reduce your expectation of getting a whole series of LL rewards rather than just the one currently at stake.

When you think your resistance to the current temptation is both necessary and sufficient to maintain the cooperation of your successive selves, you will have an incentive that constitutes willpower, without the involvement of any other faculty or organ. Since your estimate of necessity and sufficiency depends on your interpretation of past choices and on your prediction of future selves' interpretation of your current choice, your decision cannot be manipulated or even predicted with certainty from the incentives you face. Thus your will is imponderable, and, arguably, free.² However, the more your current choice matters as a precedent rather than as an experience in its own right, the more your evaluation of your choices will become lawyerly and, in the extreme case, compulsive. The observational benchmarks that predict whether you will go on cooperating with future selves come to define internal laws or *personal rules*, which are self-enforcing because your expectation of whole bundles of LL reward depend on your seeing yourself obey them.

Since this model of will depends on the recursive feeding back of choices and even of mooted choices, it will not be directly testable by controlled experiments until it is possible to observe the internal stages of choice. Neuroimaging has made some progress in locating the brain sites that are active when there is a prospect of reward and even when a subject is resisting an urge to consume a reward, but the results so far have not gone much beyond brain geography: When human subjects evaluate expected cash

rewards, they not only discount them hyperbolically but generate correlated gradations of neural activity in identified brain centers such as the ventral striatum and posterior parietal cortex (Glimcher & Kable, 2005). These regions are active also when smokers anticipate smoking (Monterosso, Ainslie & London, 2006). As both Berridge (1999) and I (1992, pp. 101-114) have pointed out, a rewarding component is necessary for even aversive processes to be “attention grabbing” (Berridge, 2004, p. 195), a component that he calls “incentive salience” and that I call very short range reward. There is evidence that the reward component that is common to both pleasurable and aversive processes involves GABAergic neurons in one part of the ventral striatum, the nucleus accumbens (*ibid*). Similar is Bechara’s concept of “primary induction,” for which he finds the amygdala necessary (2006). There are even some imaging studies of self-control. For instance, subjects choosing between SS and LL coupons with monetary exchange values have more lateral prefrontal activity when they choose the LL coupons (McClure et al., 2004). When smokers are instructed to not take available puffs on a smoke-delivery apparatus, the dorsal anterior cingulate cortex and supplementary motor area become differentially active (London et al., 2006). The medial prefrontal (orbitofrontal) cortex has been implicated in both temptation and its opposite, rational planning, depending on the method of observation (Davidson et al., 2000; McClure et al., 2004; Rolls, 1999, pp. 124-144; Volkow & Fowler, 2000).

These neurophysiological data are encouraging. They confirm that some of the functional components of motivation that have been derived from behavioral research have specific locations. They increasingly support the notion that there is a marketplace in the brain of all choosable processes, i.e. that there is a common currency of reward, which makes all options comparable (Shizgal & Conover, 1996). However, the direct observation of how the components of temptation and self-control interact still looks distant.

Meanwhile, I have argued that a favorite tool of the philosophy of mind, the thought experiment, can make common intuition usable to test models of willpower (Ainslie, 2007). Kavka’s problem will serve as an example: You are offered a large sum of money just to intend to drink an overwhelmingly disgusting but harmless toxin. Once you have sincerely intended it, as verified by a hypothetical brain scan, you are free to collect the money and not actually drink the toxin (Kavka, 1983). Philosophical discussion has revolved around whether you have any rational motive actually to drink the toxin once you have the money, and whether, foreseeing a lack of such motive, you can sincerely intend to drink it in the first place, even though you would drink it if that were still necessary to get the money. People usually say that it would not be rational to drink the toxin, but are then bothered by the counterintuitive possibility of defrauding themselves in this way.

The service of this thought experiment is to point out that there is a conceptual piece missing in the common theory of how people intend difficult behaviors. It is probably not possible to intend to drink if you expect to renege, but it is possible to commit yourself, more or less, not to renege. You do this by putting up a pledge of sufficient value; and the only pledge available to put up irrevocably in this situation is the credibility of your pledges in difficult choices in the future. This kind of pledge is recursive: The more you

believe that you will keep it the more you can keep it and the more you will subsequently believe you will keep it; the less you believe you will keep it the less you can keep it, etc. The pledge need not be deliberate, or even conscious. It is enough to notice that this choice resembles a kind of choice that comes up at other times. In that case your present choice will affect your prospect of making the similar future choices, whether you want it to or not. By the same token, your memory of similar past choices will tell you whether or not you have any such prospect to lose.

The current pledge need not put the set of all future pledges at risk, but if you don't relate it to a broader class of pledges—not just those involving toxins-- you will probably expect it to be inadequate from the start. You won't feel as if anything else in your future is at stake. You will then have to throw in more collateral, as it were-- to put a larger category of pledges at stake-- if your intention is going to register on the scanner. This is a step that people facing stubborn urges speak of in many ways: You “get serious,” you “solemnly vow,” you “really mean it,” you are “in earnest this time,” etc. But whichever way you speak—or think—about it you have thereby related your current choice to the major category of choices where you are serious, earnest, etc., and thus have put a big chunk of your credibility at risk. Once you have perceived your credibility with yourself to be at risk you have a rational incentive to drink the toxin. I submit that this unique ability to solve Kavka's problem supports the intertemporal bargaining model of willpower. Furthermore, the fact that people who think about the problem usually experience the incentive to drink the toxin only as a vague discomfort with the RCT solution (renegeing) provides an illustration of the tacit nature of most intertemporal bargaining.

Higher processes. To recap my model thus far: Hyperbolic discounting predicts a population of processes in partial conflict with each other, with outcomes determined by which of them is dominant when an irreversible choice can be made. These processes have grown to obtain rewards, and remain viable as long as they sometimes succeed. Processes that obtain quick rewards will be robust, and will dominate as long as they are not forestalled by earlier processes. These earlier, necessarily foresighted processes will be based on more attenuated reward values, but values that are more nearly proportionate to the “objective” rewards at stake. They have the advantage of acting before their shortsighted competitors, and, as I have described above, can often dominate even immediate alternatives when they seem to lead to whole bundles of rewards-- the phenomenon of willpower.

There is nothing in this model to prevent the growth of foresighted processes that are generalists, brokers of reward that do not depend on specific sources of reward but forage for adequately long-lasting reward of any kind. Piaget described similar processes in his “tertiary circular reactions (1937/1954);” and, of course, if ego functions must compete in the internal marketplace they do so on this basis. These generalist processes are the financiers of the reward economy. The foresighted processes discern the greatest of rewards discounted over a relatively long time (the LL rewards) and plan how to forestall competing, faster rewards that are recognized as smaller from the perspective of distance (the SS rewards). The negatives in this model are always SS rewards, rather than

nonrewards or inverses of reward, since in the absence of reward a process cannot compete for a person's participation. Addictive rewards such as drugs or self-destructive thrills turn negative over periods from hours to days; itch-like urges such as tics or irritating habits ("wanted" but not "liked" in Berridge's theory, 1999) turn negative within seconds; urges for processes that never feel voluntary, like panic or the emotional component of physical pain, repeatedly lure and turn negative within a fraction of a second, as I have argued (Ainslie, 2001, pp. 51-61); but all compete for expression with the same positive currency. This should be called reward because its defining property is to make processes that it follows more likely to occur; but in modalities where it seduces only briefly it is radically different from pleasure.

Relevance to orthodox economics. Intertemporal bargaining connects the special case of rationality within competitive *interpersonal* markets to the general case of hyperbolic valuation—in somewhat the way that relativistic physics defines a place for Newtonian physics while correcting its anomalies. The *intertemporal* bargaining model of decision-making that I have called *picoeconomics* suggests an infrastructure for the economics of *interpersonal* transactions. The model of impulsiveness and will just described depicts the individual agents of microeconomics as populations in their own right, which, like collectives of individuals, must take their own diversities of interest into account. They will try to limit their choices to those that will stand up both to intermittent temptation and to compulsive application of personal rules. Someone trying to sell to such populations will discover markets both for rapidly paying goods that are harmful to them in the long run and for devices that restrict this market. These complex agents can sometimes follow personal rules to choose *as if* their discount curves were exponential, but often cannot do so. The lower the exponent (interest rate) demanded by these rules, the greater the strain on their resolve. *Interpersonal* bargaining among members of a culture will result in a rough consensus about what degree of this strain a person can normally be expected to withstand (Ainslie, 2001, pp. 100-104).

More importantly, people depend on interpersonal influence to help restrain their impulses; as technology develops fast-paying new goods such as television, video games, designer drugs, credit schemes, and simply an increased velocity of delivering whatever is chosen, culture lags behind in recognizing and teaching dangers. Just as the Amerindian culture of previous centuries had no time to respond to the bewildering lure of distilled liquor, we are baffled by the effects of our newly evocative inducements to emotion. Where our culture identifies problems it tends to blame the physical parts of the process—specific program content or the specific properties of drugs—but in an affluent society reward does not depend greatly on specific physical stimuli.

Behavioral science has not had much to say about intangible reward. As important as intertemporal conflict is in determining people's choices among conventional goods, it is even more important in the choice of rewards that do not physically depend on external events. *Picoeconomics* suggests an approach to the analysis of transactions whose value is based on emotion.

Bounded self-interest, the entryway to emotional reward

Self-interest is said to be bounded, i.e. limited, because of cases where individuals have behaved altruistically without any prospect of reward, or, as the argument about this case has progressed, without any increase in the prospective fitness of their genes. Even though altruistic choices can be required by personal ethics—in effect by personal rules (Rachlin, 2002)—a strict utility theorist will still demand what source of reward such rules are protecting. The obvious answer is that the immediate basis of altruism is emotional; our motive is to enjoy the beneficiary's feelings vicariously, or at least not to suffer vicarious pain. Where someone's rules are lax it even happens that people give alluring but destructive presents so as to enjoy the immediate evidence of pleasure they seem to produce, despite credible information that later the beggar will indulge his drug habit or that the fed animal will get sick. However, there is no generally accepted model of this emotional process. If altruism is mostly a way of getting occasions for vicarious pleasure, we need to ask what is the nature of vicarious pleasure.

But this question opens an even larger topic in turn. Vicarious pleasure is just one of many positive emotional experiences³—experiences that are not governed by bodily sensations. And emotions are the preponderant goods of any modern society. As physical discomforts (hunger, cold, pain) have become trivial problems, most human effort has been directed toward obtaining some kinds of emotional experience and avoiding others. Even activities that aim at getting bodily sensations have a strong admixture of getting components that are not bound to these sensations. Food is enjoyed within the context of particular tastes that differ from one culture or one person to another. Pain is more or less aversive depending on its context, and the person's attitude toward it. (Beecher describes some extreme forms of this—1959)..) It is true that people work for their money in order to buy facts—a good that is delivered or not, a service that is performed or not—but their valuation of what they buy often depends more on the emotion that it engenders than on any objective measure of well-being. Even money itself is apt to be valued for more than the goods and services that it will buy, a well-documented value that has been likened to a drug effect and that is not accounted for by standard economic theory (Lea & Webley, 2006).

Thus the difficulty with RCT extends far beyond its excessive endorsement of selfishness. I submit that much of this difficulty comes from its anchoring basic value in the hard currency of external events—no longer a single font of value like gold or land, but nevertheless a set of external stimuli that are held even by psychologists to control the reward process. On the contrary, hyperbolic discounting raises the possibility that much reward is freely available to the internal marketplace, and constrained mainly by the properties of intertemporal bargaining. Such a model makes no sense in a world of exponential discounting, of course, since that lacks a rationale for intertemporal conflict. Given hyperbolic discounting, the elusive behavior of emotion can be particularly well explained.

Emotion as reward-dependent. It is easy to study behavior toward durable rewards, events that reward reliably as long as the physiological potential (*drive*) exists, like food, sex, and relief of discomfort. When the internal process of reward is less tied to physical events, the valuation process becomes harder not only to observe but to conceive. What selective mechanism determines the value of processes that begin and end in the mind? Most value that does not arise from sensation arises from emotion, but emotions can be summoned deliberately. It is true that summoned emotion is usually paler than the kind that surprises us, but that property itself needs explanation. A process that is both a behavior subject to reward and a reward itself is in danger of falling into a positive feedback loop. What determines the competition of emotions in the internal marketplace, against each other and against more tangible sources of reward? This is a question that has begun to be answered at the neurophysiological level, but despite the birth of “neuroeconomics” the interaction of brain regions has so far told us little about how the currency of reward behaves.

In the meantime the hyperbolic form of future discounting offers at least a possible explanation (Ainslie, 2001, pp. 65-69, 161-197; 2005), which, again, is only summarized here. Reward that can be had just by opting for it will be limited by your hunger for it. This phenomenon, too, is easiest to see in the case of tangible rewards. A person with continuous, easy access to food will get significant enjoyment from eating only if she restricts herself from “grazing” and lets hunger build up. Hyperbolic discounting makes her value SS pleasures over LL ones when the SS ones are close, so her tendency will be to cash in small amounts of hunger as it develops rather than wait until the hunger is intense. In response people create strong restrictions by keeping food at a distance or making personal rules such as not to eat between meals. The motivation for such rules need have nothing to do with dieting; it may be enough that they permit sufficient build-up of hunger to make meals pleasurable.

The case of emotional reward is less obvious, but desirable emotions must be infrequent to be strongly rewarding. However, unlike food, emotions are hard to restrict. You have continuous, easy access to them, and your attention moves too quickly to be controlled by personal rules, which themselves have to wait for the attention it takes to evaluate each proposed choice by the rules’ criteria. A primitive kind of restriction develops naturally, in that a pattern of continuously opting for a given emotion—opting without occasion--will extinguish. The real competition will be among emotions cued by intermittent occasions—as well as between these emotions and tangible rewards, of course. Strings of such occasions are supplied by *texts*, a necessary term despite its abuse by deconstructionists, which covers fictions, gambling games, news reports, memories-- any sequential experience. All of these can pace the generation of emotions. However, some pacing patterns are the equivalent of grazing to satisfy hunger-- “light fiction” or casual daydreams-- while others have the power to build a high degree of suspense or longing. Having an emotion will be a more rewarding activity in the long run when the occasion is uncommon. News items, feats in sporting events, objects of collection, and victories in romance incite feelings in proportion to their perceived rarity.

Of course the rarity of an occasion may be subject to change precisely because it has served as an occasion for emotional reward—People learn to go looking for such instances. Occasions that you can voluntarily sap the strength of emotions. People-- and presumably nonhuman animals-- wind up experiencing as emotion only those patterns that have escaped the habituation of voluntary access, by a selective process analogous to that described by Robert Frank for the social recognition of "authentic" emotions (1988): Expressions that are known to be intentionally controllable are disregarded, as with the false smile of the hypocrite. By this process of selection, emotion is left with its familiar guise as passion, something that has to come over you.

The role of facts. Belief has been increasingly recognized as a behavior in modern times, but the constraints that separate goal-directed beliefs such as self-delusions from what is experienced as make-believe have not been clear. Obviously the occurrence or non-occurrence of tangible rewards—those that satisfy hungers—will strictly shape beliefs about what will obtain them, as will the success or failure of instrumental beliefs—those that deal with getting testable goals. These shaping factors are experienced as facts—call the *instrumental* facts—and they leave little room for psychological construction. There are other facts that lack such practical tests of their validity but that are still useful in a way that make-believe is not (Ainslie, 2001, pp. 175-179). The property of factuality gets its non-instrumental importance from its selection of few occasions for emotional reward from among many candidates. That is, non-instrumental facts get their importance by maintaining the rarity of occasions.

Other things being equal, texts that qualify as facts (by any stringent selective process, including communal folklore) are more potent than fictions, but facts can become cheap as well. News programs comb the world for facts which, if they happened close to you, would be overwhelmingly moving. There are all-sports channels, all-shopping-offer channels, all-history channels, and so on. The impact of facts is reduced to that of fictions as they become infotainment, that is, when they stop being relatively rare. Ironically, one of the most potent factors that limit this cheapening of non-instrumental facts is instrumental value. The set of facts that are tools for tangible gain are often excellent disciplines for pacing emotional reward, partly because they inspire your neighbors to challenge you by competing for them, but more basically because they are limited in availability. Thus instrumentality, the value of facts for getting other goals, confusingly becomes a source of non-instrumental value. For instance, gambling for money has more kick than gambling for points, even when we gamble for money purely as recreation. And in the United States, at least, the variability of gasoline prices among stations makes the search for cheaper gas a challenging game; several acquaintances have admitted to a temptation to drive uneconomically far out of their way just for the sensation of winning at this game, even though they would not be playing it if it did not ostensibly save them money. Once we authenticate money as a prize, it becomes a tool for occasioning emotion as well, as does any text selected from the general ruck of texts by an adequately stringent process.

But rarity alone is not enough. Attention ranging freely as it does, it will inevitably move forward in a text in order to anticipate its occasions, and may thereby return the

emotional reward sequence to a pattern of grazing. You can make a personal rule not to read ahead in a book, but memory or imagination will necessarily grow stale as repetition leads to anticipation and hence premature satiety; it must be refreshed by surprises--turnings that you cannot anticipate. With most emotional rewards, the only way to stop your mind from rushing ahead is to avoid approaches that can be too well learned. Thus the most valuable occasions will be those that are either 1. uncertain to occur or 2. mysterious-- too complex or subtle to be fully anticipated, arguably the goal of art. To get the most out of emotional reward, you have to either gamble on uncertainty or find routes that, although certain, will not become too efficient. In short, your occasions have to stay surprising-- a property that has also been reported as necessary for activity in brain reward centers (e.g. Hollerman *et.al.*, 1998; Berns *et.al.*, 2001). Accordingly, surprise is sometimes said to be the basis of aesthetic value (Berlyne, 1974; Scitovsky, 1976). In modalities where you can mentally reward yourself, surprise is the only commodity that can be scarce.

Vicarious reward. The source of the most robust occasions for emotion is not (mainly) a source of instrumental reward, but a source of patterns that are readily synchronized to our own emotional rhythms: the apparent experience of other people. I have argued elsewhere that the richest source of emotional occasions is to gamble on vicarious experience (Ainslie, 1995, 2001, pp. 179-186). Although a person is free to sample many sources of this experience, thus risking arbitrariness, the emotions suggested to her by a given perception are fixed—either the same emotions as her object is experiencing, or, in the case of negative empathy, an obvious converse emotion like gloating at the object's chagrin. Recent neurophysiological data suggest that just watching another person generates highly specific signals about what she is experiencing via the stimulation of "mirror neurons" in your own cortex (Iacoboni *et.al.*, 1999). Empathy seems to be the hedonic exploitation of such a process, the modeling of another person's emotional choices by using your own (see a detailed hypothesis about this by Barnes & Thagard, 1997). You adopt the criteria that you think the other is using to occasion emotion; for the time being, you entertain what you think would be her emotions. But of course, they are hers only in the sense that you are having them according to a theory about her. They are happening in your brain. If you keep your model close to your observations, you can use it to occasion emotions just as you use your own situation.

Since emotions don't require a turnkey, just available appetite and adequately rare occasions to preserve this availability, you can sometimes experience the emotions you're modeling in the other person as substantially as the ones you have as yourself. To model the other person is to have their expected feelings; and nothing makes these "vicarious" feelings differ in kind from "real" ones. This, I argue, is the basis of the altruism that does not otherwise enhance your well-being. However, the impact of this phenomenon will be limited by the uniqueness of your relationship with the other person, just as the impact of texts in general is limited by their factuality. Your vicarious experiences from strangers picked for the purpose will be little more than daydreams.

De gustibus disputare: An addition to orthodox economics I began this chapter by repeating the truism that economics is the study of choice among limited resources. I then

summarized some important implications of the hyperbolic discounting of expected value, leading to the point that intangible (emotional) reward is physically limited not by the availability of commodities but by internal states equivalent to hungers—call them the appetites or drives for these emotions—which permit self-reward at will as long as they are present. Because memory cannot be well controlled and positive emotions attenuate readily, good occasions for them must be adequately rare and surprising. Such occasions are the basic goods of the emotion-based sector of any economy.

Emotional occasions differ from other goods of commerce in several important properties:

1. They cannot be in assured supply without eventually losing their value. To maintain their freshness they must be at least partially unpredictable. To seek them deliberately you must accept gambles: works of fiction that have not become too familiar, chances for an exciting relationship, challenging tasks or sports, objects of collection that are competed for.
2. Instrumental tasks often make excellent pacers of emotional reward, but the best strategies for instrumental effectiveness are apt to differ from the best strategies for emotional reward. The most efficient way of making a product may not be one that permits craftsmanship, and the most efficient way to solve a problem may not permit the savoring of the theoretical possibilities encountered. Efficiency experts delight in uncovering the flaws of methods that have been shaped by workers' tastes—and this criticism may sometimes be necessary not just for material efficiency but to maintain the specificity of the pacing criteria—but to be realistic economics needs to recognize the rationale of the emotional strategies, too. Actually this has somewhat happened, This has begun already in analyses of what lottery structures are most popular or how game-like presentations enhance sales of tangible goods.
3. Because modern society values progress over mere cycles of appetite and satisfaction, people often feel obliged to control their impulses to seek risks. Adequately risky activities may then be selected only if they seem objectively productive. The result is a market for activities that have some rationale as productive but actually make their ostensible objects less certain: lotteries as ways to get rich, fights to impose peace, or complex methodologies that people are unaccountably loath to simplify. There often results an asymmetrical competition between making progress and putting that progress at risk, which cannot be resolved by weighing these purposes against each other because intermittent loss of the ostensible goal is necessary for maintaining the emotional payoff of this goal. The attempt to get satisfaction is called hedonically rational; the attempt to refresh appetite is stigmatized as irrational. The problem may be that society does not recognize that the value of such incentives as wealth or “objective” future prospects is subject to appetite. However, recognition that you are intentionally incurring losses is apt to undermine your perception of the task as instrumental and thus as unique. Our cultural ignorance may have itself been shaped by differential reward.
4. People develop tastes for emotional rewards just as they do for satisfying hungers. We are apt to find incitements to our favored emotions with the regularity of

- eating meals. Just as dictators are said to need an unbroken succession of enemy threats, there are people who seem always to have something to be angry about⁴ or, more rarely found, people who always find something at which to rejoice. Emotional goods depend more on individual tastes than instrumental goods do, but for an individual they may have a steadier value.
5. Emotional goods depend on the probability structure of their occasions rather than on specific, turn-key powers as foods or drugs do. Desirable structure could be called *texture*, the availability of satisfactory patterns with which to occasion emotional reward. Hence on one hand emotional goods are much more substitutable for one another than are tangible goods as one good becomes too predictable (loses its texture) or common or uncommon. Without a component that restrains premature satiation over time, they are highly susceptible to fashion.
 6. When they have such a unique component, on the other hand, they may become unsubstitutable. Since a history of having been chosen is one feature that can make an emotional occasion unique, an occasion for emotional reward may acquire the status of a fact just by having been consistently chosen. Among many potential religious tenets with adequate textures, for instance, belief in one soon becomes self-confirming, since it both serves its purpose and has come to stand out from the others by being part of the person's history. Similarly, if someone gets emotional gratification from particular habits of hospitality to neighbors or frugal housekeeping or just a daily routine, these habits may initially be shaped by an instrumental purpose; but after some years they no longer need this purpose, because the details of these specific habits have come to stand out as occasions.

Analysis of how emotional occasions are governed may certainly lead to the discovery of marketable goods, beyond what merchandizers have already discovered empirically. However, welfare economics may have the greater need for this analysis. Recognition of the self-generated nature of reward may lead to analysis of how the increasing systemization of modern institutions—trends toward uniform “best practices,” thorough review, and zero risk-- may be having a negative impact on the texture of peoples' lives. The monotony of factories has been notorious for a century, and many of the poor have always preferred life in the streets to the regimentation of institutions, but now automation has made it possible for centralized managements to impose data collection and customer relations routines on increasing numbers of employees, including doctoral level professionals and ex-entrepreneurs, who, driven out of business by superior systemization, must seek jobs with their former competitors. Recognition of the role of texture may gradually transform welfare economics from a concern with inequalities of wealth to a concern that a person or system that can control the texture of others' lives often makes their days monotonous, without being held accountable for this cost.

Summary

The history of economics comprises increasingly sophisticated observations of people's choices in markets and theoretical models suggested by these observations. Because successful negotiation in markets depends on assigning each good a value that depends on its scarcity and declines exponentially with delay, economic theories of choice have

converged on the model that assigns this value most effectively, rational choice theory (RCT). However, as physical discomforts (hunger, cold, pain) have become trivial problems in modern societies, most effort has been directed toward obtaining certain kinds of emotional experience and avoiding others. Events that are in limited supply still provide the occasions for these experiences, but since a person has some ability to assign meaning to these events and even to have the emotions without them, they have not been understood as goods (or bads) in a market. This understanding should be re-evaluated in light of mounting evidence that all of an individual's choices are determined in a single internal marketplace, and that this marketplace, in contrast to conventional markets, discounts delayed events hyperbolically rather than exponentially.

Hyperbolic discounting can explain the higher mental processes ("ego functions") that have heretofore not seemed derivable from elementary reward-seeking processes. It predicts a need for will, and offers a rationale for how willpower can arise simply from a person's interpretation of her own response to existing incentives. Hyperbolic discounting also leads to a theory in which emotions both generate reward and are selected by reward. Its implications make risk a positive factor in determining value, as well as giving scarcity a value beyond that created by its conventional role in the relationship of supply to demand. These constraints on emotional reward differ radically from the constraint of needing conditioned stimuli, the one that seems to be generally assumed. They make vicarious experience valuable in its own right, quite apart from the instrumental value of human relationships. Thus the *intertemporal* marketplace can be expected to create specific patterns of interaction with *interpersonal* markets, patterns that may be responsible for many anomalies such as bounded willpower and bounded self-interest that contradict RCT.

References

Ainslie, G. (1975) Specious reward: A behavioral theory of impulsiveness and impulse control. *Psychological Bulletin* 82: 463-496.

Ainslie, G. (1986) Beyond microeconomics: Conflict among interest in a multiple self as a determinant of value. In J. Elster (ed.), *The Multiple Self*, Cambridge University Press, pp. 133-175.

Ainslie, G. (1992) *Picoeconomics: The Strategic Interaction of Successive Motivational States within the Person*. Cambridge University Press.

Ainslie, G. (1995) A utility-maximizing mechanism for vicarious reward: Comments on Julian Simons "Interpersonal allocation continuous with intertemporal allocation" *Rationality and Society* 7: 393-403.

Ainslie, G. (2001) *Breakdown of Will*. Cambridge University Press.

Ainslie, G. (2005) Précis of *Breakdown of Will*. *Behavioral and Brain Sciences* 28: 635-673.

Ainslie, G. (2007) Can thought experiments prove anything about the will? In D. Spurrett, D. Ross, H. Kincaid and L. Stephens, eds., *Distributed Cognition and the Will: Individual Volition in Social Context*. MIT Press.

Barnes, A. and Thagard, P. (1997) Empathy and analogy. *Dialogue: Canadian Philosophical Review* 36: 705-720.

Bechara, A. (2006) Broken willpower: Impaired mechanisms of decision-making and impulse control in substance abusers. In N. Sebanz and W. Prinz, eds., *Disorders of Volition*. MIT Press, pp. 399-418

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

Beecher, H. (1959) *Measurement of Subjective Responses*. Oxford University Press.

Berlyne, D.E. (1974) *Studies in the New Experimental Aesthetics*. Hemisphere.

Berns, G. S., McClure, S. M., Pagnoni, G. and Montague, P. R. (2001) Predictability modulates human brain response to reward. *Journal of Neuroscience* 21: 2793-2798.

Berridge, K. C. (1999) Pleasure, pain, desire, and dread: Hidden core processes of emotion. In Kahneman, D., Diener, E. and Schwartz, N, eds., *Well-Being: The Foundations of Hedonic Psychology*. Sage.

--- K. C. (2004) Motivation concepts in behavioral neuroscience. *Physiology and Behavior* 81: 179-209.

Davidson, R. J., Putnam, K. M. and Larson, C. L. (2000) Dysfunction in the neural circuitry of emotional regulation: A possible prelude to violence. *Science* 289: 591-594.

Frank, R. H. (1988) *Passions Within Reason*. W.W. Norton and Company.

Gibbon, J.. (1977) Scalar expectancy theory and Weber's law in animal timing. *Psychological Review* 84: 279-325.

Glimcher, P. W. and Kable, J. W. (2005) Neural mechanisms of temporal discounting in humans. paper presented at the *Third Annual Meeting of the Society for Neuroeconomics*, Kiawah Island, SC, September 16.

Green, L. and Myerson, J. (2004) A discounting framework for choice with delayed and probabilistic rewards. *Psychological Bulletin* 130: 769-792.

- Harris, C. and Laibson, D. (2001) Dynamic choices of hyperbolic consumers. *Econometrica* 69: 535-597.
- Hollerman, J. R., Tremblay, L., and Schultz, W. (1998) Influence of reward expectation on behavior-related neuronal activity in primate striatum. *Journal of Neurophysiology* 80: 947-963.
- Iacoboni, M., Woods, R. P., Brass, M., Bekkering, H., Mazziotta, J. C. and Rizzolatti, G. (1999) Cortical mechanisms of imitation. *Science* 286: 2526-2528.
- Jolls, C., Sunstein, C. R., and Thaler, R. (1998) A behavioral approach to law and economics. *Stanford Law Review* 50: 1471-1550.
- Kahneman, D., Slovic, P. & Tversky, A. (eds.) (1982) *Judgment Under Uncertainty: Heuristics and Biases*. Cambridge University Press.
- Kavka, G. (1983) The toxin puzzle. *Analysis* 43: 33-36.
- Kirby, K. N. (1997) Bidding on the future: Evidence against normative discounting of delayed rewards. *Journal of Experimental Psychology: General* 126: 54-70.
- Lea, S. E.G. and Webley, P. (2006) Money as tool, money as drug: The biological psychology of a strong incentive. *Behavioral and Brain Sciences*.
- Lerner, J. S., Tiedens, L. Z. and Gonzalez, R. M. (in press) Portrait of the angry decision maker: How appraisal tendencies shape anger's influence on cognition. *Journal of Behavioral Decision Making*.
- London, E.D., Monterosso, J., Mann, T., Ward, A., Ainslie, G., Xu, J., Brody, A., Engel, S. and Cohen, M. (2006) Neural activation during smoking self-control. poster at the 68th annual meeting of the *College on Problems of Drug Dependence*, Scottsdale, AZ, June 20.
- McClure, S. M., Laibson, D. I., Loewenstein, G., and Cohen, J. D. (2004) The grasshopper and the ant: Separate neural systems value immediate and delayed monetary rewards. *Science* 306: 503-507.
- March, J. G. and Simon, H. A. (1958) *Organizations*. Wiley.
- Monterosso, J., Ainslie, G., and London, E.D. (2006) Delay discounting based on activation in the ventral striatum. poster at the 68th annual meeting of the *College on Problems of Drug Dependence*, Scottsdale, AZ, June 20.
- Piaget, J. (1937/1954) *Construction of Reality in the Child*. M. Cook, trans. Basic.
- Poundstone, W. (1992) *Prisoner's Dilemma: John von Neumann, Game Theory, and the Puzzle of the Bomb*. Doubleday.

Rachlin, H. (2002) Altruism and selfishness. *Behavioral and Brain Sciences*. 25: 239-296.

Rolls, E. T. (1999) *The Brain and Emotion* Oxford University Press.

Samuelson, P.A. (1937) A note on measurement of utility. *Review of Economic Studies* 4: 155-161.

Scitovsky, T. (1976) *The Joyless Economy: An Inquiry into Human Satisfaction and Consumer Dissatisfaction*. Oxford.

Shizgal, P., and Conover, K. (1996) On the neural computation of utility. *Current Directions in Psychological Science* 5: 37-43.

Volkow, N. D. and Fowler, J. S. (2000) Addiction, a disease of compulsion and drive: Involvement of the orbitofrontal cortex. *Cerebral Cortex* 10: 318-325.

¹ Many situations ranging in importance from ordinary courtesy to whether wars will be escalated have been negotiated as repeated prisoner's dilemmas, but the formal game was described only in 1950 (Poundstone, 1992).

² As it becomes clearer that physical indeterminacy would make choices feel random rather than willed, authors trying to fit the experience of an originating will have searched for a process that would make choice among known incentives unpredictable, even by oneself. I have shown how recursive self-prediction can put the observing self in the middle of a chaotic process and thereby satisfy this test (background and argument in Ainslie, 2001, pp. 129-134, and in press).

³ I will not deal with negative emotions here. I have argued elsewhere that the urge for grief or panic is like the urge for an addictive substance, only the cycle of reward and nonreward is extremely condensed (Ainslie. 2001, pp. 173-174).

⁴ Although anger is usually thought of as a negative emotion, it has many of the properties of a positive one (Lerner et.al., in press) and some people certainly cultivate it on a daily basis. There are people for whom even emotions that they consciously avoid seem to have a regular market value: the person who is constantly afraid of something despite the variability of actual threat (the customer of fear "mongers"), or always grieving at something.