

A Theory of Addiction

RICHARD J. HERRNSTEIN AND DRAZEN PRELEC

We would often be sorry if our wishes were gratified.

Aesop

No man ever became extremely wicked all at once.

Juvenal

THE COMPLEXITY of addiction is mirrored in the many disciplines that study it. The chemistry of addictive substances falls in the domain of *biochemistry*, tolerance and withdrawal belong to *physiology*; various personality or hereditary predispositions, and the role of stressful events, are jointly addressed by *psychology* and *human genetics*; the relation to the poverty, community structure, and the "social matrix" are problems of *sociology* and *political science*.

Alongside these various approaches, however, there must also be a theory of addiction that reconciles the ostensibly self-destructive consequences of addiction with the central *economic* assumption that human action can be understood as the rational pursuit of self-interest, or, if reconciliation is not possible, to examine what the implications are for that central economic assumption. The economic aspect of addiction provides the focus of this essay.

Addiction resolves into two separate, but related, paradoxes for any theory that assumes behavior to be generally utility maximizing. First, addiction is perceived to be harmful to the person who consumes the substances (above and beyond its effects on family members and society). The addict's revealed preferences are inconsistent

with society's view that addiction is a losing proposition for the addict. Second, many addicts claim that they wish to change their behavior but are unable to do so. Their stated preferences are inconsistent with the preferences that they reveal through behavior. What shall we conclude about the relation between revealed preferences and utility in light of addiction?

Four Interpretations of Addiction

Economic theory can deal with addictive behavior in four distinct ways, as follows:

(1) *Addiction as disease, not choice*: A drug addict may be viewed as having lost the power to choose whether or not to indulge his habit, in which case the addict's behavior would not need to be accounted for by *any* theory of choice, including the economic theory. The historical shift from addictions as vices to addictions as diseases was a shift to this first approach. Similarly, criminal law has formalized a principle of "no choice" in some of its tests of culpability. The *irresistible impulse* rule, for example, allows someone to be acquitted of responsibility for an act if it can be proved (in the legal sense) that, because of overpowering emotion, the perpetrator lacked the power of choice when he committed it. Criminal law also excuses people when they act without conscious intent, or *mens rea*, on the grounds that choice requires consciousness.

Physiology has also been invoked as a reason for classifying certain behaviors as outside the domain of choice. Murderers have been defended on the grounds that their decision-making ability had been destroyed by eating too much refined sugar (the "Twinkie" defense) or by a brain tumor. The premenstrual syndrome has been offered as a defense for some crimes. Everyday theorizing about unacceptable or unconventional behavior often excuses it by calling it "physiological." The no-choice approach says, in effect, that the behavior in question was not controlled by its potential consequences.

Such arguments reflect a tendency to subtract the volitional component from choice in proportion to our knowledge of the physical reasons for the behavior. Obesity, for example, might become classified as an involuntary condition, upon discovery of its physiological correlates or genetic predispositions.

Whatever the merits of these considerations for determining personal responsibility, the "no-choice" approach does not clearly identify addictive behavior. First, behavior always has physiological rea-

sons, whether or not we know what they are. Some relatively harmless rewards—for example, sweets—have a relatively well-understood physiology, while the rewards provided by some addictive behaviors, such as gambling, are obscure.¹ According to the no-choice approach, the former behavior would not involve choice, but the latter would, at least until we discover the physiological basis for gambling. Second, the drives for addictive behavior are not always intense, as they ought to be if the irresistible impulse criterion is to apply. Experiments have shown, for example, that obese people are more easily deterred from eating than people of normal weight by the presence of minor physical obstacles to the food. Schachter (1971), for example, describes experiments in which overweight people are more deterred from eating nuts by having to shell them or from eating sandwiches by having to get them from a refrigerator than nonoverweight people are.

Finally, addictive behavior is not distinguished by the absence of conscious deliberation. Acquiring illegal drugs, or purchasing a package tour for gambling at Las Vegas, requires more planning than many "normal" consumer choices, like watching television or hailing a taxi.

(2) *Addiction as rational self-medication*: The opposite view of addiction sees it as part of a rational lifestyle, which only appears unusual and self-destructive because we do not understand its environmental and constitutional context. A sophisticated example of this approach is provided by Becker and Murphy (1988), who treat addictive behavior as part of an optimal intertemporal consumption plan. Behavior is, in effect, *perfectly* controlled by its consequences. Addicts take full account of the impact of their current behavior on the future, including their future taste for the addictive substance, according to this theory. Although addicts may be unhappy by normal standards, they would be "even more unhappy if they were prevented from consuming the addictive goods" (Becker and Murphy, 1988, p. 691).

In all but name, Becker and Murphy depict addiction as a form of medication: The addictive commodity or activity is an expensive, inconvenient, *but nevertheless rational* treatment for special psychological conditions, such as depression, stress, and low self-esteem. This outlook seems, at least on the surface, to be inconsistent with addicts'

¹ Although recent studies (Goleman, 1989; Roy, Adinoff, Roehrich, Lamparski, Custer, Lorenz, Barbaccia, Guidotti, Costa, and Linnoila, 1988; Roy, De Long, and Linnoila, 1989; Shaffer, Stein, Gambino, and Cummings, 1989) suggest that chronic gamblers may have low levels of activity in the noradrenergic system, abnormalities in cerebrospinal fluids, and high levels of extroversion in their personality profiles.

trying to free themselves of their habit, as well as with the relative excess of young, rather than old, addicts despite the higher long-term risks to the former. But, in any event, if this view is correct, then policies that restrict access to drugs and criminalize their purchase are misguided and cruel, because they penalize people whose personal welfare is already extremely low.

(3) *The primrose path*: Addictive behavior is sometimes viewed as a trap into which one is lured, because the latent costs of addiction are initially hidden or because the underlying behavioral process is deficient at making rational choices. This is an approach that splits the difference between the first two: It holds that addiction truly does depend on a person's choices, but that those choices can sometimes fail the test of rationality. Behavior is controlled by its consequences, according to this approach, but the result may be far from perfectly adaptive. Theories that attempt to make precise this view of addiction may be labeled *primrose path* theories. The key observation is that the typical addict goes down the primrose path believing that there is little danger of losing control (Goldstein and Kalant, 1990). The danger arises because the availability of certain substances or activities creates a situation in which a person's normal behavioral rules are inadequate. The semblance of rationality that our normal, that is, nonaddictive, behavior exhibits is then the consequence not of a utility maximization process, but rather of a good match between the behavioral rules and particular circumstances. When this match is not good—as with addictive goods—then the same behavioral rules produce poor results. We are, for example, following the same fundamental rules when we develop an appetite for golf as when the developing appetite is for heroin, according to this theory. The theory developed in the second section is a primrose path theory, but there could be others.

(4) *The divided self*: The final approach starts with the observation that the same person has different preferences at different times. In the morning, for example, a person may know that he does not want to eat dessert after supper and would order appropriately if the person could bind himself to do so; at supper, he or she succumbs. Or the person may awake daily, filled with resolve not to smoke, drink, dawdle at the water cooler, snarl at the children or spouse, etc., yet fail at virtually the first chance. It is not a matter of faulty knowledge, for the scene may be reenacted daily for years.

The self seems to be divided whenever preference depends on vantage point, which is often the case in addiction. The person discovers that too much eating or drinking has been a primrose path, leading ultimately to trouble, but this discovery fails to protect the

person from succumbing when the "undesirable" alternative is at hand. Consequently, persons are in the paradoxical situation of knowing how to act in their own interest in the absence of occasions to do so, but failing to do so when the occasions arise. The person seems to know just how unwise it is to drink except when he passes a saloon.

In an article about the transmission of AIDS by unsafe drug use or sexual practices, *Science* magazine quotes Marshall Becker of the University of Michigan School of Public Health on the difficulty of changing this highly dangerous behavior: "We're asking people to make these crucial decisions over and over again at the exact moment when they're most vulnerable, which is to say when they're about to have sex or right when they're about to stick a needle in their arm" (Booth, 1988, p. 1237).

How individuals and societies do or do not resolve the paradox has inspired much of the recent work on the subject of self-control (Ainslie, 1986; Elster, 1986; Schelling, 1980; Thaler and Shefrin, 1981; Winston, 1980). The common element in formalized divided self models is that the individual is viewed as a collection of rational subagents, jockeying for control of behavior. Although individually the subagents are utility maximizers, the behavior of the collective can be severely suboptimal, as numerous examples from game theory demonstrate (i.e., the Prisoner's dilemma). The third section will discuss the divided self approach to addiction, and will also show its natural links to the primrose path theory presented in the second section.

The remainder of the essay attempts to characterize addiction itself, to show its relation to primrose path and divided self theories, and, in the fourth section, to draw some implications of this analysis for public policy.

A Behavioral Model of Addiction

Four Diagnostics for Addictive Behavior

As Becker and Murphy (1988) point out, the range of activities that can at one or another time be considered addictive is extremely broad: "People get addicted not only to alcohol, cocaine, and cigarettes but also to work, eating, music, television, their standard of living, other people, religion, and many other activities" (p. 675f).

Drugs, eating, television, music, a standard of living, human relationships: What do these diverse activities have in common? Some

of the activities on this list would not usually be called addictions in ordinary speech, as opposed to habits, perhaps, or acquired tastes. Habits, acquired tastes, and addictions share the characteristic that they refer to activities that become more likely with repeated choice (Becker and Murphy, 1988; Leonard, 1989). But an activity is usually called an addiction only if this change seems to be a trap of some kind, locking the person into a behavior that he would avoid if he could only view it "objectively," as may well be the case with, for example, certain personal relationships and watching television.

We agree with Becker and Murphy that addictions include many activities beyond substance abuse: gambling, spending beyond one's means, compulsive buying of particular items, such as, for example, cosmetics,² some deviant sexual acts like exhibitionism, probably also the "type A" (hypercompetitive) personal style, excessive temper, a tendency to form self-destructive love relationships, and many others. We could perhaps include computer hacking as a new form of addiction. If it is taken as axiomatic that *all* behavior is physiologically grounded, then the everyday characterization of addictions as being "physiological" no longer makes sense. What, then, are addiction's defining features?

Although we propose here four criteria for addictions, we do not argue for a strict logical connection between the conditions listed and the use of "addiction" in ordinary language. The conditions are meant to capture most, if not all, of the denotations of the word.

(1) Addiction is normally not produced by a single action, but is rather the result of a long stream of choices.

Here, we merely note the obvious, namely, that addiction is a *habit*. A rash decision to enlist in the Foreign Legion does not constitute an addiction, no matter how long and unsatisfying the subsequent lifestyle. Addictions, in other words, are ordinarily built up from many smaller decisions; in this respect, they are closely related to what we recognize as defects in character or personality.³ A lazy

²From the *New York Times Magazine* (Wells, 1988): "next to the club soda, Perrier and cat food in the refrigerator are 150 tubes of lipstick. There's hardly room for such clutter in the bathroom. That is crowded with 100 or so bottles of fragrance and uncounted cases of eyeshadow, blusher and other necessities. This shrine to makeup is Margot Rogoff's apartment. . . . This is the land of the cosmetics addict."

³One could, in principle, become addicted from a single encounter with the addictive commodity—a super-addictive drug, for example—but even here, as should become clear, many of the characteristics of addiction, especially its harmful effects, will only materialize as the addictive behavior occurs repeatedly. An issue not dealt

person, for example, is not one who *once* chose to rest rather than work, but one who is consistently predisposed to choose rest over work. We could say that he is addicted to loafing. Likewise, a hygienic person is not someone who once chose to bathe or to wear clean clothes, but someone who is predisposed to doing so repeatedly and consistently. We are likely to refer to cleanliness or idleness as addictions only if they meet at least some of the other criteria to be described, particularly the next one.

(2) Addictive behavior has significant negative intrapersonal side effects—costs that are caused by addiction but that appear in the context of other, ostensibly unrelated activities.

From the outside, it is usually clear that addictive behavior is having profound effects on the rest of the addict's life. But, for the addict, the pattern is for the psychic benefits to be directly associated with consumption of the addictive good, while the psychic costs are to be spread all around. Thus, a heavy gambler or heroin addict craves and welcomes gambling or heroin, but finds that work and family life are not as satisfying as they used to be. The negative effects of the addiction on other activities are perceived as deteriorating personal relationships or careers or the like, rather than as part and parcel of the addictive behavior—as much a part of it, say, as the pleasures of gambling or a drug "fix."

However, activities in which the primary side effect is often perceived as beneficial can also become addictive, although the addiction is then a compulsive overdoing of an otherwise worthwhile activity. With cleanliness, as with exercise or dieting, for example, the relevant benefits of greater fitness or more attractive appearance constitute a side effect, because they are absorbed as an increase in satisfaction derived from a wide range of other activities, at least initially. Beyond a certain rate of engaging in the activity, however, the side effects become negative, in that so much time is spent washing, working, or exercising that other valued activities are placed at risk. It is only at this point that the washing, exercising, or working is likely to be considered addictive.

(3) The benefits of each instance of an addictive behavior are generally more immediate than the costs.

with in this chapter is the relation between whether or not a given addiction is thought of as betraying a defect of character and how many episodes of indulgence it takes a typical person to become an addict (but see Prelec and Herrnstein, 1991).

Activities whose initial impact on the occasions when it occurs is negative (i.e., painful) are not usually regarded as addictive. If, for example, the pleasure-withdrawal cycle in opiate addiction is reversed, so that significant pain precedes pleasure, then opiates would most probably cease to be addictive. This time dependence, over rather short intervals of time, seems to play some role in many addictions. The exceptions seem mostly to be cases, such as exercise, dieting, and cleanliness, for which the natural tendency is underindulgence, perhaps because of the dominance of short-term costs over long-term benefits, but become addictive when they shift to overindulgence, at which point they, too, may again have short-term benefits and long-term costs, albeit over longer intervals than opiates and the like.

(4) Addictive behavior, if not at first then eventually, displays temporary preference—it is anticipated with apprehension, looked back on with regret, and engaged in nevertheless.

Economic approaches to addiction invariably focus on changes in tastes—the developing of appetites for, say, alcohol, tobacco, or cocaine. While the tastes of an addict differ significantly from the tastes of nonaddicted persons, and perhaps differ, too, from his own tastes prior to repeated encounters with the addictive substance, this can also be said about anyone who regularly indulges in commodities of which the values depend on how often they are consumed. The class of such consumers is so broad that it probably includes everyone. The tastes of skiers differ from those of the general population, as do the tastes of stamp collectors, vegetarians, philanthropists, or rodeo fans. Those who love pasta have different tastes from those who love rice. The critical question is why among all these varied “acquired tastes” that constitute the preferences of any person, only a special class would be labeled addictive.

For a behavior to be called an addiction, rather than just a personal bent or appetite, it must be *unwanted*. The person must want to stop but fail to do so, or at least an onlooker suspects that the addict lacks the ability to stop should he ever want to. Most present addicts would have refused the addictive lifestyle if it had been presented to them as a one-shot choice that locks in a specific consumption program, like a regimented vacation plan.

Although our four conditions are not hard to state informally, it is difficult to characterize them within a rational choice model. This, in itself, is a significant limitation on rational modelling. To start with, the distinction between a one-shot choice and a habit does not appear

in the normative theory, because a rational agent should be able to make a once and for all decision to choose a particular rate of consumption (even when the marketplace does not offer a "subscription" that locks him into that rate). Second, it should not matter to a rational individual whether the costs or benefits of an activity accrue while the activity is in progress, or whether they spread over other activities: The concept of *side effects* has no role as such in the rational model, inasmuch as the model in principle includes all effects of behavior. Third, in order to make time dependence relevant over the relatively brief cycles of pleasure and pain that occur in addiction, one would have to assume nonnegligible discount factors for intervals of days or hours or even minutes, which would then, by the logic of compound discounting, imply a complete insensitivity to consequences more than a few days or weeks away. Finally, as noted already, the idea that a person might dread engaging in certain behavior in advance, then engage in the behavior "voluntarily," only to regret having done so afterward, is on the face of it inconsistent with rationality.

Given the difficulty of expressing the four conditions within a rational model, we will turn to an alternative set of theoretical building blocks, which have been developed in the context of psychological research on human and animal choice behavior. The remainder of this section describes these concepts, and formulates a theory of addiction based on them.

Distributed Choice and Addiction

As we have just stated, standard economic theory does not draw a distinction between consumption variables that are decided with a single action (such as a car purchase), and variables that are the aggregate consequence of a series of individual, small-scale decisions, such as becoming obese or a habitual smoker. Our central hypothesis, derived from a growing body of experimental evidence, is that in this second class of situations, choice is guided by a particular sort of limited optimization, one that is fairly efficient in some situations and markedly inefficient (suboptimal) in others.

According to this idea, called the principle of *melioration* (Herrnstein, 1982; Herrnstein and Prelec, 1991; Herrnstein and Vaughan, 1980; Prelec, 1982), a person's behavior in situations of repeated choice eventually distributes itself over alternatives in the choice set so as to equalize the returns per unit invested, in time, effort, or some other constrained dimension of behavior. Behavior at the equilibrium

point, where returns per unit invested are equal, conforms to the *matching law*. We have described melioration and matching elsewhere in this book (see Chapter 10) and will not therefore repeat it except to apply it to the question of addictive behavior. For present purposes, it should suffice to identify some commodity that is at issue—alcohol, desserts, loafing, gambling, etc.—as one of two alternatives, without saying just which one it is. The other alternative we interpret as the collection of activities and commodities that are mutually exclusive and exhaustive with respect to the commodity at issue. Thus, for drinking, one value function is for drinking, the other is simply for not drinking. The two categories of consumption are thus complementary by definition. The case for viewing alcohol consumption, and, by implication, other addictive goods, as a choice of this sort has been forcefully made by Vuchinich and Tucker (1988).

In Figure 13.1, the x axis represents *distributed choice*, namely, the rate with which the behavior associated with the commodity or activity at issue is engaged in during some appropriate interval of time—say a week, a month, a year. If one assumes a fixed budget constraint, and as long as prices are constant, this rate is proportional to the

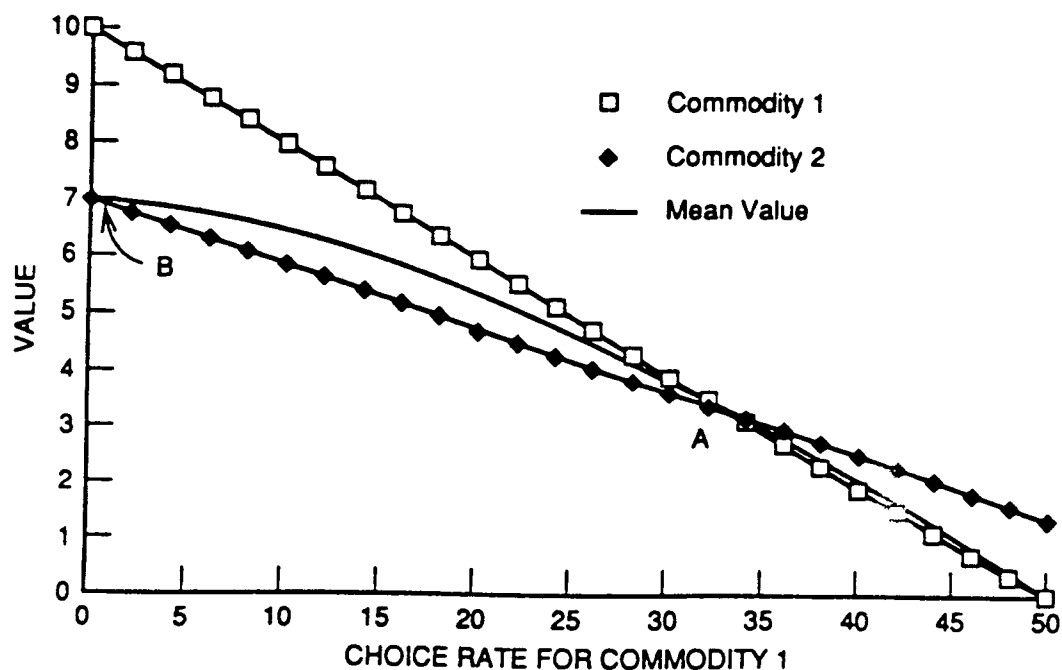


Figure 13.1. Value as a function of allocation to alternative 1, for alternative 1 (squares) and alternative 2 (diamonds). The weighted average value over both alternatives is given by the solid curve. The matching point is at A; optimal allocation is at B.

expenditure on the item. Expenditure on everything else is proportional to the distance to the right corner of the x axis. Changes in price have been shown to produce the expected changes in consumption rates, even for the most powerfully addictive substances, such as cocaine, alcohol, and tobacco (Goldstein and Kalant, 1990).

The heights of the value functions at each allocation show the average returns at that allocation, taking into account *all* the relevant parameters affecting current value, such as drive level, past history (as embodied either in pathological physiological condition or learned valuations), reinforcement-delay, price, risk, etc. The lines through the squares and diamonds represent the value of the commodity in question and its complement at any particular long-run division of choices between them (i.e., with a pair of long-run consumption frequencies). As drawn in Figure 13.1 the value functions indicate that the taste for either commodity is a linearly diminishing function of how often commodity 1 is chosen. The solid line is a weighted average of the two individual value functions, with the weight given to each function equal to the *relative* frequency of consumption of the corresponding commodity.

The process of melioration, without further elaboration, implies that choice is guided by the heights of the two value functions at any given allocation. The equilibrium pair of consumption rates is at the point where the two value functions intersect, which is also the point where the values obtained from both alternatives exactly match the mean value obtained from both (at A in Figure 13.1). This *matching* point does not necessarily coincide with the optimal pair of choice rates, which, in this example, would require total abstinence from commodity 1 (at B).

The configuration in Figure 13.1 is representative of many settings that people perceive as being problematic. The return to each commodity depends on how much of the total behavioral investment it receives. Relatively more of commodity 1 (squares) means less return to both commodities. In the context of choices between 1 and 2, average returns to 2 (diamonds) rise with its increased consumption; those to 1 decrease sharply with its consumption.

The crossing value functions in Figure 13.1, at A, define a stable matching point (Herrnstein and Prelec, 1991 and Herrnstein and Prelec, Chapter 10, this book), in that deviations around the point are self-canceling. Matching yields much less aggregate utility than the allocation at the optimum, that is, the maximum point, B, on the average curve, where commodity 2 would be chosen exclusively, even though commodity 1 is more desirable at this allocation.

We assume that a person does not "see" the overall picture, as it is plotted in Figure 13.1, because he is psychologically embedded in it, at a certain point along the abscissa. A person who currently favors alternative 1 (e.g., at the matching point in Figure 13.1) cannot quickly test the values that would be enjoyed at the optimal distribution—or at any other distribution, for that matter. If one is in the habit of loafing, one cannot casually and instantaneously sample the fruits of long and hard labor. The interactions between consumption and taste are here modeled as a "black box," gradually shifting the value of the commodities as patterns of allocation change, and driving distributed choice toward a matching point.

It may help to think of Figure 13.1 as a representation of how a person evaluates, in reinforcement or utility terms, the two mutually exclusive and exhaustive answers, namely, "yes" and "no," to the question of consuming commodity 1 at each allocation. The question is, in principle, continuously on the table during the time interval under observation.⁴

The rate of saying yes or no intersects various of the factors affecting value. There is the obvious interaction with drive state—the average value of a slice of chocolate cake varies inversely with the rate of consumption over some range of rates. For many commodities, external availability mimics the motivational interaction. The more one visits a given berry bush, the poorer and scarcer the berries one finds.

An important class of interactions between the utility of returns and consumption rate is mediated by delay of gratification. If we say yes to cake only when it is set before us, the pleasures of eating it are likely to be forthcoming sooner than if we say yes more often, including when we are far from food, thereby initiating a chain of activities with cake at its end. It is not that we cannot initiate at any time a chain of activities with cake at the end, but that the reinforcement for doing so may be long deferred at the moment of initiation. If deferred returns are discounted, then, other things being equal, high rates of saying yes produce lower average benefits, independent of any interaction with drive or availability.⁵

⁴Later we show that certain cultural practices may be construed as ways of taking the question off the table at certain times, and thereby constraining the rate of "yesses."

⁵An equivalent formulation of this point, more obviously linked to the principle of melioration, is that a given reinforcer spread over a longer time constitutes a lower rate of reinforcement, inasmuch as reinforcement rate is given by the ratio of the absolute value of the reinforcement to the time allocated to obtaining it (Herrnstein, 1982; Herrnstein and Vaughn, 1980).

How Addictions, and Addicts, Differ

Pictures like the one in Figure 13.1 may be fairly charged with hiding more than they disclose. Their advantage is that they provide a framework for taking into account the variables that affect distributed choice for meliorators, particularly for those commodities and activities that are generally agreed to be self-destructive. We can sample only sketchily from the vast clinical literature describing the phenomenology of addictive substances, such as smoking (tobacco), drinking (e.g., caffeine, alcohol), and the many varieties of chemical abuse (psychostimulants like cocaine or amphetamine, opiates like morphine or heroin; see Irwin, 1990, for a useful classification of addictive substances), let alone the even larger literature on all sorts of repetitive, harmful behavior. The following points reflect certain general patterns that are evident in this large literature.

In what follows, it may seem that we are simply postulating functions to fit the literature, but the point is that doing so is a natural extension of melioration, rather than a contrived use of it. This is because, unlike rational choice theory, the melioration framework makes the value functions for the competing behavioral alternatives decisive in controlling choice. In a maximization framework, a person's behavior is taken as a whole.

Individual Variation Individuals vary in their susceptibility for falling into self-destructively high or low rates of indulgence, which is to say they vary in their propensity toward a stable matching point at which the aggregate returns are significantly suboptimal. In terms of Figure 13.1, the placement and slopes of the two value functions vary from person to person. For alcohol, the indications point toward genetic dispositions toward excessive use. But individual differences need not be a matter of narrow, substance-specific susceptibilities, genetic or otherwise. Much of the clinical literature suggests that some people are launched on the (primrose) path toward addiction or other self-destructive behavior in the spirit of problem solving (e.g., Jellinek, 1960; Orford, 1985; Pattison, Sobell, and Sobell, 1977).

A person may use a psychostimulant to deal with the problem of depression (or, in some cases, obesity). Someone else may solve a shyness problem with alcohol. A youngster may find that he or she gains the admiration of peers by smoking tobacco or dope. A gambler may have been trying to solve a genuine financial problem when he fell into the bottomless pit of compulsive gambling. In short, the motivational structures relevant to addictions and other suboptimalities can extend far beyond the specific drives involved in the activities

at issue. Indeed, abusers of any given substance or activity are disproportionately likely to be abusing other substances or activities, or to be otherwise suffering from psychiatric illness (Lesieur, 1989).

Knife Edges The value functions for people at risk for abusing some substance or activity may have subsidiary features that Figure 13.1 omits. It is part of the lore of certain addictions, such as that of alcoholism, that something akin to a knife edge describes the risk of succumbing (Koob and Bloom, 1988). Low levels of indulgence can go on indefinitely, but, at some threshold, the pattern tips over, and consumption seems to break free of control. Figure 13.2 is a simple alteration of the value functions consistent with this story; drinking (or some other comparable good) is represented by commodity 1 (squares), not drinking by commodity 2 (diamonds). The weighted average value is given by the solid line.

Figure 13.2 postulates two stable equilibria, at A and C, and one unstable one, B, which locates the knife edge (disregarding the two additional unstable points at exclusive choice of either alternative). The matching point at A would likely be encountered first, because

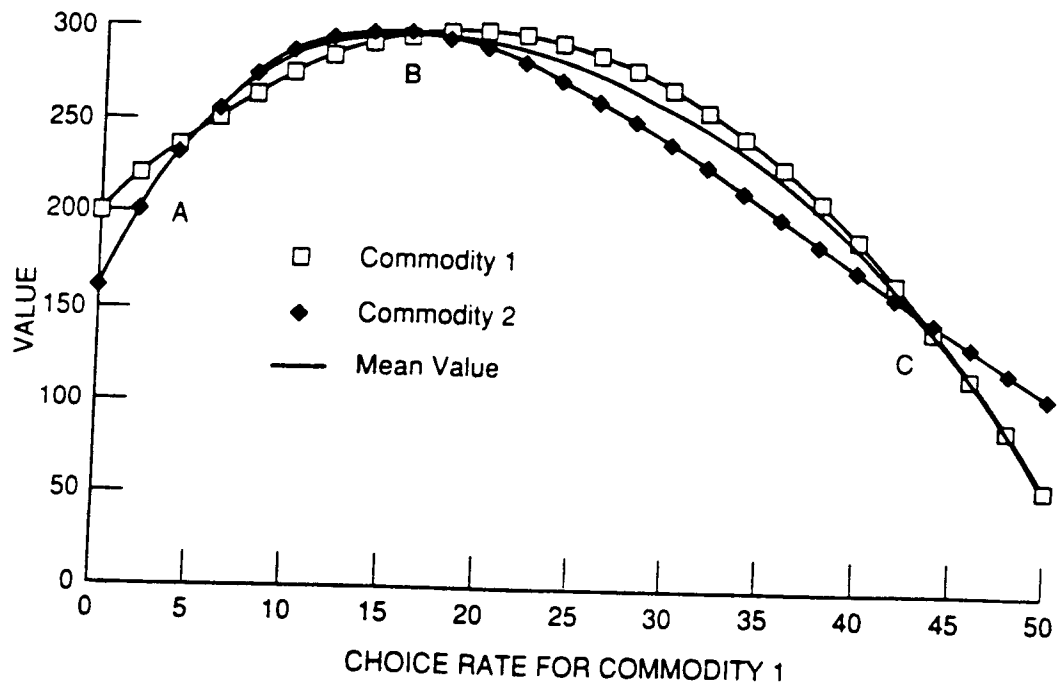


Figure 13.2. Value as a function of allocation to alternative 1, for alternative 1 (squares) and alternative 2 (diamonds). A and C are stable matching points; B is the "knife edge." The solid curve, the weighted average, virtually overlaps with alternative 2 to the left of A and with alternative 1 to the right of C.

it requires only a low rate of indulgence. In this region of allocation, the person is still developing a taste for drink, or still honing whatever skills may be involved in its successful use, hence, the positive slope of the value function at this level. At the same time, using alcohol may still be solving some problem in the person's life, enhancing enjoyment of activities in addition to drink itself, hence, the positive slope of the value function for alternative 2. Thus, at and around A, the average returns from drinking and nondrinking are rising with increased use of alcohol. At this matching point, drinking would seem to be under good control.

However, should something change, either in the environment or in the person's drives, so that an ongoing pattern of allocation now falls to the right of the knife edge (i.e., B in Figure 13.2), another dynamic process takes hold. Between B and C, melioration drives allocation toward dangerously high alcohol consumption, consistent with the impression of loss of control. Increased drinking begins to cast a hedonic shadow on nondrinking as the value function for commodity 2 declines. Alcohol begins to interfere with the pleasures to be had from other activities. The person may begin to experience the "anhedonia" or "dysphoria" referred to in the literature of alcoholism. Concurrently, the value function for drinking again exceeds that for not drinking, as it did at the lowest levels of use.

At some point, the returns to drink itself start declining on the average. Average returns to 1 fall off because tolerance develops to alcohol (as to most but not all addictive substances; see Koob and Bloom, 1988). The effect of tolerance is to reduce the value gained from a unit of the substance. But even without tolerance in the familiar physiological sense, it is likely that, at higher consumption rates, longer chains of activities are involved in drinking, hence, they are lower on the delay of gratification gradient (also see footnote 5). At C, which is the stable matching point at high rates of indulgence, both drinking and not drinking have lost much of their capacities to please. Neither "yes" nor "no" works well in the vicinity of C, yet there can be a stable matching point here.

Fully Constrained Addiction Figure 13.2 shows how certain hypothetical configurations of functions imply equilibria that are qualitatively reminiscent of the clinical literature of alcoholism and other forms of substance abuse. In Figure 13.3, the picture is adapted to the literature of gambling. Given the disposition of the value functions here, the choice distribution would not stabilize at any point between the left or right corner, where no or all discretionary finan-

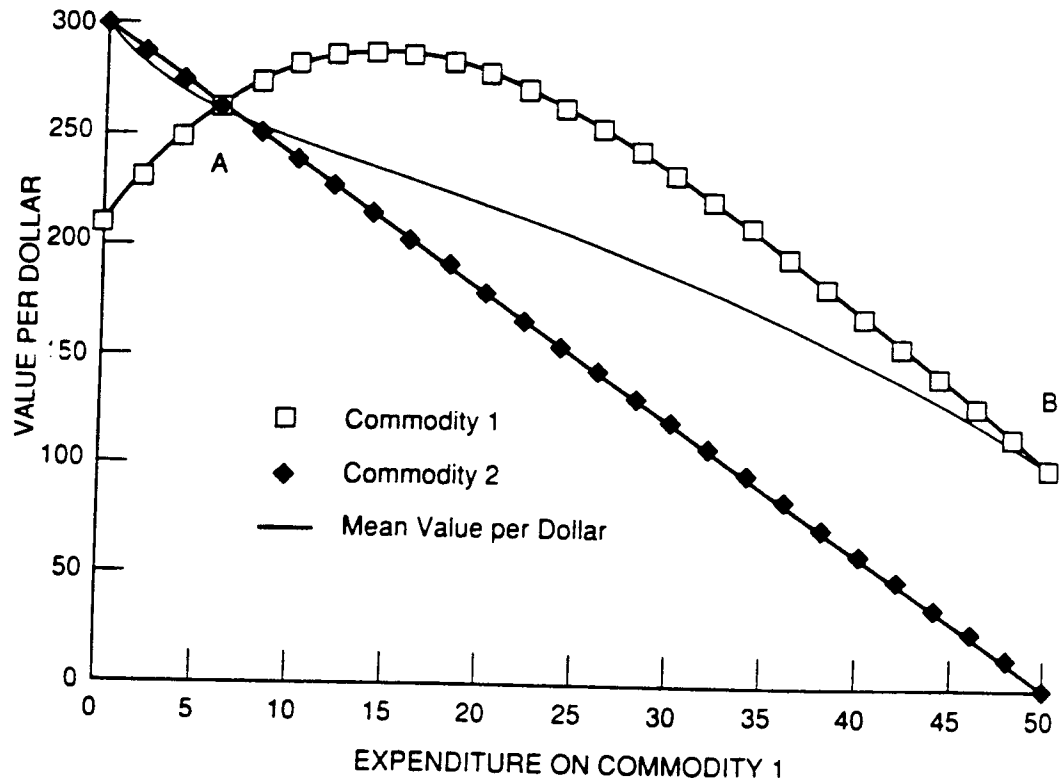


Figure 13.3. Value as a function of allocation to alternative 1, for alternative 1 (squares) and alternative 2 (diamonds). A is an unstable matching point, a knife edge, and B is the stability point for a fully constrained addiction.

cial resources, respectively, are consumed by gambling.⁶ The crossing functions, at A, are at an unstable matching point, which is to say, at a "knife edge." The right corner, B, is the point of *fully constrained addiction*, where expenditure on the addictive activity is only held in check by the sheer lack of additional resources. In hedonic terms, however, the person is still not at equilibrium.

Figure 13.3 assumes that as more financial resources are plowed into gambling (squares), the average returns from gambling will initially rise because the learning of skills or the development of tastes for the paraphernalia of gambling are rate-dependent. But at high rates of indulgence, the returns will decline if riskier bets are taken or if social disapprobation becomes part of the harvest of the activity.

Likewise, the returns from not gambling (diamonds) will also decline with increased allocation to gambling. The increasingly small

⁶To show the hypothetical value functions for a person who also has a stable matching point for low, rather than zero, levels of gambling, we need only complicate the chart slightly, borrowing the left-most end of the chart shown in Figure 13.2.

resources left to the activities other than gambling are swept up by utter necessities, enough food to live rather than to enjoy, mere shelter rather than a home to take pleasure from, and so on. Little is left for other recreational activities. Instead of vacations, gifts for one's loved ones, and so on, the shrunken complement to gambling is preempted by the need to attend to often harrowing crises in one's life, many of which are themselves the by-products of gambling. Like the alcoholic or drug addict, the compulsive gambler slides toward allocations where neither gambling nor not gambling can give much pleasure.

In constructing Figure 13.3, we charged the negative by-products of gambling against activities other than gambling, rather than to gambling itself, because this is the way people tally the consequences of their actions, according to the clinical literature. Quarrels with one's spouse, triggered by gambling, are more likely to poison domestic life than to ruin gambling. In the literature of gambling, people reform precisely when they adopt a more sophisticated approach to mental bookkeeping, and place the blame where it really belongs (see Orford, 1985; Shaffer, Stein, Gambino, and Cummings, 1989).

The distinction between constrained (as in Figure 13.3) and naturally equilibrating (as in Figure 13.2) addiction is important, yet difficult to draw precisely. Financially draining addictions, such as gambling or expensive drugs, most probably leave the addict fully constrained. Food, tobacco, or video games are good examples of unconstrained addiction, for most people, at least. Alcohol consumption could be constrained or not, depending on whether the person can still draw on a steady income.

Some addictions are constrained by sheer physical restrictions on choice rate. A baseball "addict" can only watch his team play 162 different games during the regular season; a workaholic or computer addict is constrained by the clock, as are individuals caught in a "relationship addiction."

Transitions Other things being equal, a person who is fully constrained (Figure 13.3) should find it more difficult to break out of addiction than someone who is at an intermediate equilibrium (Figure 13.2). At or near B in Figure 13.3, not gambling would feel hedonically inferior to gambling even though virtually all resources are going into gambling. Moving leftward, toward not gambling, would require one to endure a hedonically inferior alternative for some extended period of time (until point A is reached).

The situations in Figures 13.1 or 13.2 are quite different. In Figure 13.1, a person who starts to move leftward from point A (or from

point C in Figure 13.2), toward less consumption of the problematic alternative, should immediately experience an increase in satisfaction derived from *both* activities. Consequently, it may seem that only unawareness of the shape of the value functions could keep a person stuck at an intermediate, unconstrained addictive equilibrium.

But the charts have so far shown value functions only in the long run, that is to say, after each level of allocation has lasted long enough so that tastes at that point are no longer changing. Tastes interact with consumption, but the interaction may be slow, because the controlling variables for values—the basic taste changes, learning processes, social reactions to our behavior, and so on—may themselves be only slowly driven by distributed choices. For some suboptimal behavior, much, if not most, of the explanation of its tenacity probably resides in more rapid interactions with rates of consumption, the transients in the value functions as a person shifts from one allocation to another.

Goldstein and Kalant (1990) state two “special characteristics” of chemical addictants: The first is the development of tolerance, which is a decline in the hedonic yield of a given amount of the substance. Tolerance is represented in our theory as a steeply negative relation between *long-term* consumption rates and hedonic yield per unit of consumption, as shown in the earlier figures. The second characteristic is physical dependence, which leads to the *short-term* symptoms often called *withdrawal*, which we introduce at this point.

Transitory effects are especially relevant to relatively unconstrained addictions, such as smoking. The steady smoker does not experience sharp hedonic fluctuations over time, as long as cigarettes are consumed at a steady rate. The smoker, therefore, is at a point where his value functions intersect. Yet, interrupting consumption would rapidly bring on withdrawal symptoms—headaches, anxiety, tremors, and the like—which only the resumption of smoking can alleviate.

These temporary changes in value are not captured in any of the value functions considered so far. In our conceptual framework, withdrawal should appear as a shift in the value functions that is produced by a relatively brief interruption in consumption. Figure 13.4 describes a situation in which the impact of the interruption is felt as a severe reduction in the value of the nonaddictive activity. The diagram contains two value functions for this activity: The higher value function for alternative 2 (crosses) represents the satisfaction derived from nonaddictive activities *if* the consumption of the addictive substance over some relatively short-time period has exceeded a threshold level. If consumption does not meet this level, then with-

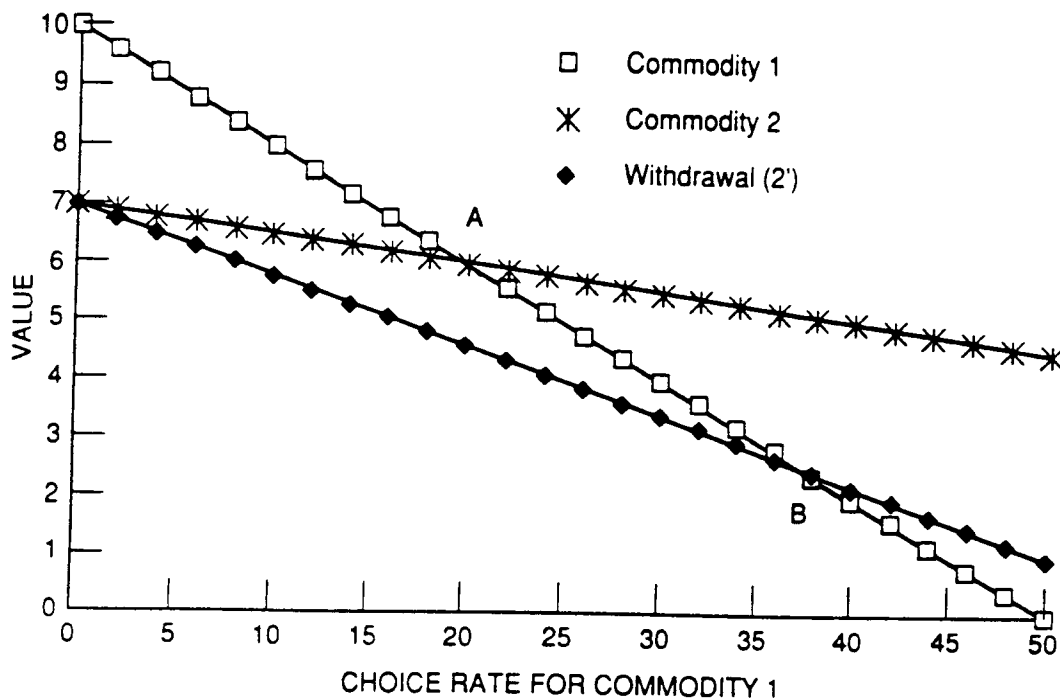


Figure 13.4. Value as a function of allocation to alternative 1, for alternative 1 (squares) and for alternative 2 if consumption of 1 is above (crosses) or below (diamonds) a threshold level. The threshold level is the level of consumption of 1 below which withdrawal symptoms appear. A is the matching point if consumption of 1 is above the threshold; B is the matching point if it is below the threshold.

drawal symptoms appear, and nonaddictive activities are evaluated according to the lower function (the *withdrawal curve*, i.e., diamonds labeled 2').

Figure 13.4 assumes, therefore, that the cost of saying no is temporarily escalated by withdrawal, hence, its net returns are down just as the person is trying to say no more often. Melioration under these conditions dictates shifting toward more of the addictive activity. The precise track of this process depends both on the time course of the short-term withdrawal symptoms and the speed with which the distributed choice process readjusts long-term tastes. In effect, the melioration process that drove the person down the primrose path to the addictive matching point is thus stacked even further against him should he try to reduce the level of indulgence.

A comparable transient influence may act on the addictive commodity itself. It has been observed that large changes in the nicotine and tar content of cigarettes may have large compensating short-term effects, but, in the long term, their effect on the consumption rates of smokers is relatively small (Gori, 1980; Russell, 1979). If smokers

were regulating their habit so as to attain a preset level of nicotine or tar concentration in their bodies, as is approximately the case with certain parameters of food (e.g., calories, proteins), then it would follow that reducing nicotine (or carbon monoxide or any of the other habit-forming components of tobacco smoke) should permanently increase the number of cigarettes smoked. Instead, the long-term rate of smoking evidently sets the target levels for the habit-forming components of smoking within a broad range of consumption rates. One "gets used to" a certain physiological level of the addictant. Smoking is a self-regulating activity, but the regulatory levels are arbitrary across a range. For other "habit-forming" activities as well, a given level of indulgence sets the values of parameters around which behavior is regulated.

The opponent-process theory of motivation (Solomon and Corbit, 1974) attempts to explain this interaction between past consumption level of certain substances or activities and present motivational state by postulating homeostatic mechanisms. Substances such as nicotine, for example, induce a change in affective state. A homeostatic mechanism in the body tends to counteract that state, as part of the general tendency of the body to hold certain variables within a narrow range. If the initial affect is pleasant, the counteracting affect is unpleasant, and vice versa. The "opponent process" is affectively opposite to the original agent, but it is slower acting. As a result, after repeated exposures, it leaves the organism with a diminished initial affective reaction to the original agent, but a lingering opponent reaction.

Opponent-process theory, or something like it, handles some of the phenomena of addictions or other habit-forming commodities (Solomon, 1977; Solomon and Corbit, 1973). Rozin and his associates have suggested (Rozin, 1977; Rozin, 1982; Rozin and Fallon, 1981; Rozin and Schiller, 1980) that, for example, the taste for certain foods, such as that for chili pepper, as well as for tobacco, caffeine, and other substances, seems to conform to an opponent-process theory. The initial burning effect of chili is disagreeable, but, after one habituates to it, the lingering opponent process is pleasant, and it is the opponent process that sustains its use among habitual users. A substance that is a source of negative affect at low rates of consumption has become a source of pleasure at high rates. For substances that produce a disagreeable opponent process (known as withdrawal), such as morphine and alcohol, further indulgence in the substance itself is the antidote (as already implied by Figure 13.4 in the gap between the lines labeled 2 and 2'). Here, a high rate of consumption creates a potentially potent source of reinforcement: the avoidance of withdrawal symptoms. Whether the initial affective state is positive

or negative, any attempt to decrease an established level of indulgence temporarily intensifies the desire for the substance and presumably also the reinforcement gained from consuming it.

Social and Divided Selves

The first time we let the lawn get out of hand, we may not have realized how much we will dislike the contrast between a neighbor's green grass and our green weeds. But after a few summers, we realize it all too well. Likewise, we may be warned that we are overeating, but still not know just how unpleasant it feels to catch a glimpse of our overweight selves in a mirror. Once we get there, we know. Even a meliorator, insensitive as he is to marginal returns, may know what the aggregate returns look like across a range of choice allocations.

The question is how to incorporate the effects of knowledge about the aggregate returns into the present theoretical framework. Ulysses was worldly wise enough to know the risks, ahead of time, of sailing too close to the Sirens' song and, therefore, had himself tied to the mast. Besides such individualistic solutions, the accumulated wisdom of a culture is full of warnings about the pitfalls of distributed choice. The wisdom is built into many of a culture's institutions.

Isolated and in a state of nature, we would frequently distribute choices suboptimally. The interactions between tastes and allocation are not, as a general rule, known innately, so they must be discovered by experimentation across ranges of allocations. As meliorators, however, nothing impels us to spend the time and effort looking into the black box of our taste changes. And even if we did look, we often seem to lack a natural way to keep a record of our observations. Our memories of the returns from past allocations do not usually enable us to calculate, say, just how often we should eat caviar versus hamburger in order to optimize across those alternatives. We are thus likely to ramble down the primrose path to various suboptimal equilibria, many of them far worse than overindulgence in caviar.

But in a community of meliorators, knowledge about at least some of those primrose paths would accumulate. For meliorators, it matters whether we are observing other people's behavior or just our own. Observing just our own, we are in the clutches of our own current value functions and our impoverished memories of past ones. Observing others, it would soon be noted that, as pleasing as wine may be, it can become "too much of a good thing." We would soon be saying "enough is enough," "all things in moderation," or, simply,

"know thyself." Those familiar aphorisms exemplify how human culture provides meliorators like us with rules to help us cope with our suboptimality. What follows is a sketchy attempt to classify how meliorators cope.

Social Supplementation

Society routinely exploits our sensitivity to social reinforcers. For example, we praise hard work and deplore indolence. Most people are attuned to the social reinforcers attached to their choices, and many of those reinforcers are indexed to the rates of consumption of the commodity in question. Eyebrows rise when we ask for the second, let alone third or fourth, slice of cake. Without the social supplements, many of us might work less and eat more cake than is optimal for us individually, let alone optimal for society as a whole. People who are insensitive to social reinforcers, the sociopaths, are dangerous not only to society but to themselves (Wilson and Herrnstein, 1985).

The social supplements are often keyed to circumstances—"love and marriage go together like a horse and carriage," an old song said, affirming that love (i.e., sex), like many other activities, has its proper time and place. Activities are labeled as vices, sins, evil, virtuous, etc., which is the language used to signify whether the social supplements are positive or negative and how intense they are. The concept of salvation itself is a kind of social supplement, promising benefits in the hereafter for right conduct here. When experts decide that such behaviors as gambling, excessive alcohol consumption, habitual criminality, and so on are better dealt with as "diseases" than as the "vices" they were long taken to be (Orford, 1985), an inadvertent result may be more gambling, drinking, and crime, as if a tax has been lifted from them, unless the stigma of disease is itself an incentive that fully replaces the moral incentives lost by abandoning the language of vice and virtue. By disapproving of behavior, we deter it, insofar as our disapproval has net, negative incentive value. This disregards any compensating benefits of the disease approach to bad behavior, which may be considerable.

In the framework of standard economic theory, the need for social supplements would be most readily accounted for in terms of the fundamental problem of welfare economics—that individually rational behavior may be collectively damaging. We tax the grazing of livestock on the commons and the use of other public goods, so as to avoid the "tragedy of the commons," which is to say, to deter

individual freeloading. But meliorators need social supplements, too, in order to make their choices individually more nearly optimal than they would otherwise be. As Herrnstein and Prelec show (Chapter 10), the problem of externalities in welfare economics has an analogue in the allocations of individuals.

Because distributed choices often go haywire by stabilizing at the wrong rate, society invents many *rate-setting customs*. The optimal amount of alcohol consumption is probably not zero, but some relatively low level, at which one can enjoy some of its disinhibiting and analgesic effects, without wrecking the rest of one's life. For some people, the process of melioration fails to provide a stable matching point in this range of allocations, and the primrose path would lead to trouble, except for social practices that lend support to the desirable range of allocations. A daily cocktail or two before dinner, beer in 12-ounce bottles, a few holidays enlivened by drink, occasional religious observances with a taste of wine, are examples of rate-setting customs that detour the primrose path away from danger. A beer commercial on television advises us to "know when to say when."

There will still be some who, because of physiological constitution, past history, or present circumstance, succumb to excess indulgence even in the face of social protections against it. For alcohol, total prohibition remains a possibility, a potentially unstable matching point that has some hope of enduring when supported by the techniques of self-management that Ainslie and others have discussed. But for many bad habits, the "just say no" option is not feasible. We cannot say no to all eating, and it seems too harsh and costly to say no to all leisure, all sex, all unessential spending, and so on. Rate-setting practices, other than total prohibitions, therefore abound.

Here are some commonplace examples. For eating, the norm is three meals a day, and the meals themselves fall within a roughly prescribed range of sizes and nutritional contents. For people whose value functions might lead them to overeat, undereat, or otherwise malnourish themselves, the norm guides them toward a more desirable range of choices. Someone who eats six eggs and a cheeseburger for breakfast is likely to get immediate negative feedback from his dining companions; likewise for someone who eats no breakfast at all. For the "right" rate of leisure, we have the 9-to-5 working day, 5 days a week. The sabbath itself is an ancient and hallowed version of a rate-setting device for leisure, and also for religious observances. For charitable donations, tithing or other rules of giving help us maintain a level that we would like to choose but may find it difficult to attain because we meliorate. Daily clean clothes and a bath may stipu-

late a level of personal hygiene that people might otherwise frequently undershoot, to their own regret as well as ours.

In these, and many other, cases, the mechanism of control is the application of social incentives, pro or con, to constrain behavioral allocation beneficially, at least on average. But insofar as individuals vary, in value functions and in environmental circumstances, the social supplements are likely to be only crudely helpful in particular cases. Methods of self-management more finely tuned to the individual's own problems, therefore, need to be considered.

Buying a magazine or concert subscription is a way to transform one's own distributed choice into something closer to a once-and-for-all choice, once we realize that the distributed choices are suboptimal as we reckon our own utility. In effect, we buy a certain rate of allocation. Subscriptions often discount unit cost, but they probably need not do so in some instances. If we think we *should* go to the symphony or to the exercise clinic more often than we find we do, we may be willing to pay a premium for a package deal, as a way of buying an incentive to control our future behavior.

Divided Self

Once we buy an incentive to control our future behavior, we have begun to act as if more than one person, pulling in different directions, lives in our skin. Person A within us spends money for a subscription or some other rate-constraining device, hoping to influence the other person, B, also within us, because he knows from past experience that B may otherwise act against A's wishes. At different times, our preferences can be so inconsistent that we may entertain notions of demonic possession, loss of will, multiple personality, and the like. Modern theorists, particularly Ainslie, Elster, Schelling, Strotz, Thaler, and Winston, have written illuminatingly about intertemporal inconsistency. The theories explaining the inconsistency vary, but the fact itself is too much a part of everyone's experience to be in serious doubt.

Only saints and liars can say they have not experienced such problems as the following: Early in the semester, we resolve to keep up with the homework assignments, but then, having failed to do so, we run out of time to catch up and fail to earn the grade we think we could have earned. It may be a surprise to discover our failure the first time, but some students go through this disappointment semester after semester; likewise resolutions to exercise more, eat less cake, save more money, read better books, watch less television,

drink less beer, and so on. These are all problems of distributed choice, involving rates of consumption that violate our own assessments of utility and rationality. They are also problems that we may, in time, discover about ourselves, and we often work hard at trying to solve them. Solutions may become extreme. For example, a newspaper story (Pitt, 1988) described a man who tried to cure his drug habit by getting himself arrested and locked up in jail, where he would be treated for drug addiction while being held in a relatively drug-free environment. He broke windows in two police stations, displayed hypodermic needles and empty crack vials in another, but found "himself back on the streets each time, his pleas ignored." A judge finally helped out by persuading a private religious organization to take him into its drug rehabilitation program.

Not only have we gone down, or at least looked down, the primrose path, but we know that we must not let ourselves be guided by the current value functions. Even after the value functions have been molded by social conventions, religious teachings, and ethical principles, we may still face trouble with our choices. It is when we know our distributed choices are suboptimal that we become conscious of a process of decision making, one that is attuned to our personal irrationalities.

Ainslie (1975) has depicted the problem of conscious decision making as one involving two hyperbolically discounted delay-of-reinforcement curves. On a given occasion, we may confront a pair of alternatives, with the larger payoff more delayed than the smaller. Because time discounting is hyperbolic (rather than exponential; see Chung and Herrnstein, 1967; Mazur, 1987; Prelec, 1989), preference may switch from the later, larger reinforcement to the earlier, smaller one, as the earlier one becomes imminent. When the alternatives are both remote, the later, larger one is again preferred.

While this theory accords with both subjective experience and experimental data about time discounting, it needs to be extended to distributed choices and primrose paths, where the problem is not so much that two alternatives have temporally displaced consequences (although they may, in many cases), but that the value functions depend on levels of allocation, as illustrated in the earlier figures, and one cannot rapidly sample from other points across the range of allocations. Someone who continually watches television 12 hours a day has a hard time discovering that life would improve with a smaller ration of watching time, because the value functions for different levels of allocation are not simultaneously available for comparison. Prelec and Herrnstein (1991) have discussed some of the complexities of making choices "in the small"—for example, a drink

versus no drink—as opposed to making them “in the large”—for example, drinking just one drink before dinner versus starting the drinking day at lunch. The former is the time discounting problem in the small; the latter is that in the large.

Addictions evidently belong to that class of behaviors for which the time discounting problem arises in the large. A person who knows he has a drinking problem, for example, is likely to want to choose a lower rate in general, but the rate he wants in the large is lower than the rate he gets by aggregating the consequences of his answers in the small. Many of the prudential rules or principles of conduct that people construct to guide their own behavior are attempts to resolve discrepancies of this form.

Conclusion: Policy Implications

It may be useful to distinguish the four approaches to addiction described in the first section by the class of policy recommendations that flow from them.

The *no-choice* approach implies a search for interventions to cure the “disease” presumably underlying the unwanted behavior. Rather than rearranging the incentives that govern choice, this approach would focus on the “root causes” of the troublesome behavior and attempt to alter them. The person should, in any case, not be held accountable for the behavior or afflicted with a sense of guilt for engaging in it, for doing so would be tantamount to trying to control behavior by incentives when incentives are, presumably, irrelevant.

The *rationality* approach implies that some people are addicts because of their endowment (where endowment includes material, constitutional, and psychological aspects). Consequently, the most efficient way to increase their welfare is to give them more money. The public policy goal for drug abusers, for example, would be to make it possible for addicts to self-medicate more easily. Disregarding the possible offsetting social welfare penalties for allowing addicts freer rein to pursue their habits, the logic of this approach argues against criminalization of addictive substances, or any other form of government action that raises prices or otherwise impedes their use (see, e.g., Becker and Murphy, 1988; Friedman, 1989; Nadelmann, 1989). It also argues against attempts to change demand by education, restrictions on the advertising of addictive substances, and the like. There should be a market for addiction insurance, that is, for insuring against the high costs of drugs, should stressful events lead to a need

for drug medication. Even the restriction of substances to children violates this approach in its most thoroughgoing form. If there is an overriding social welfare argument against drug abuse even though it is individually optimal, the "rationalists" have, to our knowledge, not yet made it. The argument would take the form of showing that, even though any particular addict is making optimal choices, the utility of society as a whole is reduced by the addiction. This would not be an easy argument to make rigorously, because for every compassionately suffering loved one of an addict,⁷ there may be more than one supplier of goods and paraphernalia benefiting from the addiction, along with the addict.

The *primrose path* approach suggests that society should at least provide people with more information, on the grounds that they are less likely to go down the path if they know where it is headed. A useful thought experiment in this regard is to ask whether the propensity toward addiction would be reduced by biofeedback that makes the buildup of tolerance or the imminence of physical dependence more visible. If the answer is yes, then the *primrose path* notion is strengthened at the expense of both of the foregoing approaches. More fanciful types of feedback might include actuarial information (vividly presented) about one's career, health, family, prospects, as a function of current behavior. Besides altering the information a person has about the long-term effects of choices, public policy could attempt to influence the value functions themselves, perhaps by applying social reinforcers to make the addictive activity less attractive in relation to its alternatives. Again, this approach splits the difference between the preceding two. It accepts the reality and significance of incentives in relation to addiction, but it does not assume that the individual can perform a rational calculation in weighing those incentives. If the first approach rejects criminalization because it is irrelevant, and the second approach rejects it because it is against the addict's self-interest, the *primrose path* approach may favor it for providing incentives where they could be truly helpful.

Finally, the *divided self* approach extends the analysis of the *primrose path*. It acknowledges that individuals can sometimes rearrange the value functions controlling their own behavior. The public policy implication is to make self-control mechanisms more accessible and more available than they evidently are now for those who succumb. Addiction, in other words, is like market failure, in that the market

⁷And why are they compassionately suffering, if their addicted loved one is maximizing his or her utility?

does not supply convenient self-control devices (and the legal system prohibits some of them). In the absence of external precommitment mechanisms, the individual is forced to invent personal strategies for self-control (as Ainslie and Schelling, among others, have described: see Ainslie, 1975, 1986; Schelling, 1978, 1980). Public policies could be devised to make the teaching and dissemination of self control more effective.

We wish to thank George Ainslie for detailed comments on an earlier draft. Thanks are also owed to the Russell Sage Foundation for support.

References

- Ainslie, G. "Specious Reward: A Behavioral Theory of Impulsiveness and Impulse Control." *Psychological Bulletin* 82 (1975): 463-509.
- . "Beyond Microeconomics: Conflict Among Interests in a Multiple Self as a Determinant of Value." In J. Elster (ed.). *The Multiple Self*. Cambridge, UK: Cambridge University Press, 1986.
- Becker, G.S., and K.M. Murphy. "A Theory of Rational Addiction." *Journal of Political Economy* 96 (1988): 675-700.
- Booth, W. "Social Engineers Confront AIDS." *Science* 242 (1988): 1237-1238.
- Chung, S.-H., and R.J. Herrnstein. "Choice and Delay of Reinforcement." *Journal of the Experimental Analysis of Behavior* 10 (1967): 67-74.
- Elster, J. (ed.). *The Multiple Self*. Cambridge, UK: Cambridge University Press, 1986.
- Friedman, M. "An Open Letter to Bill Bennett." *Wall Street Journal* (1989): September 7 A16.
- Goldstein, A., and H. Kalant. "Drug Policy: Striking the Right Balance." *Science* 249 (1990): 1513-1521.
- Goleman, D. "Biology of Brain May Hold Key for Gamblers." *New York Times* October 3 (1989): C1, C11.
- Gori, G.B. "Less Hazardous Cigarettes: Theory and Practice." In G.B. Gori and F.G. Bock (eds.). *A Safe Cigarette?* Cold Spring Harbor, NY: Cold Spring Harbor Laboratory, 1980, pp. 261-279.
- Herrnstein, R.J. "Melioration as Behavioral Dynamism." In M.L. Commons, R.J. Herrnstein, and H. Rachlin (eds.). *Matching and Maximizing Accounts*. Cambridge, MA: Ballinger, 1982, pp. 433-458.
- , and D. Prelec. "Melioration: A Theory of Distributed Choice." *Journal of Economic Perspectives* 5 (1991): 137-156.

- , and W. Vaughan, Jr. "Melioration and Behavioral Allocation." In J.E.R. Staddon (ed.). *Limits to Action*. New York: Academic Press, 1980, pp. 143–176.
- Irwin, S. *Drugs of Abuse*. Tempe, AZ: Do It Now Foundation, 1990.
- Jellinek, E.M. *The Disease Concept of Alcoholism*. New Haven, CT: College and University Press, 1960.
- Koob, G.F., and F.E. Bloom. "Cellular and Molecular Mechanisms of Drug Dependence." *Science* 242 (1988): 715–723.
- Leonard, D. "Market Behavior of Rational Addicts." *Journal of Economic Psychology* 10 (1989): 117–144.
- Lesieur, H.R. "Current Research into Pathological Gambling and Gaps in the Literature." In H.J. Shaffer, S.A. Stein, B. Gambino, and T.N. Cummings (eds.). *Compulsive Gambling: Theory, Research, and Practice*. Lexington, MA: Lexington Books, 1989, pp. 225–248.
- Mazur, J.E. "An Adjusting Procedure for Studying Delayed Reinforcement." In M.L. Commons, J.E. Mazur, J.A. Nevin, and H. Rachlin (eds.). *The Effect of Delay and of Intervening Events on Reinforcement Value*. Hillsdale, NJ: Erlbaum, 1987, pp. 55–73.
- Nadelmann, E.A. "Drug Prohibition in the United States: Costs, Consequences, and Alternatives." *Science* 245 (1989): 939–947.
- Orford, J. *Excessive Appetites: A Psychological View of Addictions*. Chichester, UK: Wiley and Sons, 1985.
- Pattison, E.M., M.B. Sobell, and L.B. Sobell, *Emerging Concepts of Alcohol Dependence*. New York: Springer, 1977.
- Pitt, D.E. "Judge Hears Addict's Plea to Help Him." *New York Times* December 23 (1988): B1–B2.
- Prelec, D. "Matching, Maximizing, and the Hyperbolic Reinforcement Feedback Function." *Psychological Review* 89 (1982): 189–230.
- . *Decreasing Impatience: Definition and Consequences*. Harvard Business School Working Paper 90-015, 1989.
- , and Herrnstein, R.J. "Preferences or Principles: Alternative Guidelines for Choice." In R.J. Zeckhauser (ed.). *Strategy and Choice*. Cambridge, MA: MIT Press, 1991, pp. 319–340.
- Roy, A., B. Adinoff, L. Roehrich, D. Lamparski, R. Custer, V. Lorenz, M. Barbaccia, A. Guidotti, E. Costa, and M. Linnoila. "Pathological Gambling: A Psychobiological Study." *Archives of General Psychiatry* 45 (1988): 369–373.
- , J. De Long, and M. Linnoila. "Extraversion in Pathological Gamblers: Correlates with Indexes of Noradrenergic Function." *Archives of General Psychiatry* 46 (1989): 679–681.
- Rozin, P. "The Significance of Learning Mechanisms in Food Selections: Some Biology, Psychology and Sociology of Science." In L.M. Barker, M.R. Best, and M. Domjan (eds.). *Learning Mechanisms in Food Selection*. Waco, TX: Baylor University Press, 1977, pp. 557–589.
- . "Human Food Selection: The Interaction of Biology, Culture and Individual Experience." In L.M. Barker (ed.). *The Psychobiology of Human Food Selection*. Bridgeport, CT: AVI, 1982, pp. 225–254.

- , and A.E. Fallon. "The Acquisition of Likes and Dislikes for Foods." In J. Solms and R.L. Hall (eds). *The Role of Food Components in Food Acceptance*. Zurich: Forster, 1981, pp. 35–48.
- , and D. Schiller. "The Nature and Acquisition of a Preference for Chili Peppers." *Motivation and Emotion* 4 (1980): 77–101.
- Russell, M.A.H. "Tobacco Dependence: Is Nicotine Rewarding or Aversive?" In N.A. Krasnegor (ed.). *Cigarette Smoking as a Dependence Process*. Rockville, MD: National Institute of Drug Abuse, 1979, pp. 100–122.
- Schachter, S. "Some Extraordinary Facts About Obese Humans and Rats." *American Psychologist* 26 (1971): 129–144.
- Schelling, T.C. *Micromotives and Molar Behavior*. New York: Norton, 1978.
- . "The Intimate Contest for Self-Command." *The Public Interest* No. 60 (Summer 1980): 94–118.
- Shaffer, H.J., S.A. Stein, B. Gambino, and T.N. Cummings. (eds.). *Compulsive Gambling: Theory, Research, and Practice*. Lexington, MA: Lexington Books, 1989.
- Solomon, R.L. "An Opponent-Process Theory of Motivation: IV. The Affective Demands of Drug Addiction." In J.D. Mazer and M.E.P. Seligman (eds.). *Psychopathology: Laboratory Models*. San Francisco: W.H. Freeman, 1977.
- , and J.D. Corbit. "An Opponent-Process Theory of Motivation: II. Cigarette Addiction." *Journal of Abnormal Psychology* 81 (1973): 158–171.
- , and ———. "An Opponent-Process Theory of Motivation: I. Temporal Dynamics of Affect." *Psychological Review* 81 (1974): 119–145.
- Thaler, R., and H.M. Shefrin. "An Economic Theory of Self-Control." *Journal of Political Economy* 89 (1981): 392–410.
- Vuchinich, R.E., and J.A. Tucker. "Contributions from Behavioral Theories of Choice to an Analysis of Alcohol Abuse." *Journal of Abnormal Psychology* 97 (1988): 181–195.
- Wells, L. "Conspicuous Consumers." *New York Times Sunday Magazine* November 20 (1988): 90.
- Wilson, J.Q., and R.J. Herrnstein. *Crime and Human Nature*. New York: Simon and Schuster, 1985.
- Winston, G.C. "Addiction and Backsliding." *Journal of Economic Behavior and Organization* 1 (1980): 295–324.