

Getting Hooked

Rationality and Addiction

Edited by

JON ELSTER
OLE-JØRGEN SKOG

Introduction	1
JON ELSTER AND OLE-JØRGEN SKOG	
Addiction and Social Interaction	30
KARL OVE MOENE	
Addiction, Weakness of the Will, and Relapse	47
OLAV GJELSVIK	
The Dangers of Willpower	65
GEORGE AINSLIE	
The Neurobiology of Chemical Addiction	93
ELIOT L. GARDNER AND JAMES DAVID	
To Legalize or Not to Legalize: Is That the Question?	137
HELGE WAAL	
Rationality, Irrationality, and Addiction – Notes on Becker's and Murphy's Theory of Addiction	173
OLE-JØRGEN SKOG	
Gambling and Addiction	208
JON ELSTER	
A Visceral Account of Addiction	235
GEORGE LOEWENSTEIN	
Epilogue: Rationally Coping with Lapses from Rationality	265
THOMAS SCHELLING	

 **CAMBRIDGE**
UNIVERSITY PRESS

1999

The Dangers of Willpower

GEORGE AINSLIE

Addiction and dissociation are puzzling phenomena that are not usually thought of as related, although they are often prominent in the same individuals. Both violate our conventions of rationality. Addiction entails the voluntary choice of options that the person himself reports he does not want. Dissociation is a temporary reversal of preference so extensive as to change what the person acknowledges as his "self"; such reversal, for which the person is amnesic, is seen to an extreme degree in multiple personalities and to a lesser extent in fugue episodes, "blackouts," and the experience of spirit possession. Economic Man, the conventional utilitarian model of how people evaluate choices (Stigler 1980), explains neither, much less suggests any reason why they tend to appear together.

Empirical research on the relationship of dissociation with volitional lapses such as addictions has been sparse. There is a high correlation between dissociation and other self-control disorders: Of 100 substance-abuse patients, 39 percent have been reported to also have a dissociative disorder (Ross et al. 1992), as have 41 percent of 265 male veterans in a substance abuse program (Dunn et al. 1993). Patients with eating disorders also have high rates of dissociative disorder (McCallum et al. 1992, Zerbe 1993). Furthermore, a large proportion of drug addicts have post-traumatic stress disorder (e.g., 59 percent of a female drug rehabilitation population [Fullilove et al. 1993]), which entails many dissociative symptoms. However, these findings only suggest an important relationship between addiction and dissociation; they do not shed light on what it is.

The boundaries of the concept of addiction are in dispute, but the core meaning of the term is evident from its etymology: The term originally meant judicial enslavement, as when a person was sentenced (addicted) to serve another (*Oxford English Dictionary*), so that as a metaphor it clearly means enslavement by an appetite. Ample experience with addicts has shown that this enslavement comes not from a fear of withdrawal symptoms but from the continuing wish for a high, which often causes relapse after years of sobriety. Modern science has had increasing success in isolating the reward mechanism that makes the substance-based addictions extraordinarily attractive (Gardner 1992 and this volume), but this attractiveness does not explain the sense of slavery. Missing

is an explanation for the addict's ambivalence, that is, why he regularly pursues a goal that he says he wants to get away from.

Scientific theories have avoided confronting the conundrum of ambivalence: They hold either that the addict does not realize the contingencies he faces or that the pain is worth it to him; Economic Man has so little trouble maximizing his prospects that he cannot knowingly be enslaved. True, he can suffer from erroneous expectations and thus stumble innocently into a downward spiral of conspicuous pleasure followed by insensibly diminishing options, the "primrose path" described by Herrnstein and Prelec (1992b); and he can have a very steep temporal discount function, which makes him simply take no interest in the future, like the "rational addict" described by Becker and Murphy (1988). However, continual maximizer that Economic Man is, he cannot have the familiar human experience of looking forward with apprehension to a behavior that he knows he will regret; nor does he ever have a reason for buying disulfiram or naloxone, the means for limiting his own future choices to forestall such a behavior. In short, he cannot have motivational conflict beyond an uncertainty about magnitudes.

Thus, conventional utility theory could explain without further assumptions both a naive subject's entry into addiction and an experienced user's unambivalent preference for his substance use over sobriety; but it does not explain an addict's internal conflict, as evidenced in the extreme case by his attempts to restrict his own future freedom of choice.

Despite these difficulties, addiction could theoretically be explained by a variability in the addict's estimate of his future prospects, however unaccountable, in someone with consistent values. Dissociation is more mysterious by far; it represents a regular reversal of those values, a change that may block access to all memory of times when the person's values were different. Such changes of character might seem to be out of the person's control, a neurological calamity such as epilepsy, but actually they are well known to be motivated (Putnam 1989). If utility theory has had difficulty explaining addiction, it has not even attempted to analyze something as irrational as the harboring of whole contradictory selves.

I will show that it is possible to reconcile both addiction and dissociation with strict maximization of utility, in light of recent evidence that a basic assumption of conventional utility theory has been wrong. I will also deduce a relationship of these two seeming paradoxes with a reliance on willpower for self-control and suggest a relationship between them.

1. Picoeconomics Predicts Self-Defeating Behaviors

I have argued elsewhere (Ainslie 1975, 1992) that the exponential discount curve by which people evaluate delayed goods in conventional utility models cannot predict irrational choice. However, controlled

experiments have found a substantially different curve describing how the effect of a reward declines with delay, a curve that can be elicited in all higher organisms; this curve is not exponential but hyperbolic (Ainslie 1974, Ainslie and Herrnstein 1981, Ainslie and Haendel 1983). It is an aspect of Herrnstein's (1961) matching law, which has been confirmed by a number of researchers (Green, Fry, and Myerson 1994; Kirby and Herrnstein 1995; Stevenson 1986).

Hyperbolic curves predict that people will regularly form temporary preferences for the smaller but earlier of many possible pairs of alternative rewards, a preference pattern that in turn can account for a number of phenomena that had previously seemed paradoxical (an analysis I have called *picoeconomics*; Ainslie [1992]).¹ Many people balk at accepting a hyperbolic discount curve as fundamental, since at first glance it seems incompatible with the exponential discount curves that accurately describe a great deal of human economic behavior. However, the aggregate effect of hyperbolic curves on whole classes of choice can produce the appearance of exponential curves under some circumstances (Ainslie 1991). The identification of exponential discount curves as special cases within a framework of hyperbolic discounting liberates motivational theory from these curves' relentless prediction of rationality.

I cannot present a comprehensive exposition of *picoeconomics* here; but since several of its predictions combine to offer a rationale for addiction and dissociation, I will summarize them and refer the reader to specific parts of a fuller discussion (Ainslie 1992).

The effect of regular temporary preference is to make all motivated processes compete for survival against incompatible processes and to give them an incentive to strategically forestall these competitors. The "mind" or "self" is a population of reward-seeking operations that survive insofar as they actually obtain reward. The mental operations selected for by a particular reward or class of rewards constitutes the person's *interest* in that reward; interests within the person are very like interests within a community – those factions that are rewarded by ("have an interest in") the goal that names them (e.g., "the petroleum interest," "the arts interest"). Since a common marketplace of reward² keeps a person's purposes coherent except when conflicting rewards dominate at successive times, it makes sense to name an interest only where that conflict exists, that is, where the success of one process depends on its forestalling another process that would otherwise become dominant and undo the work of the first process. I would not be said to have separate chocolate and vanilla ice cream interests, even though these are often alternatives, because at the time when I prefer chocolate I do not increase my prospective reward by forestalling a possible switch to vanilla. However, I may have an ice cream interest and a diet interest, such that my expectable reward from the diet is intermittently threatened by an immediate prospect of ice cream. I do

not increase my prospective reward in either the long or the short range by defending my choice of chocolate against the possibility that I may change to vanilla; but I increase my prospective long-range reward by defending my diet against ice cream, and I increase my prospective short-range reward by finding evasions of my diet for the sake of ice cream. Whichever faction promises the greatest reward, discounted from its expected time of occurrence to the present moment, gets to decide my move at this moment; the sequence of moves over time determines which faction ultimately gets its way.

Where alternative rewards are available at different times, each will build its own interest, and one interest will forestall the other only if it can leave some enduring commitment that will prevent the other reward from occurring: If my diet interest can arrange for me not to get too close to ice cream, the discounted prospect of ice cream may never rise above the discounted prospect of the rewards for dieting, and the diet will have effectively won. However, whenever the value of ice cream spikes above that of dieting, the ice cream interest may suddenly undo the effect of many days of dieting. The ultimate determinant of a person's choice is not his simple preference, any more than the determinant of a legislature's action is simple voting strength; in both processes, strategy plays a major part.

This model of hyperbolic discounting might seem to make self-defeating behaviors such as addiction a straightforward problem. Once a person has identified a recurring preference as temporary, he should be able to commit himself not to give in. However, the solution is not that simple, as we will see.

Self-commitment is certainly a familiar behavior. Even pigeons have some ability to do this: After varying amounts of experience with their own strong tendencies to choose the poorer but earlier of two alternative food rewards, they can learn to forestall these tendencies somewhat. If a particular route to the largest available reward goes through some point A, where they can choose a poorer, earlier alternative reward, many subjects learn to take the longer or more effortful route through a different point B that avoids this peril (Ainslie 1974). To some extent, pigeons can even achieve this internally, that is, without the experimenter offering them a route through point B that physically keeps them away from point A (Mazur and Logue 1978).

This last finding raises the question of what other committing tactics are possible, for pigeons or for people. Perhaps these subjects have learned to keep their attention away from cues that, once seen, offer the choice of a temporarily more valued smaller reward; perhaps they are learning to think "cool thoughts" instead of building their appetites for the reward, as the Mischels' five-year-old children report doing (Mischel and Mischel [1983]; more discussion in Ainslie [1992, pp. 130-42]). The most

interesting possibility is that they are forming elements of willpower – resolutions or intentions as the philosophers of mind picture them (Bratman 1987). That is, they may be evaluating choices as members of larger categories.³ For exponential discounters, this would have no effect on choice, but for hyperbolic discounters it would let the disproportionate effectiveness of an immediate temptation be averaged in with the more objective, albeit discounted, evaluations of many similar prospective choices in the future. Categorical choice might allow subjects to evaluate and reject temptations with "both alternatives steadily held in view," as William James (1890, p. 534) characterized the effort of will.

The stabilizing effect of deciding "on principle" has been recognized since Aristotle (Barnes 1984, 1,147a24-b17) and was well known to the Victorian psychologists (see Sully's quote in note 3, as well as James [1890, p. 534]). It has recently been verified by the same kind of experiments that demonstrated the matching law: Even pigeons will choose a larger, later reward over a smaller, earlier one more often if this choice is part of a series of choices than if it stands alone, an effect that Rachlin (1995) has attributed to his subjects' adoption of a "molar" (i.e., global) framework of the choice in place of a "molecular" (i.e., one-by-one) framework. Similarly, Heyman (1996) has demonstrated how a consistent signal associated with the better long-range choice leads pigeons to increase that choice, a change he explains in terms of the signal's having supplied his subjects with an "overall" versus a "local" perspective on the outcomes.

However, the increase in both pigeons' and children's capacity to wait for the larger reward is small in these experiments (usually less than double). There is reason to believe that even grouping choices together in series, while necessary for genuine willpower, will not be sufficient. That is, there is nothing in the process of molar or overall categorization per se that prevents a short-range interest from distinguishing each current choice from the molar category. Choices in potential series will all have unique as well as common features, which can be a basis for excepting them; addictive motives are notorious for subverting general plans of conduct by establishing the example at hand as a special case (as in James's famous list of drunkard's excuses [James 1890, p. 565]). Somewhere, perhaps in the long exploration of social rules that takes place in the grade-school years (Piaget 1965), humans discover the power of the test case, of focusing on their own choice between molar and molecular interpretations of a current example as a predictor of what they can expect of themselves in similar cases in the future.⁴ People soon learn that the outcome of such tests weighs on subsequent choices that look similar and that the threat of losing one's expectation of future self-control can marshal a great deal more motivation than depends on the single case at hand. Intention policed by such a strategy is experienced as

something transcendent: will, or resolve, or moral determination. Perhaps this is even the "knowledge of good and evil" that gave Man self-will and mythically ended his career in the Garden of Eden (Genesis 3).

The perception of test cases, in individuals or in interacting groups, gathers motivation remarkably. Relevance to a diet greatly changes the incentives for eating versus not eating a single piece of candy, just as fluctuations in the price of a "bellwether" stock have a disproportionate effect on the whole stock market.⁵ Furthermore, it can be shown that perception of a series of choices as test cases for a common rule should shape the relationship of successive choosing selves into a familiar relationship, limited warfare (Schelling 1960, pp. 21–80), the properties of which are described by a well-known bargaining game, the repeated prisoner's dilemma (Ainslie 1992, pp. 142–79). This is to hypothesize that willpower achieves deterrence against each individual impulse through the fear of setting precedents for whole strings of future impulses. The *personal rules* that define what is to be seen as cooperation in this bargaining process seem identical to the principles specified in Kant's categorical imperative (Kant 1960, pp. 15–49) and Kohlberg's highest stage (VI) of morality (Kohlberg 1963), and thus they appear at first glance to represent rationality itself. Certainly they give molar or overall views of reward great leverage over molecular or local ones. It might indeed seem that they should solve the problem of addiction.

2. Willpower Is an Awkward Expedient, Not the Ultimate Rationality

Unfortunately, a person's perception of the relationship of the prisoner's dilemma and the willpower that results from this perception do not really solve the problem of temporary preference. Willpower may be the best expedient we have⁶; but it turns out to have serious side effects that become major factors in people's lives. I argue that these side effects are what shape the problem of addiction as the clinician encounters it in a developed society; that they link addiction to other phenomena, particularly dissociation, that seem anomalous in any rigorous theory of behavior; and that they are little recognized, leading even the best-intended advice to produce perverse results both in individuals and in social policies.

Although personal rules can sometimes lead an individual to approximate an exponential discount curve for future events, they do not give him the prelapsarian trait of maximizing his expected utility whenever he makes a choice, nor do they preserve the experience of spontaneity in choice making. Rule-governed choice has four properties that cause it to fall short of these utilitarian ideals:

1. Categories of choice come to overshadow their individual members, resulting in a legalistic manner of choice making.

2. Rules make lapses disproportionately damaging and tend to turn them into permanent symptoms.
3. Rules create a motive to distort the perception of reality.
4. Rules need not be in the person's longest range interest.

Categories of Choice Overshadow Their Individual Members

The perception of choices as precedents often makes a choice much more important for its expected effect on future choices in the intertemporal prisoner's dilemma than for the rewards that literally depend upon it. Insofar as this is true, choice becomes detached from the properties of its objects and takes on an aloof, legalistic quality. In the parlance of existential therapists the person becomes "inauthentic" (Kobasa and Maddi 1983).

A person is often hard put to predict how his own future selves will perceive the precedent his current choice is setting. This makes cooperation in intertemporal prisoners' dilemmas both rigid and tenuous. Unless chance provides clear lines that can be used as boundaries it may be difficult to tell whether, facing a choice in the future, one will look back at the current choice and judge it to have been a lapse.

The difficulty of this task depends on topography of the range of options. A person trying to give up a heroin habit at least benefits from the *bright line*⁷ between some heroin and no heroin; but a person trying not to overeat has to make judgment calls continually about what food to allow himself, even if he has committed himself to one single diet. In consequence a short-term interest within the individual can usually offer a colorable exception to a diet, and it may escalate such proposals by degrees until the diet has been rendered useless without ever having been clearly violated.

Under the pressure of temptation, a person may claim an exception to a rule but later see the claim as unsupportable, that is, see himself as having had a lapse. Conversely, he may be cautious beyond what his long-range interest requires, for fear that he will later see his choice as a lapse – a rationale that exacerbates compulsiveness. Errors in either direction impose costs that would never result from the exponential curves of Economic Man, since these curves would create no reason for discerning precedents in the first place.

Rules Magnify Lapses

When a person violates a personal rule, the cost is a fall in his prospect of getting the long-range rewards on which it was based. This prospect is what he uses to stake against the relevant impulses; a lapse suggests that his will is weak, a diagnosis that may actually weaken his will in the recursive pattern described long ago by James, Lange, and Darwin (James

1890, vol. 2, p. 458; Ainslie 1992, pp. 200–5). To save his expectation of controlling himself generally, he will be strongly motivated to attribute the lapse to a particular aspect of his present situation; the consequence may be that he abandons his attempts at willpower whenever that aspect is present. That is, he may create a class of exceptions rather than suffer a general loss in the credibility of his will – to decide that he *can't* resist the urge to panic when speaking in public, or to lose his temper at incompetent clerks, or to smoke after meals. Thereafter the occasions for these urges will seem to bring on the behaviors automatically, without an intervening moment of choice.

In other words, exceptions to personal rules need not arise through deliberate rationalization; an exception can impose itself on the person's intentions in the most awkward of places, wherever a lapse has threatened a broad loss of impulse control. When she has been using her willpower to control panic, or rage, or smoking in a particular setting, and her will has been overwhelmed, she will be under pressure to distinguish that setting from all others so that she will not see her lapse as a precedent for her choices, in general. Her discrimination of this special circumstance has a perverse effect, since whenever she faces it she sees only failure predicting further failure. I have called this area, where the person dares not attempt efforts of will, a lapse district, by analogy to the vice districts in which Victorian cities tolerated the vice that they lacked the political will to suppress (Ainslie 1992, pp. 193–7). Where the encapsulated impulses are clinically significant, clinicians call a lapse district a symptom. In this way the perception of repeated prisoner's dilemmas stabilizes not only long-range plans but lapses as well.

The availability of boundaries to circumscribe a potential lapse district may determine whether it will form in the first place. Lapses are goal-directed behaviors, and thus they are not beyond the influence of long-range rewards; lapses endanger these rewards less when there is a boundary that can limit their implications. Short-range interests prosper most when they parasitize, as it were, rather than kill, long-range interests. Thus the uniqueness of a vacation trip may make a spendthrift more apt to go off a budget; likewise, an alcoholic may find lapse districts forming in time-limited circumstances (e.g., only when she has just completed a project or a school year or an election campaign).

The boundary may be intrinsic to the kind of reward involved. Many "out of control" consumption patterns commit the addict to withdrawal sickness followed by a period of abstinence before the substance becomes rewarding again. The crude periodicity of such a cycle may permit the addict to stably apportion his ambivalence into two successive, contradictory activities. Then the trigger for a binge may be nothing more than a perceived lapse itself: His willpower governs until it stumbles but, as soon as he has a drink or fix, he experiences a "loss of control" and winds up on a bender (Evenson et al. 1973; for a similar loss of control in dieting,

see Polivy and Herman 1985); his will when sober will be preserved by his perception of the drunken state as a lapse district. In the period of exhaustion that ends the bender, his "normal" will can reassert itself, but only until a regenerated appetite for the addictive activity again asserts itself in a lapse. He will have two distinct selves, an intoxicated one and a sober one.⁸ By contrast, a rapidly satiating activity such as smoking will never reach the proportions of a binge; here, any regular alternation between control and loss of control must be cued by something other than intrinsic exhaustion and recovery.

Rules Motivate Misperception

Personal rules depend heavily on perception – on noticing and remembering one's choices, the circumstances in which they were made, and their similarity to the circumstances of other choices. Because personal rules organize great amounts of motivation, they naturally create an incentive to suborn the perception process. When a lapse is occurring or has occurred, it will often be in both the person's long- and short-range interests not to recognize that fact: His short-range interest is to keep the lapse from being detected so as not to invite attempts to stop the forbidden behavior. His long-range interest is also at least partially to keep the lapse from being detected, because acknowledging that a lapse has occurred would lower the expectation of self-control that he needs to stake against future impulses.

After a lapse, the long-range interest is in the awkward position of a country that has threatened to go to war in a particular circumstance that has then occurred. The country wants to avoid war without destroying the credibility of its threat and may therefore look for ways to be seen as not having detected the circumstance. The person's long-range interest will suffer if he catches himself ignoring a lapse, but perhaps not if he can arrange to ignore it without catching himself. This arrangement, too, must go undetected, which means that a successful process of ignoring must be among the many mental expedients that arise by trial and error; these are retained simply because they make the person feel better without his realizing why.⁹ Money disappears despite a strict budget, and people who "eat like a bird" mysteriously gain weight. Clouding of consciousness in the face of temptation has been reported by observers from Aristotle (Bogen and Moravcsik 1982) to Sjoberg and Johnson (1978). Here is a motivational pattern that could easily create a black market, indeed an underworld, of those behaviors that include blockage of notice and recall.

Rules Need Not Serve the Person's Longest Range Interest

There is a puzzling class of behaviors that are more stable than the addictions, but that a person still perceives not to be in his "best" interest.

Whereas bouts of addictive behavior are preferred for periods of minutes to days, these more stable behaviors are preferred for years, or indefinitely, but are still accompanied by a sense that they will be regretted. Compulsions, dieting in anorexia nervosa, and making oneself impervious to emotion in alexithymia (Nemiah 1977) are clinical examples, but many highly systematic "character flaws," such as miserliness, fastidiousness, and overcontrol, never come to a clinician's attention. These maladaptively narrow behaviors lack a generally accepted name; I have called them *sellouts*.¹⁰ Like addictions, sellouts create a sense of enslavement, but they are apt to be more systematic and may even seem rational, because they are often based on personal rules. That is, sellouts can combine against both longer and shorter range interests.

Although the combinatory mechanism of personal rules obviously requires them to serve longer range interests against shorter range ones, they need not serve the person's longest range interests. Some rules may even group such mid-range payoffs together so as to steal dominance from longer range payoffs, without violating the strict logic of hyperbolic discounting (Ainslie 1992, pp. 216–24). The relationship of a person's longest range interest with sellout- and addiction-range interests is illustrated by Jon Elster's example:

I wish that I didn't wish that I didn't wish to eat cream cake. I wish to eat cream cake because I like it. I wish that I didn't like it, because, as a moderately vain person, I think it is more important to remain slim. But I wish I was less vain. (1989, p. 37 note)

Assuming that his wish to be less vain does not merely serve his wish for cream cake, it reveals a perception that vanity is not in his longest range interest – that he will regret it even though it dominates his current choice. That is, he sees vanity as an overly narrow trait, a sellout. His vanity is in turn threatened by a tendency to eat cream cake; assuming that episodes of seeking cream cake dominate his choice for a matter of minutes to hours, they represent temporary preferences of roughly the same duration as those in addictions.¹¹ Under these circumstances the vanity will be able to foster a personal rule against eating the cake (i.e., a diet), even though it may contradict the still longer range interest in not being vain. This will be true as long as this longest range interest cannot find effective criteria that would make a rule against the vanity practical. It is on the availability of such criteria (bright lines) that contests among relatively long-range interests seem to depend.

In seeming inversions of the interests' power relationships – where a sellout-range interest that is protected by a rule dominates the person's longest range interest – the decisive factor may not be how well a rule would pay if enforced, but how well it can be enforced. Enforceability depends in turn on how clearly each choice is either a member or not

a member of the relevant series.¹² That is, personal rules operate most effectively with countable criteria. Thus rules against acts of vanity are apt to be much less enforceable than rules against overeating.

The impact of having rewards marked by discrete stimuli is illustrated by recent experiments on "melioration": Human subjects had to choose between amounts of money such that choice A produced a conspicuously larger reward than choice B, but choice A led to poorer subsequent payoffs for both choices A and B (Herrnstein and Prelec 1992a). Where choice of a larger amount reduced the *amounts* to choose from on subsequent turns, most subjects soon discovered the strategy of picking the smaller amount in the current choice. However, where choosing the larger amount led not to smaller amounts but to greater *delay* before subsequent choices, thus reducing total income in a game of fixed duration, subjects tended to keep picking the larger amounts and getting smaller returns (Herrnstein et al. 1993). Amounts are eminently countable; delays are vague unless someone specifically points to the seconds. As we might expect, when Herrnstein thinks to count the subjects the greater delay that came from choosing the larger reward, they also generally started choosing the smaller.¹³

Thus when sellout-range interests are based on well-marked rewards, and their richer, longer range alternatives are harder to define, the sellouts may win the protection of personal rules. Rules that serve sellouts are particularly apt to become dominant when the person has had a conspicuous addiction-range interest that has motivated strict controls, as in the case of the anorectic patient who finds his draconian diet worthwhile to end a history of overeating. It is easier to enforce specific rules about diet than more subtle rules such as "eat what you need" or "eat what you will be glad of in retrospect," though if the latter were enforceable rules they would permit the most reward in the long run. When a person seeks the comparative safety of the most clearcut criteria for his personal rules, he may be forestalling not only short-range impulses but also the richest long-range rewards.

Extreme commitments to personal rules are widely recognized as compulsions. But even short of frank compulsiveness, the systemization that lets personal rules recruit the greatest motivation may undermine the very purposes that inspired them. The attempt to optimize our prospects with rules confronts us with the paradox of definition: That to define a concept is to alter it, in this case toward something more mechanistic. A person who concludes that he should maximize money becomes a miser; one who rules that he should minimize his openness to emotional influence becomes alexithymic (Nemiah 1977); if he concludes that he should minimize risk, he becomes obsessively careful; and so forth. The logic of rules may so come to overshadow a person's responsiveness to experience that his behavior becomes formal and inefficient. A miser

is too rigid to optimize his chances in a competitive market, and even a daring financier undermines the productiveness of his capital if he rules that he must maximize each year's profit (Malekzadeh and Nahavandi 1987). Similarly, strict autonomy means shielding oneself against others' ability to invoke his passions, but the alexithymics who have accomplished this cannot effectively use the richest strategy available for maximizing reward, the cultivation of human relationships (Ainslie, 1995 and in preparation). And avoidance of danger at any cost is poor risk management.

Furthermore, beyond the inefficiency of hidebound rules in obtaining their ostensible goals, picoeconomic analysis predicts a reduction in their emotional payoff that occurs insofar as they *do* succeed: Optimal exploitation of the available reward for any gamelike activity requires the maintenance of suspense, but hyperbolic discount curves create a short-sighted impatience for resolution that can be effectively resisted only by committing oneself to be surprised (Ainslie 1992, Ch. 7). That is, imagination attenuates through an innate impatience to entertain its high points at the cost of prematurely resolving emotional tension, thus increasingly wasting the relevant appetite as the person gains experience with a given situation. The more efficient a person learns to be at overcoming a challenge, or the more familiar he becomes with a story, the more he will anticipate its resolution and undermine the longing needed for relatively intense satisfaction. Thus the various mental diets that let the person maximize wealth, or autonomy, or safety reduce by their very efficacy the amount of surprise he is subject to. Since surprise is the key factor in appetite for emotional experience, the maximization process is apt to become stale, even if it is working well. This seems to be how nature keeps organisms exploring novel areas instead of resting on their laurels. A consequence is that clarity per se predisposes rules for pacing emotional rewards to serve sellout-range interests.¹⁴

Of course, in many cases a person never regrets maximization schemes such as anorexia and miserliness, so that these are held to be inferior choices only in an observer's estimation (raising, as with alleged addictions, that the person himself does not recognize, the question of tastes); but often the person regrets them intermittently, or blames them for spoiling his happiness without being able to devise a rule structure sufficient to motivate their abandonment. Perhaps the person initially tried less concrete rules, such as to do what he would be glad of in retrospect, and found they could not hold their own against more specific prescriptions; his "consistent preference" for the more concrete rules is then just acquiescence in their seeming inevitability.

Personal rules convert changes of preference that depend on proximity to temptation into the simultaneous conflicts that seem more numerous (Ainslie 1992, p. 193). Getting drunk no longer depends on how close one gets to a bottle but on whether an ongoing search for excuses finds a

believable one. With characterologic choices such as anorexia or miserliness, the temporal basis of the conflict is harder still to see, but the payoffs even for these abstemious behaviors are clearly a current sense of some kind of well-being, at the expense of any realistic expectation of feeling glad of one's choices later.

In these four ways, elementary motivational conflict is sharpened and its stakes raised when it is structured by personal rules. After the need for clarity has taken its toll on subtlety, and overcaution on flexibility, and undercaution with its consequent failures on resolve and self-observation – in short, after the makeshift nature of our attempt at global perception has caught up with us – the conflict between global and local approaches to choice making is no longer simply one of long- versus short-range interests. Given the distorting effects of both hyperbolic discounting and the personal rules that compensate for it, "rationality" becomes an elusive concept. Insofar as it depends on personal rules demanding consistent valuation, rationality means being systematic, but systemization readily goes too far and winds up advancing a person's mid-range interests against her longest range ones. Sometimes a particular global approach is less productive in the long run than local, myopic alternatives, as anorectics and misers discover to their bewilderment.

3. Side Effects of Willpower Foster an Interaction of Dissociation and Addiction

Experiential underworlds (classes of goal-directed behavior that the person cannot report) have often been described, but theorists have been hard put to make sense of them in motivational terms.¹⁵ Modern parlance characterizes the formation of underworlds as "dissociation," a person's blockage of information about one experience while he is having another. Hypnosis experiments (Hilgard 1977) and observations of unusual natural phenomena such as multiple personalities (Putnam 1989) and psychic amnesia (Schacter and Kihlstrom 1989) have characterized the experiencing entities (the various "selves") as parallel and potentially autonomous, with radically divergent value systems. One of these value systems is almost always conspicuously impulsive, whereas another is conventionally prudent.

Current belief about dissociation attributes it to a combination of innate cognitive predisposition and a strong acquired motive to avoid intrusive traumatic memories (Spiegel 1991). The relation between these two factors is in dispute, but there is usually agreement that dissociation is motivated by the dysphoria of intrusive thoughts. I will refer to this as *hypothesis A*.

This hypothesis suggests only a minor relationship between dissociation and addiction: Concomitant abuse of some substances may physiologically augment the repression, but nonsubstance addictions should

neither cause repression nor be caused by it. I will briefly comment about recent research on this mechanism and then suggest two additional motives for dissociation that arise from personal rules and have a more intimate relation to addiction-range motives: Dissociation can protect useful personal rules (*hypothesis B*) and may offer relief from oppressive ones (*hypothesis C*); in both cases it would do this by strengthening the encapsulation of lapse districts, which would in turn provide occasions for the dissociations. Any of these three mechanisms might operate side by side with any of the others.

Dissociation Avoids Intrusive Thoughts

Some individuals seem to have more of a bent for dissociating than others. Perhaps 10 percent of a general population have spontaneously entered experiences that they could not report afterward or that they discovered only by external evidence – “waking up” in a different place or being told of it; 2 percent do so repeatedly (Ross, Joshi, and Currie 1990). The latter are said to have a dissociative disorder. Observers have long noticed that dissociation seems to go along with an unusual aptitude for directing one’s attention. High dissociators do not readily compare one part of their experience with other parts but become totally “absorbed” in one to the exclusion of others (Crawford, Brown, and Moon 1993, Roche and McConkey 1990). Hypnotizability is another associated trait. Whether from heredity (Crawford and Gruzelier 1992) or learning, about a fifth of the population can follow a suggestion to experience events that they cannot remember at other times (Hilgard 1965). These highly hypnotizable people also report a great degree of absorption (Smyser and Baron 1993).

The relationship of dissociation with these attentional factors is not surprising. In motivational terms, hypothesis A states that dissociation represents an extreme use of attention control – the restriction of information gathering – as a tactic for controlling very short range urges¹⁶: Those trauma victims who have predispositions to dissociate use their powers of absorption to divert their attention from the triggers of intrusive memories, until the strength of the unsatiated intrusive urge is enough to attract it back; and during their subsequent surrender to these memories they protect their long-range interests from association with them by repressing those interests in turn.

The source of the intrusive memories has seemed obvious: A large percentage of dissociators report a history of trauma (Putnam 1989), and a large percentage of people with post-traumatic stress disorder (PTSD) dissociate (Branscomb 1991). PTSD is increasingly recognized. It is now said to affect 12 percent of the population, or at least of its females, at some time, and 5 percent at any given time (Resnick, Kilpatrick, and Dansky 1993). However, the role of trauma in dissociation is complicated by the relationship of PTSD with another trait, fantasy proneness. There is

a group of people who readily generate vivid fantasies of both rapture and dread; 1 to 2 percent of the population are said to be capable of fantasy experiences comparably vivid to those occasioned by actual physical stimuli, even food and sex (Wilson and Barber 1983, Rhue and Lynn 1987). Like PTSD patients, the fantasy-prone have intrusive thoughts, which are easily triggered and which may undermine long-range activities that require sustained attention.¹⁷ However, the fantasy-prone are not merely like PTSD patients: They often *are* PTSD patients, reporting histories of trauma at an extremely high frequency (Lynn and Rhue 1988).

It has been impossible to say to what extent fantasy-proneness is a way people learn to cope with urges to entertain real traumatic memories and to what extent an innate fantasy proneness lures the person into elaborating or fabricating trauma fantasies, which then have the vividness of memories. Certainly some people form convincing memories of demonstrably impossible experiences such as prior lives, and they can do so either spontaneously or under hypnotic direction (Spanos et al. 1991, Labelle et al. 1990); however, a large proportion of people who report childhood traumata can find independent corroboration of them (Herman and Schatzow 1987), although this has not been studied specifically in people who are fantasy-prone. Whatever the interaction of trauma and fantasy, we know that fantasy proneness is closely associated with attentional absorption (Rhue and Lynn 1989), raising the additional question of whether an aptitude for directing attention creates exceptional fantasy abilities or whether a proneness to exceptionally vivid fantasies motivates attention-controlling strategies such as high absorption. Hypothesis A is obviously complex and in need of systematic study. In any case, it should probably be broadened to include intrusive thoughts that do not come from trauma.

Dissociation Protects Personal Rules in the Person’s Longest Range Interest

We would expect the processes of partitioning volitional lapses (side effect 2 of willpower, above) and repressing them (side effect 3) to reinforce each other, in that each limits the scope of personal rules. Either of these responses (the lapse-district or the information avoidance) may occur by itself, but there is no reason that one need exclude the other. The cue that warns the person not to attempt willpower may tell him to block access to memory as well: Defining a lapse district will not completely protect the person’s other prospects for impulse control, since even after distinguishing the circumstances of a lapse the person cannot be sure of how important his future selves will think the distinction is; it is hard to be sure that future selves will not see predictive implications in that lapse. However, forgetting a lapse will end its particular danger, albeit at the expense of strengthening the person’s tendency to forget lapses in

the future. Thus there is the potential for a synergy of repression and lapse-district formation.

Once a lapse district has become established, its circumscribed extent will invite the person's other short-range interests to exploit it as an opportunity for expression. For instance, the alcoholic's binges offer not only a limited period of drinking but also an occasion for a vacation from other rules. Other temptations that have heretofore been under tenuous control can thrive using the period of the binge as an exception to the rules that forbid them, a holiday like Saturnalia. Modified rules such as, "I will be faithful, respectful, and prudent, but not when the alcohol is talking," are supported by a reduction in the pressure of urges between binges, which may make up for the aftereffects of the periodic adultery, rage, and/or dissipation, particularly if other people accept the attribution to the alcohol. The drunken and sober selves will take on a whole panoply of different characteristics, often under the general rubric of hero and villain.

The greater the scope of the rules threatened by the person's impulsive behaviors, the more necessary the blockage of memory to maintain his confidence in his will when sober, that is, the more he will be motivated to use dissociation to further separate his lapse district from the rest of his life. Although a person may learn to accept a circumscribed weakness of will against a "normal" temptation such as drinking, he might find that a disgusting paraphilia or outbursts of vicious rage at someone he loves require ignorance. A person who is vegetarian because of a concern for world food supplies might not be threatened by episodes of eating meat and might even recognize periodic occasions for it; but for one who is vegetarian out of a tenet that killing animals is murder, eating meat would threaten a larger belief structure, and he would be motivated to "not be himself" during any such episode. Dr. Jekyll could not accept Mr. Hyde, even as a lapse. A person with enough blockage will have a "split ego."

Thus some dissociations may be extreme forms of that spontaneous pressure valve, the lapse district. Just as a lapse district itself offers other impulses the occasion of a natural exception to rules, the concealment from these rules afforded by dissociation will further attract the other major addiction-range interests that the person has tried to outlaw. Once an exception is thus doubly protected from the review of personal rules, it becomes a formidable focus for organizing other impulses. Thus dissociated selves are apt to become very bad citizens, but primary selves will be proportionately virtuous.

Dissociation May Evade Rules That Serve Sellout-Range Interests

There is a more speculative possibility that sometimes appears in the lore of addictionology: That a person's will may have become so confining that

a pattern of regular lapses actually makes him better off in the long run. This lore attributes bingeing to a patient's inhibitedness in the rest of his life; his general overcontrol is said to set up periodic episodes of breaking loose.

The model of intertemporal bargaining predicted by hyperbolic discount curves provides a specific rationale for this pattern: Ruling out any large source of emotional reward will create a proportional motive for the person to bypass or break his rules. If these rules have, in William James's phrase, "grown too narrow for the actual case" (James [1890, p. 209]; side effect 2, above), even his long-range interest will lie in partially escaping from them. Thus personal rules that serve sellout-range interests potentially create alliances between long- and short-range interests. The person's occasional binge comes to serve as a correction to the comparative sterility of such rules, a means of providing richer experiences while still limiting the scope of behaviors that pay off in the addiction range. The longest range interest of an alcoholic who is too rigid when sober may be to tacitly foster the cycle of drunkenness and sobriety, rather than be continuously imprisoned by his rules.

An alcoholic is sometimes described who becomes nicer, or more genuinely creative, or more fully human when drunk. Furthermore, some addicts plan binges in advance. Such people may believe that their binges are undesirable (indeed, "rationality" will almost certainly dictate such a belief), but the therapists they hire find them mysteriously unresponsive to treatment. The patient who arranges for drinking several days in advance – goes off his disulfiram, for instance, or brings bottles with him to his rehabilitation program for later use – cannot simply be yielding to a short-range impulse. This is behavioral evidence that he experiences his rational plans as sellouts, which, even at a distance, appear to need hedging, although he may be unable to report any such thing.

Evasion of sellout-based rules is thus a third incentive for dissociation, beyond the avoidance of intrusive thoughts and the protection of useful personal rules.¹⁸ Insofar as this mechanism is important, a simplistic policy of "the more willpower, the better" will contradict the experience of many addicts. They are able to "listen to reason" only when they have somehow learned to ease the strictures that starve their own long-range interests.

Pervasive personal rules may arise in response not only to extraordinary temptation or ineptness in dealing with temptation but also to an environment's extraordinary demands for systematic decision making. As a society develops a complex, interdependent economy, it generates both better means for individuals to maximize their efficiency and more pressure to do so. This is not to say that one has to obey particular people or otherwise narrow one's choices; on the contrary, the range of substitutable behaviors grows enormously. But a person who

chooses these substitutions on a basis that ignores economic advantage loses conspicuous benefits.

Even as parents and rulers have less control, the logic of an increasingly comprehensive marketplace has more. The implications of a person's choices comes to extend beyond a set of related behaviors in a circumscribed context (how to satisfy this customer, make love to this partner, fix this machine) to most of one's behaviors in a common universe, and the behaviors of many other people whom he may never meet. I list three of the many manifestations of this progression: First, cash prices and wages, which make disparate decisions comparable, have penetrated choice making ever more minutely, so that goods and services that used to be bartered as part of relationships are increasingly paid for. It is indicative that the smallest unit of money in use was half a day's wages in medieval times but is now barely worth picking up from the floor (ironically, the penny in both cases; Burnett 1969). Second, and similarly, the long-term records that result from one's behavior, which once consisted (beyond a few major documentations such as births, marriages, and deeds) of the emotional memories of neighbors, are now automated, quantified, and increasingly collated. The consequence is greater comparability of past and current choices; job reports, credit ratings, and traffic tickets issued a decade ago are increasingly available and used to predict a person's behavior. Third, formal acculturation by schooling is increasingly lengthy and gets increasingly audited according to standards. A person's performance in one subject or at one time can be compared to her performance elsewhere, or to another person's, with increasing precision, comparisons that are apt to determine her career choice, her advancement, or at least her morale. In these and many other ways, each of a person's choices is made comparable to, and thus predictive of, a range of other choices. Attention to this aspect of choice is rewarded by greater efficiency in any systemizable endeavor; but, as we have just discussed, it may lead to less productive occasioning of reward in the long run.

These changes do not feel forced upon us by alien interests. On the contrary, they are means to the ends we ourselves have defined, means that are increasingly selected for efficiency by the competition of the marketplace. We freely adapt our personal rules to be compatible with them; but these rules may come to serve sellout-range interests.

Attuned as we are to modern efficiency in the developed world, we do not recognize the oppressiveness of an environment so rationalized that much of our natural idiosyncrasy has been anticipated and harnessed or selected out. Yet newcomers accustomed to more backward economies find our relationships superficial; and our own fictional heroes are increasingly those who rage against systems. The costs of basing decisions only on countable outcomes never appear in cost-benefit analyses; indeed they are recognized only obliquely, as anomie or other imponderable

moral problems, part of the "X-inefficiency" of workers that somehow limits their responsiveness to monetary incentives (Leibenstein 1976). Only within the past two decades has there been retrenchment where systemization is most advanced – in large corporations – in the form of decentralization, quality circles, self-directed work teams, and other retreats from the conventional approach of time/motion efficiency (Macdonald and Piggott 1993).

The systemization of modern society may be making the evasion of personal rules an increasingly important motive for dissociation. Certainly the apparent increase in multiple-personality disorder in modern times has been perplexing. The condition was known in Victorian times and as demonic possession even earlier (Ellenberger 1970, pp. 13–22). Granted that it has recently received a great deal of publicity, if there were a substantial incidence of it in earlier days, one would expect there to have been more discussion of it. The supposed causative agent, trauma in childhood, must have been more widespread in earlier generations, an increase in reportage notwithstanding: Historical cultures encouraged physical discipline, neglected children's experiences (Aries 1965), and had scant means or will to interfere with abusive families. The violence of the life children saw in earlier times is attested to by the cruelty of their games.¹⁹ If trauma were the main cause of dissociative disorder, we would expect the incidence to be decreasing, and this, at least, does not seem to be happening. The evasion of an increasingly systematic world is a speculative cause at best, but it makes some sense. Dissociation in such a cause would represent another alliance of convenience between addiction-range and long-range interests.

4. Summary

The strategy of controlling temporary preferences by perceiving one's choices as sets of precedents is far different from the rationality of exponential discounting envisioned by conventional utilitarians. Instead of being an elementary process, that rationality seems to be a legal fiction, a convention that has been shaped by practical experience to be the most useful goal of the intertemporal bargaining process. We undertake to make choices as if this goal represented our spontaneous preferences, and perhaps even to suppress our awareness of this very undertaking. We do so to give our long-range interests an advantage over shorter range interests and to give ourselves an advantage over economic and social competitors, where self-control is a factor in the competition. However, this rationality is a somewhat clumsy artifice, seemingly cobbled together in the form of personal rules without having evolved in lower species, as soon as the human race realized that our spontaneous preferences did not maximize our long-range prospects.

I have described four side effects of personal rules that interact to produce two behavior patterns that we find pathological, patterns that in turn interact to stabilize each other. The four properties are

1. rigidity leading to legalism;
2. magnification of lapses, leading to encapsulated symptoms;
3. perceptual distortion, leading to limited self-knowledge;
4. inordinate dependence on concrete criteria, which suborns personal rules away from the person's longest range interests.

The two pathological processes are addiction and dissociation. My analysis of their interaction uses the piceoeconomic concept of durations of temporary preferences, in particular, addictions (hours to days) and sell-outs (months to years).

The artificiality of personal rules (side effect 1) may make them seem like prisons. Formation of lapse districts (side effect 2) leads to stable behavioral "symptoms" with an addiction range of preference. Distortion of perception (side effect 3) may both conceal and strengthen lapse districts. The selective advantage that rules with concrete criteria have over those without (side effect 4) often leads to the dominance of rules that serve sellout-range interests, an effect that enhances the sense of imprisonment. When a personal rule no longer serves a person's longest range interest, lapse districts that have broken away from this rule ironically may start to serve this interest, forming an alliance that can explain why some addictions are highly resistant to therapy.

For these reasons, two incentives for dissociation may well be added to the conventional one of avoiding intrusive thoughts: The protection of desirable personal rules and the evasion of the personal rules that serve sellout-range interests. Incidentally, fantasy proneness may be added to traumatic memories as a source of such thoughts.

Neither addiction nor dissociation represent the straightforward functioning of an elementary principle; rather, hyperbolic discount curves can be expected to create them semistably in the course of intertemporal bargaining. Significant clinical addictions will usually occupy lapse districts (areas that the will has stabilized by its failures) and these addictions may ally with other addiction-range interests and even long-range interests. Likewise, dissociation has other motives besides the management of traumatic memories.

As theories of this complex subject proliferate, a unifying discipline will be indispensable: that of specifying how each tenet explains robust contradictory behaviors within the constraint of maximizing expected utility. This discipline seems attainable if and only if we understand the discount curves from expected rewards to be hyperbolic.

Acknowledgments

A preliminary version of this paper was presented to The Norwegian Research Council Working Group on Addiction, Collioure, France, September 28, 1994.

Coatesville Department of Veterans Affairs Medical Center and Temple University Medical School. I thank Barbara Gault, David Renfrow, and Elizabeth Ainslie for extensive help in discussing drafts of this article.

NOTES

1. People's preference for an addictive substance is sometimes not temporary, especially early in their consuming careers. An observer with superior foresight might say that it would be temporary, *if only they knew* . . .

If a person does *not* wish to quit under the prevailing circumstances, he would still be considered addicted if he found it very difficult to quit under circumstances that generated a wish to quit. This person is called a consonant addict [Skog, this volume].

However, addiction in this sense is an ordinary taste; such a taste would not puzzle us if it did not typically lead to a stage in which the person recognized his error but could not escape – in other words, to enslavement.

2. The broad substitutability of behaviors in a limited channel of expression must constrain diverse possible reward centers to be comparable along a single dimension for purposes of selection. That dimension is best called reward (Ainslie 1992, pp. 28–32; Shizgal and Conover 1996).
3. When "particular actions are united under a common rule, they are viewed as members of a class of actions subserving one comprehensive end. In this way the will attains a measure of unity" (Sully 1884, p. 631).
4. It turns out that pigeons do *not* discover the power of the test case; that is, even when they have learned some resistance to temptation, they do not come to perceive their own choices as precedents (Ainslie 1989).
5. The movements of a bellwether stock affect the market proportionately as investors are sensitive to test cases, that is, as they behave like "portfolio insurers" rather than "value-based investors" (Ferguson 1989).
Similarly, the sensitivity of an individual's self-control to the outcome of test cases may vary. A person who is not greatly worried about overeating or smoking, but is "just trying to cut down," may have frequent fluctuations into greater consumption without noticeably damaging his willpower. At the other end of the continuum, someone who is strongly motivated to cut out these behaviors and is threatened by any sign that they are returning may subject himself to such scrutiny that every relevant choice becomes a crucial test. He will stake such importance on a single lapse that, if it ever occurs, it will lead him to helpless surrender. The "first-drink" phenomenon has been such a

conspicuous cause of alcoholics' loss of sobriety that it was once thought to be a physical reflex (Evenson et al. 1973). This will be discussed in the section on how personal rules magnify lapses.

6. The empathic disengagement that characterizes a cosmopolitan society makes the most obvious alternative form of impulse control – tolerating one's vulnerability to social pressure from the community – unreliable and sometimes treacherous. The bigger the community in functional terms, the more often one encounters disengaged predators who can exploit her openness to engagement, a situation that encourages her to cultivate disengagement in turn. This disengagement is concealed by the friendly informality that is actually one of its symptoms, a sign of a "growing aversion to emotional intensity, [an aversion] that such informality requires" (Stearns 1994, p. 11). The change is well expressed by the twentieth-century cultivation of emotional "cool." Beginnings may be visible in Beau Brummel's introduction of the "stiff upper lip" to fashionable Regency society, which was at least a conspicuous milestone of our escalating competition in imperviousness to influence (Cole 1977). However, there may be important gender differences in the progress of this trend (Gilligan 1982).
7. Lawyers and other negotiators developed this useful term, meaning a boundary that stands out in some unique way from other boundaries that might be proposed. Bright lines have the effect of stabilizing negotiated settlements against shifts in boundaries from small changes in the balance of bargaining power. See Ainslie (1992, pp. 163–73).
8. This description of symptom formation as a product of intertemporal bargaining is not meant to negate the importance of biological preparedness for temptations in motivating lapses. The tastes for alcohol or thrills that make some risk-taking styles highly heritable are increasingly being traced to innate differences in the physiology of emotion (Cadoret et al. 1986, Grove et al. 1990), as are shorter range urges such as fearfulness and pain sensitivity (Kendler, Neale, and Kessler 1992, Skre et al. 1993, Panocka et al. 1986). Great urges may overcome even well-drawn rules. However, even a middling urge may overcome badly functioning rules. Rules that are overly strict or just poorly drawn may permit too little reward to sustain them against ordinary challenges. Thus bulimic patients may just be poorer lawyers than the anorectics who succeed in never bingeing, in that they cannot formulate sustainable rules for this purpose; and anorectics in turn feel a need to resist longings that people with more adaptive rules can gratify without fear of losing control.
9. As Erdelyi (1990) has pointed out, the unconscious but goal-directed effort to forget that the psychoanalysts call repression does not differ in nature from the conscious kind (suppression). I would suggest that its unconsciousness is shaped by the incentive to avoid losing the stakes of personal rules.
10. The person might be said to have "sold out" his subtler or richer interests for the reliability of a system.

Two ranges of temporary preference shorter than addictions – itches and pains – can also be discerned in common experience (Ainslie 1992, pp. 98–114), but they are not important for this discussion.

11. I use overeating and dieting as illustrations frequently, not because they seem to be pivotal, but because they are probably the best-known concerns about temporary preference in common experience.
12. Another factor – how sharply this series stands out against alternative series that could be substituted – was discussed above with bright lines.
13. Admittedly, melioration procedures with humans study subjects' awareness of rewards rather than their self-control; there should be little urge to favor small, early cash rewards when the cash can't be spent until the end of the experiment anyway. However, the advantage that countable rewards have over subtler ones in this purely informational problem should also be present when self-control is involved.
14. However frustrating this phenomenon is to the hedonist, it is liberating for motivational theory because it suggests how "process rewards," those that do not depend on a hardwired releasing stimulus, can compete on equal terms with biological need-satisfiers. A person is physically able to summon the most powerful of emotions – joy, terror, love, hate, etc. – without external stimuli, but insofar as he does so without occasions that are both adequately rare and adequately unpredictable, the absence of suspense will turn them into mere daydreams; such direct approaches will generally extinguish.
This phenomenon selects for strategies that use optimally unpredictable events to cue their process rewards (e.g., joy) and demands continual growth or replacement of these strategies as their finding of occasions becomes so efficient that these occasions become common or predictable (Ainslie 1992, Ch. 8). This model frees the observer from a search for "curiosity drives" and "conditioned emotional reactions" and brings all motives into a single marketplace.
15. Freud's "unconscious" organized centuries of puzzled observations (Ellenberger 1970, Whyte 1962), but the "repression" that he said created it confounds the impulsive diversion of attention from unpleasant truths with the impulse-controlling diversion of attention from temptations; it can be reconciled with strict utilitarian accounting, but only by the additional assumption of hyperbolic discounting (Ainslie 1992, pp. 35–9, 133–5).
16. Certainly processes with very short latency periods from suggestion to adoption – obsessions, hypochondriacal worries, vivid fantasies, traumatic memories, and other mental processes – are poorly controlled by personal rules, probably because there is not time enough between the urge and the process itself to test the urge as to whether it is permissible (Ainslie 1986). Diversion of attention from the pathway that brings up the urge should indeed work better than personal rules against very-short-latency processes.
17. Both groups also develop related self-destructive activities with a range of preference duration like that of addictions: Many traumatized patients voluntarily repeat their traumata (van der Kolk 1989), and many high fantasiers become addicted to their fantasies at the expense of richer human relationships (Firestone 1993).
18. Of course, such evasion could also be said to be preserving these rules, in that it prevents their outright confrontation. The interest that supports the rules may also support their evasion.

19. Torture of animals, for instance, is uncommon today compared with even recent times. If we go far enough back, we encounter village games where children nailed a cat to a wall and took turns trying to butt it to death with their heads (Origo 1959, pp. 40-3).

BIBLIOGRAPHY

- Ainslie, G. (1974), "Impulse control in pigeons," *Journal of the Experimental Analysis of Behavior* 21, 485-9.
- (1975), "Specious reward: A behavioral theory of impulsiveness and impulse control," *Psychological Bulletin* 82, 463-96.
- (1986), "Manipulation: History and theory," in M. P. Nichols and T. J. Paolino, Jr. (eds.), *Basic Techniques of Psychodynamic Psychotherapy*, New York, Gardner Press.
- (1989), "Internal self-control in pigeons," unpublished manuscript.
- (1991), "Derivation of rational economic behavior from hyperbolic discount curves," *American Economic Review* 81, 334-40.
- (1992), *Picoeconomics: The Strategic Interaction of Successive Motivational States Within the Person*, Cambridge, UK, Cambridge University Press.
- (1995), "A utility-maximizing mechanism for vicarious reward: Comments on Julian Simon's 'Interpersonal allocation continuous with intertemporal allocation,'" *Rationality and Society* 7, 393-403.
- (in preparation), *Empathy: A Picoeconomic Understanding of the Basic Good*.
- (in press), "A utility-maximizing mechanism for vicarious reward," *Rationality and Society*.
- Ainslie, G. and V. Haendel (1983), "The motives of the will," in E. Gottheil, K. Druley, T. Skodola, and H. Waxman (eds.), *Etiology Aspects of Alcohol and Drug Abuse*, Springfield, IL, Charles C Thomas.
- Ainslie, G. and R. Herrnstein (1981), "Preference reversal and delayed reinforcement," *Animal Learning and Behavior* 9, 476-82.
- Aries, Philippe (1965), *Centuries of Childhood: A Social History of Family Life*, New York, Random House.
- Barnes, Jonathan (ed.) (1984), *The Complete Works of Aristotle*, Princeton, NJ, Princeton University Press.
- Becker, G. and K. Murphy (1988), "A theory of rational addiction," *Journal of Political Economy* 96, 675-700.
- Bogen, J. and J. Moravcsik (1982), "Aristotle's forbidden sweets," *History and Philosophy* 20, 111-27.
- Branscomb, Louisa P. (1991), "Dissociation in combat-related post-traumatic stress disorders," *Dissociation* 4, 13-20.
- Bratman, Michael E. (1987), *Intention, Plans, and Practical Reason*, Cambridge, MA, Harvard University Press.
- Burnett, J. (1969), *A History of the Cost of Living*, Baltimore, MD, Penguin Books.
- Cadoret, R. J., E. Troughton, T. W. O'Gorman, et al. (1986), "An adoption study of genetic and environmental factors in drug abuse," *Archives of General Psychiatry* 43, 1131-6.
- Cole, Hubert (1977), *Beau Brummel*, Woodstock, NY, Beekman.

- Crawford, Helen J. and J. H. Gruzelier (1992), "A midstream view of the neuropsychology of hypnosis: Recent research and future directions," in E. Fromm and M. Nash (eds.), *Contemporary Perspectives in Hypnosis Research*, New York, Guilford, pp. 227-66.
- Crawford, Helen J., Audrey M. Brown, and Charles E. Moon (1993), "Sustained attentional and disattentional abilities: Differences between low and highly hypnotizable persons," *Journal of Abnormal Psychology* 102, 34-43.
- Dunn, Gary E., Anthony M. Paolo, Joseph J. Ryan, and Jay Van Fleet (1993), "Dissociative symptoms in a substance abuse population," *American Journal of Psychiatry* 150, 1043-7.
- Ellenberger, Henri F. (1970), *The Discovery of the Unconscious*, New York, Basic Books.
- Elster, J. (1989), *Nuts and Bolts for the Social Sciences*, Cambridge, UK, Cambridge University Press.
- Erdelyi, Matthew Hugh (1990), "Repression, reconstruction and defense: History and integration of the psychoanalytic and experimental frameworks," in Jerome L. Singer (ed.), *Repression and Dissociation: Implications for Personality Theory, Psychopathology, and Health*, Chicago, University of Chicago Press.
- Evenson, R., H. Altman, J. Sletton, and R. Knowles (1973), "Factors in the description and grouping of alcoholics," *American Journal of Psychiatry* 130, 49-54.
- Ferguson, Robert (1989), "On crashes," *Financial Analysts Journal*, March-April.
- Firestone, Robert W. (1993), "The psychodynamics of fantasy, addiction, and addictive attachments," *American Journal of Psychoanalysis* 53, 335-52.
- Fullilove, Mindy T., Robert E. Fullilove, Michael Smith, Karen Winkler, et al. (1993), "Violence, trauma, and post-traumatic stress disorder among women drug users," *Journal of Traumatic Stress* 6, 533-43.
- Gardner, Eliot L. (1992), "Brain reward mechanisms," in J. H. Lowinson, P. Ruiz, R. B. Millman, and J. G. Langrod (eds.), *Substance Abuse: A Comprehensive Textbook*, 2d ed., Baltimore, Williams & Wilkins, pp. 70-99.
- Gilligan, Carol (1982), *In a Different Voice: Psychological Theory and Women's Development*, Cambridge, MA, Harvard University Press.
- Green, Leonard, Astrid Fry, and Joel Myerson (1994), "Discounting of delayed rewards: A life-span comparison," *Psychonomic Science* 5, 33-6.
- Grove, W. M., E. D. Eckert, L. Heston, et al. (1990), "Heritability of substance abuse and antisocial behavior: A study of monozygotic twins reared apart," *Biological Psychiatry* 27, 1193-304.
- Herman, Judith L. and Emily Schatzow (1987), "Recovery and verification of memories of childhood sexual trauma," *Psychoanalytic Psychology* 4, 1-14.
- Herrnstein, R. (1961), "Relative and absolute strengths of response as a function of frequency of reinforcement," *Journal of Experimental Analysis of Behavior* 4, 267-72.
- Herrnstein, Richard J. and Drazen Prelec (1992a), "Melioration," in G. Loewenstein and J. Elster (eds.), *Choice Over Time*, New York, Russell Sage Foundation, pp. 235-64.
- (1992b), "A theory of addiction," in George Loewenstein and Jon Elster (eds.), *Choice Over Time*, New York, Russell Sage Foundation, pp. 331-60.

- Herrnstein, R. J., G. Loewenstein, D. Prelec, and W. Vaughan, Jr. (1993), "Utility maximization and melioration: Internalities in individual choice," *Journal of Behavioral Decision Making* 6, 149-85.
- Heyman, Gene M. (1996), "Resolving the contradictions of addiction," *Behavioral and Brain Sciences* 19, 561-610.
- Hilgard, Ernest R. (1965), *Hypnotic Susceptibility*, San Diego, CA, Harcourt Brace Jovanovich.
- (1977), *Divided Consciousness, Multiple Controls, and Human Thought and Action*, New York, Wiley.
- James, W. (1890), *Principles of Psychology*, New York, Holt.
- Kant, I. (1960), *Religion Within the Limits of Reason Alone* (T. Green and H. Hucken, transl.), New York, Harper and Row, pp. 15-49.
- Kendler, Kenneth S., Michael Neale, and Ronald C. Kessler (1992), "The genetic epidemiology of phobias in women: The interrelationship of agoraphobia, social phobia, situational phobia, and simple phobia," *Archives of General Psychiatry* 49, 273-81.
- Kirby, Kris N. and Richard J. Herrnstein (1995), "Preference reversals due to myopic discounting of delayed reward," *Psychological Science* 6, 83-9.
- Kobasa, S. C. and S. R. Maddi (1983), "Existential personality theory," in R. J. Corsini, and A. J. Marsella (eds.), *Personality Theories, Research and Assessment*, Itasca, IL, Peacock.
- Kohlberg, L. (1963), "The development of children's orientations toward a moral order: I. Sequence in the development of moral thought," *Vita Humana* 6, 11-33.
- Labelle, Louise, Jean-Roch Laurence, Robert Nadon, and Campbell Perry (1990), "Hypnotizability, preference for an imagic cognitive style, and memory creation in hypnosis," *Journal of Abnormal Psychology* 99, 222-8.
- Leibenstein, H. (1976), *Beyond Economic Man: A New Foundation for Microeconomics*, Cambridge, MA, Harvard University Press.
- Lynn, S. J. and J. W. Rhue (1988), "Fantasy proneness: Hypnosis, developmental antecedents, and psychopathology," *American Psychologist* 43, 35-44.
- Macdonald, John and John Piggott (1993), *Global Quality: The New Management Culture*, San Diego, CA, Pfeiffer.
- Malekzadeh, Ali, R. and Afsaneh Nahavandi (1987), "Merger mania: Who wins? Who loses?," *Journal of Business Strategy* 8, 76-9.
- Mazur, J. E. and A. W. Logue (1978), "Choice in a self-control paradigm: Effects of a fading procedure," *Journal of Experimental Analysis of Behavior* 30, 11-17.
- McCallum, Kim E., James Lock, Mary Kulla, Marcia Rorty, and Richard D. Wetzel (1992), "Dissociative symptoms and disorders in patients with eating disorders," *Dissociation* 5, 227-35.
- Mischel, H. N. and W. Mischel (1983), "The development of children's knowledge of self-control strategies," *Child Development* 54, 603-19.
- Nemiah, J. C. (1977), "Alexithymia: Theoretical considerations," *Psychotherapy and Psychosomatics* 28, 199-206.
- Origo, Iris (1959), *The Merchant of Prato: Francesco di Marco Datini*, New York, Octagon.
- Panocka, Izabella, Przemyslaw Marek, and Bogdan Sadowski (1986), "Differentiation of neurochemical basis of stress-induced analgesia in mice by selective breeding," *Brain Research* 397, 156-60.

- Piaget, J. (1965), *The Moral Judgment of the Child*, New York, Free Press.
- Polivy, J. and C. P. Herman (1985), "Dieting and bingeing: A causal analysis," *American Psychologist* 40, 193-201.
- Putnam, F. W. (1989), *Diagnosis and Treatment of Multiple Personality Disorder*, New York, Guilford.
- Rachlin, Howard (1995), "Self-control: Beyond commitment," *Behavioral and Brain Sciences* 18, 109-59.
- Resnick, Heidi S., Dean G. Kilpatrick, and Bonnie S. Dansky (1993), "Prevalence of civilian trauma and posttraumatic stress disorder in a representative national sample of women," *Journal of Consulting and Clinical Psychology* 61, 984-91.
- Rhue, Judith W. and Steven Jay Lynn (1987), "Fantasy proneness: The ability to hallucinate as real as real," *British Journal of Experimental and Clinical Hypnosis* 4, 173-80.
- (1989), "Fantasy proneness, hypnotizability, and absorption - A re-examination," *International Journal of Clinical and Experimental Hypnosis* 37, 100-6.
- Roche, Suzanne M. and Kevin M. McConkey (1990), "Absorption: Nature, assessment, and correlates," *Journal of Personality and Social Psychology* 59, 91-101.
- Ross, Colin A., S. Joshi, and R. Currie (1990), "Dissociative experiences in the general population," *American Journal of Psychiatry* 147, 1547-52.
- Ross, Colin A., Jeff Kronson, Stuart Koensgen, Ken Barkman, et al. (1992), "Dissociative comorbidity in 100 chemically dependent patients," *Hospital and Community Psychiatry* 43, 840-2.
- Schacter, D. L. and J. F. Kihlstrom (1989), "Functional amnesia," in F. Boller and J. Grafman (eds.), *Handbook of Neuropsychology*, vol. 3, pp. 209-31. Amsterdam, Elsevier.
- Schelling, T. C. (1960), *The Strategy of Conflict*, Cambridge, MA, Harvard University Press.
- Shizgal, Peter and Conover, Kent (1996), "On the neural computation of utility," *Current Directions in Psychological Science* 5, 37-43.
- Sjoberg, L. and T. Johnson (1978), "Trying to give up smoking: A study of volitional breakdowns," *Addictive Behaviors* 3, 149-67.
- Skre, Ingunn, S. Onstad, S. Torgersen, S. Lygren, et al. (1993), "A twin study of DSM-III-R anxiety disorders," *Acta Psychiatrica Scandinavica* 88, 85-92.
- Smyser, Cameron H. and David A. Baron (1993), "Hypnotizability, absorption, and subscales of the dissociative experiences scale in a nonclinical population," *Dissociation* 6, 42-6.
- Spanos, P. Nicholas, Evelyn, Gabora, Menary, J. Natalie, C. DuBeuil, Susan, et al. (1991), "Secondary identity, enactments during hypnotic past-life regression: A sociocognitive perspective," *Journal of Personality and Social Psychology* 61, 308-20.
- Spiegel, David (1991), "Dissociation and trauma," in A. Tasman and S. Goldfinger (eds.), *American Psychiatric Press Review of Psychiatry*, vol. 10, pp. 261-75, Washington, DC, American Psychiatric Press.
- Stearns, Peter N. (1994), *American Cool: Constructing a Twentieth-Century Emotional Style*, New York, New York University.
- Stevenson, M. K. (1986), "A discounting model for decisions with delayed positive or negative outcomes," *Journal of Experimental Psychology: General* 115, 131-54.

- Stigler, George (1980), "The economist as preacher," Tanner Lectures delivered at Harvard Univ., April 1980.
- Sully, J. (1884), *Outlines of Psychology*, New York, Appleton.
- van der Kolk, Bessel A. (1989), "The compulsion to repeat the trauma: Re-enactment, revictimization, and masochism," *Psychiatric Clinics of North America* 12, 389-411.
- Whyte, Lancelot Law (1962), *The Unconscious Before Freud*, Garden City, NY, Doubleday.
- Wilson S.C. and T. X. Barber (1983), "The fantasy-prone personality: Implications for understanding imagery, hypnosis, and parapsychological phenomena," in A. A. Sheikh (ed.), *Imagery: Current Theory, Research, and Application*, New York, Wiley.
- Zerbe, Kathryn J. (1993), "Selves that starve and suffocate: The continuum of eating disorders and dissociative phenomena," *Bulletin of the Menninger Clinic* 57, 319-27.