GEORGE LOEWENSTEIN

WILLPOWER: A DECISION-THEORIST'S PERSPECTIVE *

(Accepted September 3, 1999)

> The man who acts according to the rules of perfect prudence, of strict justice, of proper benevolence, may be said to be perfectly virtuous. But the most perfect knowledge of those rules will not alone enable him to act in this manner; his own passions are very apt to mislead him – sometimes to drive him, and sometimes to seduce him, to violate all the rules which he himself, in all his sober and cool hours, approves of.
>
> Adam Smith, *Theory of Moral Sentiments*, 1892: 349

During a recent camping trip in Alaska a friend (a fellow economist) and I found ourselves paddling against wind and waves on a huge lake, struggling to meet the appointed time for pickup by a float plane. Hour after hour we paddled in wet misery, making painfully slow forward progress and sometimes inching backwards when conditions got particularly nasty. At every moment we felt tremendously tempted to take a break, and, when we succumbed to this temptation, it was exceedingly difficult to motivate ourselves to resume the effort. As my companion commented, our limits were defined not by muscle power (although more would have helped) but by *willpower*.

As economists, we should have been disturbed by my friend's comment because there is no place for willpower in economics or rational choice theory. Economists and others who embrace the decision-making perspective see behavior as a matter of simply

* Paper prepared for 1998 "Workshop on Addiction and Responsibility" Sponsored by the University of North Carolina School of Law Department of Philosophy, and the National Humanities Center. I am grateful for comments and suggestions from Caroline Acker, Shane Frederick, Donna Harsch and Ted O'Donoghue. The paper was written during my sabbatical at the Center for Advanced Study in the Behavioral Sciences, which was supported by NSF grant # SBR – 960123 to the Center.

Law and Philosophy **19:** 51–76, 2000.

choosing between behavioral options. Once a particular behavioral option is selected, its implementation is assumed to be unproblematic. The notion of willpower implies that the execution of a chosen course of action is not always automatic. In some cases we need to *motivate ourselves* to carry out a desired sequence of behavior. I use the term willpower in an old-fashioned and intuitive fashion, as a kind of inner force (*power* of will) that is exerted with the purpose of controlling our own behavior. My goal in this paper is to explore some of the characteristics of willpower, begin to think about how willpower could be incorporated into decision-analytic accounts of human behavior, and discuss its implications for theoretical accounts of drug addiction.

WILLPOWER, CONSCIOUSNESS, AND THE ANIMAL BRAIN

The concept of willpower suggests that there is some part of the self that needs to be controlled to do what another part of the self wants. What part of the self is doing the controlling, and what part is being controlled by willpower? Casual scrutiny of situations that involve an imposition of willpower points to a simple answer to this question: Willpower is almost always employed to suppress or override some type of 'visceral' motivation. In previous papers (Loewenstein, 1996, Loewenstein, 1998, Loewenstein and Schkade, 1998; Loewenstein, Nagin and Paternoster, 1997) I have defined visceral factors to include three main categories of motives:

(1) *drives*, such as hunger and sexual desire;
(2) *emotions*, such as anger and fear;
(3) *somatic sensations*, such as pain.

All visceral factors serve important survival and reproductive functions. Hunger and thirst ensure that we ingest enough food and fluids, the sex drive ensures that we reproduce, anger protects us from exploitation by others, and pain protects us against tissue damage. In each case, visceral factors serve this function in part by creating an aversive sensation if the associated need is not met and in part by increasing the subjective desirability of satisfying the need. For example, when your core body temperature declines you feel uncomfortable, and activities that will warm you up, such as drinking hot liquids, become pleasurable. As visceral factors intensify,

this combination of avoiding pain and seeking pleasure can become an exceedingly powerful motivational force.

Although visceral factors perform important survival functions, they also occasionally propel us in directions that are perceived, even at the moment of acting, to conflict with self-interest. Indeed historically, visceral factors (typically under the label of 'passions') have often been seen as an exclusively corrupting force, counterpoised against the civilizing force of 'intellect', 'reason', or 'self-interest'. With easy access to high-calorie foods, for example, most of us are harmed by eating whenever we are hungry. Likewise, there are social sanctions associated with overreacting to anger, fear, or pain.

Animals and humans share visceral factors to a much greater extent than we share cognitive capabilities. Humans and many other animals experience hunger, thirst, fear, pain, and, seemingly, a range of emotions. The most significant neurophysiological difference between humans and our close animal relatives is the size of the cortex, and, reflecting this difference, the most significant difference in neural functioning are in capabilities that are closely associated with cortical functioning – most notably language and consciousness.

In general, animals appear to be 'slaves' of their drives, appetites and emotions, in the sense that they eat when they are hungry, copulate (if possible) when sexually aroused, and avoid or flee from sources of fear or pain. Due to the capacity for consciousness, humans have the ability to reflect on the broad consequences of their behavior and to identify conflicts between viscerally-driven behavior and other 'higher level' motives or goals such as physical fitness or respect from other persons.

My basic argument is that *willpower* represents *attempts to suppress viscerally-motivated* behaviors that conflict with *higher level goals*. I use the term "higher level" motives to refer to those that are cognitively mediated. Although referring to such motives as being higher might seem arbitrary, there is actually some justification for use of the term. The cortex is not only newer in evolutionary terms than the limbic system, which is the primary seat for visceral factors, but is actually situated physically above the limbic system.[1]

[1] Exactly how the cortex gets through to the limbic system is not well understood. At a neurophysiological level, there seems to be some degree of agreement

In some cases people use willpower to *inhibit* action – e.g., to avoid eating when they are hungry, refrain from having sex, expressing their own anger, or succumbing to fear or pain. In other cases, willpower is used to motivate behavior that would otherwise be suppressed by visceral factors. For example, people with stage fright require willpower to go on stage.

Willpower is commonly employed in the service of long-term objectives such as health (through dieting) or wealth (through working or saving), but it is also sometimes exercised in support of what would appear to be very short-term goals. For example, to mainline heroin for the first time (seemingly, a short-term gratification), an individual might require willpower to overcome an aversion to sticking herself with a hypodermic needle. Willpower, in this example, is used to suppress an immediate visceral reaction for an only slightly-less immediate reward. Likewise, someone who is pathologically obsessed with saving money (a long-term objective) might require willpower to splurge on expensive meal (a short-term benefit). Again, however, willpower would only be required for such a splurge if the individual's immediate visceral response to the indulgence was negative.

Although willpower can serve either short- or long-term gratifications, exertions of willpower always involve a sacrifice of *immediate* utility (pleasure, happiness). Exertions of willpower always make you feel immediately worse because visceral factors motivate specific behaviors by making them more immediately pleasurable, at least relative to *not* performing the behavior. Exercising willpower is therefore always immediately aversive.

An open question is whether exercising willpower is, in and of itself, aversive, or whether the aversive aspect of willpower results

---

that the dopaminergic system plays a role in volitional behavior. Parkinson's disease – a condition that is characterized by a progressive disintegration of volitional behavior – arises from a dopamine deficiency and can be temporarily reversed through dopamine enhancing drugs. Electrical stimulation of certain dopaminergic nerve clusters can trigger elaborate trains of actions, such as feeding or grooming behaviors. Although there are many connections between the cortex and other areas of the brain, there is considerable evidence that the prefrontal cortex, and particularly its southern sector, plays an important role in the interface between the cortex and the dopaminergic system. Damage to the prefrontal region produces impulse control disorders (Damasio, 1996).

from the fact that it is used to implement a course of action that has adverse immediate hedonic consequences. On the Alaska lake, for example, paddling resulted in extreme muscle fatigue, which is why we required willpower to continue. Was the muscle fatigue the *only* pain we felt, or did we experience additional pain from the exercise of willpower required to keep paddling? If exertions of willpower are themselves treated as aversive, there would seem to be a risk of double-counting the misery.

When Do Visceral Factors Conflict with Higher' Level Goals?

All visceral factors drive specific behaviors (although more than one visceral factor can motivate the same behavior and any single visceral factor can motivate multiple behaviors). Fear motivates flight, hunger motivates eating, sexual arousal motivates sexual behavior, etc. Visceral factors conflict with higher level goals when an individual experienced visceral motivation to perform (or refrain from performing) a particular action but consciously decides that the action (or inaction) conflicts with more important goals. Hunger, for example, is destructive for a person who is dieting, and intense fear is counterproductive when it interferes with actions that are beneficial or impractical to avoid, such as flying on airplanes or public speaking. In fact, visceral factors often exacerbate the very conditions that given rise to them in the first place. For example, fear of public speaking produces dry-mouth, which can help to fulfill one's fear of a weak performance. In rock climbers, extreme fear produces a "sewing-machine leg" reaction that seems perfectly and perversely designed to ensure that the feared fall occurs. In the climbing example, unlike the case of public speaking, the visceral and cognitive systems are aligned in perceiving a grave risk, but the visceral system causes one's leg to jiggle uncontrollably while the cognitive system recognizes that this is not a good strategy for remaining fixed to the cliff.

Constraints on Willpower

Willpower is not an unlimited resource. As the term implies, willpower involves an exercise of force, so it is not surprising that there are striking similarities between the constraints on willpower

and those on muscular exertion, mental concentration, or any other human activity that requires effort.

Limited momentary strength: Despite occasional demonstrations of phenomenal feats of willpower, there is a limit to the amount of willpower that can be instantaneously exerted in any situation. Most people who haven't slept for several days will fall asleep at the wheel of a car despite a substantial risk of death, and it is now well-established that virtually everyone 'breaks' under torture (e.g., Biderman, 1960).

Limited 'reservoir': The limits of willpower refer not only to instantaneous exertions, but also to the rate of usage. The distinction between the instantaneous and dynamic constraints on willpower is analogous to that which weight lifters make between strength and endurance. Strength is "the ability of a muscle or muscle group to exert one maximal force against a resistance" (Thaxton, 1988, p. 105). Endurance is "the ability of a muscle or muscle group to exert repeated contractions against a resistance for an extended period of time or to maintain an isometric contraction for an extended time period" (Thaxton, 1988, p. 107). Like muscle-power, willpower tends to be depleted if used too intensively over a short period of time, but replenished by nonuse.[2]

Strengthened by long-term exercise: Willpower is reputed to be strengthened by moderate usage.[3] Parents seek to develop their

[2] Suppose that the individual has a 'reservoir' of willpower, W_t, on which she can draw, with $W_t \leq W_{max}$. W_{max} can be thought of as the 'full' level of the willpower reservoir – the level which will prevail if the resource is not drawn down by exertions of willpower. In any period, t, the reservoir is depleted by the amount of willpower that is actually used, w_t, But the reservoir is also replenished – assume for illustrative purposes at a rate that is proportional to the difference between the maximum reservoir level and the current level. Under these assumptions, changes in reservoir level will be dictated by: $W_{t+1} = W_t - w_t + \delta[W_{max} - W_t]$, and the constraint that willpower is limited by the level of the reservoir means that: $W_t \geq 0$ for all t, or $w_t \leq W_t + \delta[W_{max} - W_t]$.

[3] Whether willpower is, in fact, strengthened by long-term usage may also depend on the benefits derived from its exercise. The dieter who met the partner of his or her dreams after months of self-denial is likely to have a much more positive attitude toward willpower than the addict who went 'clean' only to see his marriage break up anyway.

children's willpower through gradually increasing tests of endurance, and people do similar things to themselves in the hope of developing their own willpower capacities. Although the idea that willpower is enhanced through exercise might seem to contradict the limited reservoir assumption, the two effects operate at different time-scales. Again, these effects are analogous to those that apply to muscles, which are weakened momentarily by immediate usage, but strengthened in the long-run by repeated usage.

Undermined by extreme usage: Although long-term moderate exercise probably strengthens the will (i.e., increases the capacity of the willpower reservoir), if excessive demands are made on willpower, instead of being strengthened, the will can be "broken." A common pattern in the mountaineering literature is that, after saving their own lives or those of their companions through extraordinary acts of willpower, mountaineers become completely debilitated – that is, unable to perform the slightest actions on their own behalf. Such a breaking of the will is analogous to tearing a muscle as a result of excessive strain.

Enhanced by psychic preparation: Some degree of preparation for the use of willpower seems to be helpful. In the same way that people prepare themselves for muscle usage, they seem to be able to brace themselves for an effort of willpower. By the same token, the absence of such preparation can disarm willpower. People are often unable to exercise self-control when the impetus for action arises too suddenly and unexpectedly – e.g., to inhibit an angry response when another driver's horn interrupts one from one's commuting reveries.

Fluctuation as a function of exogenous and internal factors: Muscles are less effective in hot weather or when you missed breakfast, and can be immobilized or enervated by fear and other emotions, By the same token, willpower fluctuates as a function of external and internal factors. Many people report that depression undermines their willpower and this may be generally true of negative affect. In fact, willpower seems to be weakened by some of the same factors that undermine muscle-power, such as hunger and sleep deprivation. Some types of drugs – most notably alcohol – also disarm willpower (a point I return to below). Indeed, people some-

times seem to consume alcohol and other drugs to give themselves an excuse to lose self-control.

One reservoir or many?: An open question is whether willpower consists of a single undifferentiated reservoir, or rather a set of reservoirs each of which is applicable to a different activity (e.g., dieting willpower, willpower to resist fear, etc.). It seems likely that the reality lies somewhere in-between these extremes – i.e., that different activities draw on pools of willpower that are partly unique and partly shared.

Willpower versus Other Tactics for Self-control

Many strategies for self-control do not involve willpower. In the existing literature on self-control, some self-control tactics have been included under the heading of willpower that are, in my opinion, more properly classified as substitutes for willpower. This mislabeling has created considerable confusion.

Perhaps the most important self-control tactic is to avoid the types of situations that induce intense visceral responses. Alcoholics stay away from bars, and parties where alcohol is served. People who are prone to angry outbursts may avoid the people they are angry at or the situations that make them angry. Such tactics should not be classified as an exertion of willpower, but rather as tactics intended to reduce the need for willpower – i.e., as substitutes for willpower.

Another self-control tactic involves cognitive transformation of rewards or punishments (Mischel, 1975). The dieter can fantasize about what her life would be like with a thin body. The smoker can conjure up gruesome images of hospitals and lung-surgery. Such tactics can be seen as attempts to combat visceral factors by evoking other, competing, visceral factors. Again, these types of tactics should be interpreted as substitutes for, rather than instances of, willpower.

Other tactics that have been discussed at length include precommitment devices, such as public resolutions or side-bets (e.g., "If I eat dessert tonight I'll pay you $100"). These tactics are again not instances of willpower as I have defined the term. Precommitment devices that introduce disincentives for succumbing to visceral

factors augment willpower in at least four ways. First, and most trivially, some precommitment devices actually eliminate the possibility of committing the act one wants to avoid. For example, antabuse completely precludes alcohol consumption, albeit only for a limited period of time. Second, precommitment devices can introduce new visceral motivations that compete with the visceral factors whose actions one wants to suppress. The loss of $100 if one eats the dessert, for example, would produce a powerful visceral response. Third, precommitment tactics often disarm rationalizations. Thus, one can rationalize that "one dessert can't hurt", but it is more difficult to rationalize that losing $100 can't hurt. Finally, precommitment tactics may actually help to mitigate the visceral factors themselves. Hunger, thirst, sexual arousal, and many other visceral factors are stimulated by the availability of a reward. If precommitment tactics decrease the individual's subjective likelihood of succumbing, they may initiate a kind of self-fulfilling prophecy by reducing the hunger that served as the initial impetus for the indulgence.

Perhaps the trickiest types of tactics to classify are those that focus attention *on the behavior that one is trying to control.* William James (1890, p. 562) viewed willpower as synonymous with such attention focus:

> *Effort of attention is thus the essential phenomenon of will.* Every reader must know by his own experience that this is so, for every reader must have felt some fiery passion's grasp. What constitutes the difficulty for a man laboring under an unwise passion of acting as if the passion were unwise? Certainly there is no physical difficulty. It is as easy physically to avoid a fight as to begin one, to pocket one's money as to squander it on one's cupidities, to walk away from as towards a coquette's door. The difficulty is mental; it is that of getting the idea of the wise action to stay before our minds at all. When any strong emotional state whatever is upon us the tendency is for no images but such as are congruous with it to come up.

It is unclear whether this form of attention focus should be classified under the heading of willpower or instead as a substitute for willpower. My own opinion is that willpower involves both more and less than focus of attention. Focus of attention may be necessary for the exercise of willpower, but it probably is not sufficient. Indeed, in many cases (such as life-guarding on a boring beach), focus of attention itself *requires* willpower.

A Complication: Overestimation of Willpower

Complicating the exercise of willpower is the fact that people seem commonly to overestimate the effectiveness of their own willpower. There are at least two possible reasons for such a mistake. First, people may underestimate the influence that visceral factors will exert on their own future behavior. There is substantial evidence that, when in a viscerally unaroused state (e.g., not hungry, angry, or in pain), people underestimate the motivational force of being in a visceral state (see Loewenstein, O'Donoghue and Rabin, 1998, for a review and theoretical formalization of this effect). If so, they will tend to underestimate the amount of willpower that will be required to control their own behavior. This bias is closely related to human deficiencies in the recall and anticipation of pain that are particularly acute when it comes to decision making (e.g., people may be able to accurately report the amount of pain they experienced in the past, but such reports are poor predictors of willingness to experience the pain again in the future. See Read and Loewenstein, 1999).

Second, people may have an imperfect understanding of the dynamics of willpower. They may not recognize the extent to which their own willpower will be depleted through exertion, or they may overestimate the size of their willpower reservoir. In such cases they might squander willpower in situations where it is futile, or waste large amounts of willpower (e.g., on extreme diets), leaving them weakened in the long-run.

Overestimation of willpower is destructive because people enter into situations with an overblown estimation of their own ability to control their behavior. People enter marathons (or sign up for guided mountain adventures) in the mistaken belief that they can compensate for inadequate training through mighty exertions of willpower. They plan beach vacations, confident in the expectation that they will be able to 'slim down' for summer. And they plan long family get-togethers in the belief that they can suppress their hostile feelings longer than they actually can. In general, overestimating the effectiveness of willpower causes people to overexpose themselves to situations that require willpower and to underutilize alternative self-control tactics.

Willpower as a Decision Variable

Willpower is a resource that can be used to decrease or eliminate discrepancies between viscerally motivated and deliberatively desired behaviors. It is, however, a constrained resource and thus, to be used efficiently, it must be allocated selectively between alternative uses. The allocation of willpower, therefore, is an important decision that people face – one which interacts with other types of decisions in complex and important ways. For example, it makes no sense to "decide" that one is going to quit smoking, diet, or practice safe sex if one does not actually allocate (or possess) the requisite willpower to actually carry through with one's resolution.

The notion of willpower as a decision variable would considerably complicate decision-theoretic analyses of behavior. In the conventional decision-making framework people are assumed to compare different action alternatives and choose between them based on the answer to a single simple question: "What's best for me?" A model that incorporates willpower introduces two new types of questions that would need to be asked routinely: "How will I behave if I don't exercise willpower?", and "Can I and should I exercise the willpower required to bridge the gap (in whole or in part) between what's best and how I will behave if I don't exercise willpower?"

GENERAL IMPLICATIONS AND QUESTIONS

The idea that it takes willpower to overcome the influence of visceral factors, and that willpower is a resource with its own unique dynamic constraints, has numerous implications that are not shared by other models of self-control that have been proposed.

Dynamic limitations on the exercise of willpower: The notion of a limited reservoir of willpower focuses attention on dynamic aspects of self-control. Because exercising willpower draws down one's reservoir of willpower, and because more intense visceral factors require great amounts of willpower to control, we should expect to see an inverse relationship between the intensity of a particular visceral factor and the length of time that its behavioral effects can be suppressed. In general, people may be able to suppress the action

of most visceral factors for brief periods, but there will be severe limitations on the duration of such control. Consistent with this prediction, research on diverse forms of self-control suggests that the action, in terms of explaining variation in success, lies not in the short-term exercise of self-control, but in the longterm maintenance of self-control.

Research on dieting, for example, has shown that it is relatively easy to lose a lot of weight quickly, but extraordinarily difficult to keep it off. The success of a diet, therefore, depends more on the long-run maintenance of willpower than on the dieter's perception of the costs and benefits of dieting (as assumed in most models of self-control). The same is true for adherence to onerous medical treatments such as those involved in diabetes. Numerous studies have found that people can adhere to treatment protocols for limited periods, but that many or most eventually stop complying after some period of time. The situation is similar for exercise. Almost anyone can join an exercise club and work out regularly for a week or two. But exercise clubs bank on the fact that very few people have the willpower to maintain such a regimen. Oliver Sachs has documented analogous limitations on willpower in his accounts of Tourette's syndrome. People afflicted with Tourette's syndrome are often able to suppress their tics for limited periods of time (e.g., during a dinner party or while performing critical task at work), but the price of such suppression is a subsequently intensified outburst.

Effect of factors that disarm willpower: A second implication is that relapses of all types – failures of will – will be especially likely to occur when external or internal factors conspire to weaken willpower. There is actually a substantial amount of research focusing on precisely this point, much of which is reviewed in Baumeister, Heatherton and Tice's (1994) superb book *Losing Control*. Glass, Singer and Friedman (1969), for example, conducted an experiment in the first phase of which subjects performed a simple task under one of three conditions, one which involved exposure to stressful, unpredictable noise. In the second phase they were tested in a quiet room on their capacity to persist at a frustrating anagram task. Consistent with the idea that stress can undermine willpower, subjects who had been exposed to the stressful, unpredictable, noise, quit earlier than other subjects. Striegel-Moore, Silberstein, Frensch and

Rodin (1989), in a study of the eating habits of colleges freshmen, found that stress was associated with a worsening of eating problems for female students. Heatherton, Herman and Polivy (1991) subjected dieters to three types of stressful experiences: fear of electric shock, fear of having to make a speech in front of peers, or failure on a supposedly important task. They were then given free access to ice cream under the guise of participating in a taste-test. Compared to a control group that did not experience stress, speech threat and task failure led to increased eating.

The 'what-the-heck' effect: One important implication of treating willpower as a decision variable is that we should expect people to give up and not attempt to exercise any willpower in situations when they recognize that such an attempt would be futile. Such a what-the-heck effect has, in fact, been commonly observed in research on dieting. Ruderman (1985) found that when dieters expected to eat a highly caloric meal in the future, their resolve collapsed and they began to eat immediately. Tomarken and Kirschenbaum (1984) reported an equivalent finding for both dieters and nondieters. More generally, we should expect willpower to be applied not necessarily when visceral factors are most destructive, but rather when willpower can be applied most efficiently – where maximal benefits can be achieved at minimal cost. Such considerations of efficiency require simultaneous consideration of (1) the potential damage that could be produced by the behavior that is motivated by the visceral factor and (2) the amount of willpower that would be required to suppress the behavior.

Casual observation suggests that people often waste willpower in the sense of trying, but failing, to control their own behavior. People quit drinking, smoking, practicing unsafe sex, not once, but (as W.C. Fields put it in a famous quip) hundreds of times. One explanation for this common phenomenon is that people overestimate their own reserves of willpower and believe that they can maintain control until the craving wears off. An alternative interpretation is that they overestimate, not their willpower, but the speed at which their craving will decrease.

IMPLICATIONS FOR ADDICTION

Recently, a number of economists and psychologists have attempted to express fundamental features of addiction in decision-theoretic models. Addiction poses a special fascination to rational choice theorists as a result of its seemingly self-destructive aspect which, on the face of it, seems to fly in the face of the idea that people are making decisions based on their long-term self interest.

Virtually all decision-theoretic models of addiction assume that taking an addictive drug has two major consequences for future preferences: it increases one's subsequent desire for the drug (habituation) and decreases the utility associated with not taking the drug (dependence). Most models also assume that ever-larger-quantities of the drug are required to have the same impact on utility (tolerance). The crucial difference between the various models that have been proposed lies in their explanations for why people get addicted – why they put themselves into a position that seems patently self-destructive.

Gary Becker and Kevin Murphy's "model of rational addiction" views addiction as a form of deliberate self-medication. According to their account, people may recognize that addiction has negative long-term consequences but judge that the benefits (discounted according to when they occur) outweigh the costs. Orphanides and Zervos (1995) elaborate on Becker and Murphy by introducing the idea that people differ in their susceptibility to addiction but are unaware of their own 'type'. For Orphanides and Zervos, therefore, addiction is a rational gamble that some people lose. Herrnstein and Prelec (1992; see, also, Heyman, 1996) argue that addiction results from people's lack of awareness of the impacts of current drug-taking on future preferences – i.e., dependence and habituation – which they refer to as "internalities." O'Donoghue and Rabin (1997) argue that addiction arises from a combination of hyperbolic time discounting (which places disproportionate weight on immediate gratifications) and naivety – people's false belief that, though they may take a drug today, they will be able to desist in the future.

In a separate paper (Loewenstein, 1998) I have argued that drug addiction is a virtually paradigmatic example of viscerally-driven behavior. Although there is considerable debate on the issue, many contemporary accounts of addiction view it as driven by cue-

conditioned craving. Drug craving shares the two hallmark features of all visceral factors: it is aversive and it increases the attractiveness of drug-taking. In fact, at high levels drug craving, like other visceral factors, can take virtually complete control over behavior. Deferring again to William James (1890, p. 543):

> The craving for a drink in real dipsomaniacs, or for opium or chloral in those subjugated, is of a strength of which normal persons can form no conception. "Were a keg of rum in one comer of a room and were a cannon constantly discharging balls between me and it, I could not refrain from passing before that cannon in order to get the rum;" "If a bottle of brandy stood at one hand and the pit of hell yawned at the other, and I were convinced that I should be pushed in as sure as I took one glass, I could not refrain:" such statements abound in dipsomaniacs' mouths.

Cue-conditioned drug craving is somewhat different from other visceral factors in that it is an acquired rather than innate response. People are born with hunger, fear, anger, and the sex drive 'hardwired.' They may also be hardwired for the capacity to become addicted, but the capacity will remain latent unless the individual actually indulges. Drugs have an additional pernicious quality in relation to willpower. Many forms of addictive substances also undermine rationality by undermining will (e.g., decreasing sensitivity to punishment or narrowing focus of attention). Steele and Southwick (1985) found that alcohol has disinhibiting effects when there is an inhibitory response conflict – that is, when the person has both a desire for and an inhibition against a particular activity. By reducing self-awareness, alcohol removes the tendency to compare oneself or one's acts against norms and standards, so the inhibiting anxiety of guilt is not felt.

Willpower-related Features of Addiction

Adding willpower to theoretical accounts of addiction has the capacity to explain a wide range of phenomena that existing models are unable to address.

Importance of relapse: Some theoretical accounts (e.g., Koob et al., 1989; Solomon and Corbit, 1974) have identified withdrawal – the state of anhedonia that often begins shortly following the termination of drug use – as the major factor underlying drug-

dependence. Currently, however, there is a virtual consensus that withdrawal does not constitute the major impediment to quitting, in part as a result of the widespread availability of effective therapeutic interventions that ease the misery of withdrawal (Wise and Bozarth, 1987). Instead, the main impediment to quitting appears to be the long-term problem of *craving-induced relapse*. "During withdrawal", Gawin (1988, p. 12) comments, "most cocaine abusers can withstand postcocaine anhedonia." However, "after the period ends, episodic craving and the risk for relapse remain because of the continued role of conditioned cues." Relapse is a constant threat because craving can be initiated by almost any environmental cue that becomes associated with the drug – for example, time of day, a particular room or even the color of the room, the presence of specific individuals or paraphernalia associated with drug taking (Siegel, Krank and Hinson, 1988), sounds, and even positive or negative mood states (Gawin, 1991, p. 1582). Deconditioning – the gradual diminishment in a cue's propensity to evoke craving – is a slow process; cues may retain their ability to evoke craving even after years of abstinence (Niaura et al., 1988; Shiffinan, 1982). The constant vigilance required to avoid such cues, and to resist craving when it does occur, progressively undermines willpower, which may help to explain the distressingly low long-term abstinence rates among one-time drug addicts (Hser, Anglin and Powers, 1993).[4]

Power of weak craving: A paradox that receives periodic attention in the addiction literature is the discrepancy between the extreme motivational power, yet mild subjective aversiveness, of drug craving. Tiffany (1998), for example, reports that, contrary to the horrific popular image of heroin withdrawal, many heroin addicts report that withdrawal feels not much worse than a bad cold. Tiffany argues

[4] Long-term abstinence seems to require not only willpower, but the successful implementation of strategies that substitute for willpower. Thus, for example, successful quitting often requires a substantial investment in change of environment and lifestyle because addiction "poisons" persons, places and things associated with it in the sense of imparting them with the ability to induce craving. As Siegel (1982, p. 335) observes, "users will attempt to avoid all contact with cocaine, cocaine paraphernalia and cocaine users when attempting this self-initiated detoxification. Others engage in destruction of paraphernalia, and still others employ physical restraint by taking a vacation or even moving to another house or city."

that addicts don't experience intense craving because they typically ingest the drug they are addicted to long before craving sets in. Berridge and Robinson (1995), in contrast, argue that the subjective feeling of craving understates its motivational force because the neural systems that determine 'wanting' (motivation to take the drug) operate somewhat independently of those that determine 'liking' (the pleasure or pain diminishment that would result from taking the drug).

The concept of willpower, and the dynamic constraints associated with it, provides another possible explanation for the paradoxically strong effect of weak craving. Craving may be weak, but if the expenditure of willpower required to resist taking the drug exceeds the replenishment rate of the willpower reservoir, then it will nevertheless undermine willpower – albeit at a relatively slow rate. At any moment, craving may not be experienced as all that bad, but the individual may be unable to resist its cumulative effects, just as holding even a small weight perpendicular to your body is difficult for a prolonged period of time.

Importance of surprise: If willpower takes time to mobilize, as suggested in the subsection on constraints, then sudden, unpredicted, spikes of craving are less likely to be resisted than those that are anticipated beforehand. This is likely to be a significant problem, since craving is as unpredictable as the cues which elicit it. As O'brien et al. (1988, p. 18) write, addicts are often "surprised to suddenly feel craving, withdrawal, or even "high" when they encounter people or places associated with their prior drug use." Acker (1998), in an analysis of transcripts from interviews with opiate addicts conducted in the 1920s, similarly found that addicts often pointed to the element of surprise – as in unexpected meetings with former drug-using friends – in explaining their own relapses. The implication is that addicts who are attempting to desist from drug use should formulate plans for coping with craving in advance of actually experiencing it. As Washton (1988, p. 35) writes: "For each patient, a broad range of potential high-risk situations, including people, places, and/or things formerly associated with drug use, must be identified. Plans should be formulated in advance for successful avoidance of and coping with these situations."

Bingeing behavior: Contrary to stereotype of addicts as impulsive and short-sighted, some are actually quite deliberate and forward-thinking in their drug-taking behavior. Although probably a minority, some addicts prepare long in advance for their next binge. During the summer after graduating from college I lived in a relatively undeveloped area of Puget Sound (in Washington State). My nearest neighbor on the Sound was a young man who lived in a houseboat one cove down from mine, who was reputed to be an alcoholic, though over many visits I failed to observe him drinking. At some point during the summer, however, he began to stockpile vast quantities of beer, and told me that he was planning to go on a bender. The bender, when it finally took place, lasted about a week, during which he alternated between wildly expressive behavior and virtual catatonia. I had the sense that, during the binge, he was relieving some type of pressure that had been accumulating during his period of abstinence. Eliot Gardner describes a similar pattern of behavior by cocaine users, who save up money until they have enough for a prolonged high, then lock themselves into their apartments to remain high until their stockpiled cocaine runs out. Behavioral patterns of this type are difficult to reconcile with any of the existing decision-theoretic models of addiction, but could easily be accounted for by a model that incorporated the dynamic constraints on willpower.

Implications for Drug Policy

Traditional decision-making accounts of addiction imply that to change an addict's behavior it is sufficient to change the incentives for drug use and abstinence. This is no longer the case once willpower is introduced into the decision making equation. If willpower is required to execute desired courses of action, then incentives may change people's desire to take drugs without affecting actual drug-taking behavior. Thus, drug policies should be evaluated not only on the basis of their effect on incentives, but also in terms of how they interact with willpower. Policies that augment willpower alone can be sufficient to change behavior if drug-takers would already prefer to stop taking drugs, but policies that change incentives will have little impact if people don't have the willpower to execute the behavioral changes that they motivate. In this light the cur-

rently dominant policy strategy of attempting to deter drug-taking through delayed sanctions such as incarceration would seem to be seriously flawed. Delayed threats associated with drug taking may change people's desire to quit, but does not facilitate their efforts to quit (except to the degree that it produces an immediate visceral fear-reaction to drug-taking).

Immediacy of rewards and penalties: Policies that introduce very immediate but milder short-term sanctions for drug-use or rewards for desistance seem slightly more promising. The "enforced abstinence" policy which is being tested in various states subjects addicts who are on parole to frequent drug-tests and penalizes a finding of drug-taking with immediate incarceration (Kleiman, 1997). Punishment is, therefore, both immediate and close to certain. On the reward side, Stephen Higgins and his collaborators (1994, 1995) have demonstrated the efficacy of 'bribing' addicts to desist by presenting them daily with monetary rewards for desisting, which begin very small but increase in magnitude as a function of how long the addict has been 'clean'. The increasing size of the rewards might seem perverse. If one takes the view that abstinence should get easier as withdrawal craving wears off and cues become deconditioned one might think that it would be more effective to begin with large monetary rewards for abstinence that would shrink over time. One interpretation of the effectiveness of increasing rewards is that, if willpower is a limited resource that gets drawn down by usage, it may actually become increasingly difficult for an addict to resist drug-taking, at least for some interval of time.

Mitigation of craving: To the extent that they can be devised, however, one would expect that the most effective policies would be those that reduce the need for willpower by reducing or eliminating craving. Schelling, for example, reports that addicts experience substantially reduced craving when they reside in treatment programs that have a reputation for inviolability – for nonavailability of drugs. However, living in a closed drug-treatment setting is not a long-term solution for most people. A radical change in environment, it appears, can also substantially reduce craving by eliminating cues commonly associated with drug-taking. It is legion among drug researchers that the vast majority of the huge numbers of

soldiers who were addicted to heroin in Vietnam kicked the habit upon returning to the United States. By the same token, returning addicts to their home-neighborhoods where drugs are easily available and drug-taking partners are a constant presence seems virtually guaranteed to lead to reinstatement of the addiction.[5]

Importance of willpower dynamics: Whatever the dynamic constraints on willpower, it would be very important for people who are trying to exercise willpower to understand the nature of the constraints that they face. When people begin dieting, practicing safe sex, or trying to kick a drug habit they often go all the way – virtually starving themselves, abstaining from sex, or going cold turkey. If willpower were an unlimited resource then these types of actions might make sense, but if willpower is in limited supply then there is a danger that people will effectively "blow their wad" with such extreme measures, leaving their willpower depleted when it comes to long-term adherence. The fact that 'controlled drinking' seems to be an effective strategy for some alcoholics, contrary to the emotional claims of abstinence advocates, could be due to the willpower-depleting effects of complete abstinence (see, Sobell and Sobell, 1973; Pendery, Maltzman and West, 1982; Sobell and Sobell, 1989). If further study yields reliable insights about the dynamic constraints guiding willpower, then it might be possible for addicts and their counselors to devise more effective strategies for husbanding and utilizing willpower, much in the same way that marathon runners strategize about pacing themselves.

Phases of addiction: Well thought-out policies aimed at drug addiction must confront the fact that drug users behave very differently at different stages of addiction. The most important differences may revolve around the role played by willpower. In the early stages of addiction – i.e., in the drug-taking that leads to addiction – the decision to take or desist from drug-taking may not involve much willpower. People try drugs for a variety of reasons: curiosity, peer influences, self-medication, etc. Before they are addicted it is probably reasonable to say that they take drugs because they want to.

[5] This implication for treatment is shared with a wide range of models of addiction.

The decision may be biased, however, by underestimation of the force of the craving they will experience in the future (Loewenstein, 1998), or by overestimation of the effectiveness of willpower. As an individual becomes addicted to a drug, however, there is a progressive loss of volitional control over drug taking, and desisting requires an ever-greater exercise of willpower. Once addicted, an individual may *recognize* that abstinence is the best course of action, but his ability to abstain is powerfully constrained by limitations on willpower and its substitutes.

Such a transition in the character of decision making suggests that different policies should be aimed at drug-users situated at different points in this transition. In the early, pre-addiction, phase of drug-taking, strategies that alter costs and benefits (including the price of the drug) may effectively change behavior. If decision making is biased by an imperfect appreciation of the risk of addiction, moreover, there would seem to be an important potential role for educational programs that attempt to reduce this bias. Once the individual is addicted, however, the addiction is probably best viewed as a disease that requires treatment.

WILLPOWER AND RESPONSIBILITY

The economic approach to law is premised on the validity of the standard decision-making perspective according to which people choose, and effortlessly implement, courses of action based on assessments of costs and benefits. Perceptions of costs and benefits are, however, only part of the story. Behavior often requires the extra ingredient of willpower. People may be able to determine what is in their self-interest but implementing this ideal is not necessarily automatic. When perceptions of self-interest conflict with the motivational impetus of visceral factors, willpower is required to behave according to self-interest. Legal sanctions may thus influence assessments of self-interest, but will not have much impact if the requisite willpower is not forthcoming.

If valid, such a perspective might point to alternative policy strategies, not only for the control of drug use, but for crime-reduction more generally. Rather than trying to deter criminal conduct by raising the sanctions associated with it, it may be helpful

to formulate strategies to help people to reduce the power of visceral factors or to mobilize and husband their own willpower resources. There are already many programs in place to help schoolchildren deal with self-control problems, but these are rare among adults – even those with a history of violence.

Exaggeration of the importance of premeditation: Criminal law already acknowledges, to some degree, the importance of willpower in the distinction made between premeditated (first degree) and unpremeditated (second degree) murder. While this is perhaps a step in the right direction, the existing law seems to me to be premised on a naive view of volition and willpower. The logic behind the distinction between first and second degree murder is presumably that unpremeditated murders are more likely to involve rapid, mindless, reflexive reactions to sudden passions. As noted earlier in the paper, rapid and unexpected changes in visceral factors do indeed make greater demands on willpower. But surprise is only one of many factors that affect the strength and effectiveness of willpower, and is probably not the most important one. In fact, ongoing low-level visceral factors, such as nagging jealousy or simmering resentment, can undermine self-control just as much as sudden flashes of passion. If one wants to punish crimes differently as a function of the extent of the criminal's volition, then premeditated/unpremeditated may not be the distinction one would want to make.

Credit and Responsibility: The notion of willpower can help to shed light on how we naturally tend to confer credit and cast responsibility. In general, actions that do not require willpower tend to get less credit than those that do. If a genius rapidly solves a difficult problem, or if an athlete easily performs an impressive feat, people will express admiration. But the greatest credit is usually reserved for significant feats of willpower. Michael Jordan's fame, already great, achieved new heights when he made the winning points in a playoff game while suffering from a debilitating flu. As with any desirable trait, people are curious about their own powers of will and anxious to prove to themselves and others that they are substantial. In a recent paper titled "The Challenge of Mountaineering . . . For Decision Theory," I grapple with the age-old question of why people

climb mountains and engage in arctic exploration. The most important reason, I conclude, is to test themselves. What is being tested, however, is not one's skill with ropes and ice-ax, but one's mettle – the strength of one's will. Whereas daily life provides few compelling tests of willpower, mountaineering and polar explanation strain even the strongest wills to the limit. As Lansing (1959, p. 13) writes of the famous antarctic explorer, "Shackleton's tremendous capacity for boldness and daring found almost nothing worthy of its pulling power, he was a Percheron draft horse harnessed to a child's wagon cart. But in the Antarctic – here was a burden which challenged every atom of his strength." Polar explorer Robert Scott commented in his Antarctic diary (cited in Cherry-Gerrard, 1922, p 1xiii), "I do not think there can be any life quite so demonstrative of character as that we had on these expeditions ... Here the outward show is nothing ... Pretence is useless." The fact that mountaineering and arctic exploration are done, in part, to test or demonstrate willpower helps to explain why the most miserable trips often produce the best memories; pain and discomfort are, to some degree, the point of the trip.

At the other end of the spectrum, people are not blamed for undesirable actions, or blamed much less, when resisting the action would seem to call upon unrealistic amounts of willpower. People are probably, nevertheless, too ready to blame others, because they tend to overestimate not only their own, but also others' powers of will.

FINAL COMMENTS

For reasons discussed in this paper, willpower greatly complicates the task of modeling human behavior. Because many types of decisions do not require much willpower, decision-theorists are justified in avoiding its complexities in many or even most of their behavioral analyses. When visceral factors propel behavior in directions that are not commensurate with self-interest, however, i.e., when strong emotions, drives, and somatic sensations are in play, decision making models that do not incorporate willpower will be fatally incomplete. They predict how people wish they could behave rather than how they actually behave.

REFERENCES

Acker, C. J., Mirrors of Morphine: Scientific Constructions of Opiate Addiction in the 1920s (Department of History, Carnegie Mellon University, 1998) (working paper).

Baumeister, R. F., Heatherton, T. F. and Tice, D. M., *Losing Control: How and Why People Fail at Self-Regulation* (San Diego: Academic Press, 1994).

Becker, G. and Murphy, K., 'A Theory of Rational Addiction', *Journal of Political Economy* 96 (1988): 675–700.

Berridge, K. C. and Robinson, T. E., 'The Mind of an Addicted Brain: Neural Sensitization of Wanting Versus Liking', *Current Directions in Psychological Science* 4 (1995): 71–76.

Biderman, A. D., 'Social-psychological Needs and 'Involuntary' Behavior as Illustrated by Compliance in Interrogation', *Sociometry* 23 (1960): 120–147.

Cherry-Gerrard, A., *The Worst Journey in the World* (London: Constable, 1922).

Gawin, F. H., 'Chronic Neuropharmacology of Cocaine: Progress in Pharmacotherapy', *Journal of Clinical Psychiatry* 49 (1988): 11–16.

Gawin, F. H., 'Cocaine Addiction: Psychology and Neurophysiology', *Science* 251 (1991): 1580–1586.

Glass, D. C., Singer, J. E. and Friedman, L. N., 'Psychic Cost of Adaptation to an Environmental Stressor', *Journal of Personality and Social Psychology* 12 (1969): 200–210.

Heatherton, T. F., Herman, C. P. and Polivy, I., 'Effects of Physical Threat and Ego Threat on Eating', *Journal of Personality and Social Psychology* (1991): 138–143.

Hernstein, R. and Prelec, D., 'A Theory of Addiction', in G. Loewenstein and J. Elster (eds.), *Choice Over Time* (New York: Russell Sage Foundation Press, 1992).

Heyman, G. M., 'Resolving the Contradictions of Addiction', *Behavioral and Brain Sciences* 19 (1996): 561–610.

Higgins, S. T., Budney, A. J., Bickel, W. K., Badger, G. J., Foerg, F. E. and Ogden, D., 'Outpatient Behavioral Treatment for Cocaine Dependence: One-year Outcome', *Experimental and Clinical Psychopharmacology* 3 (1995): 205–212.

Higgins, S. T., Budney, A. J., Bickel, W. K., Foerg, F. E., Donham, R. and Badger, G. J., 'Incentives Improve Outcome in Outpatient Behavioral Treatment of Cocaine Dependence', *Archives of General Psychiatry* 51 (1994): 568–576.

Kleiman, M., 'Coerced Abstinence: A Neo-Paternalistic Drug Policy Initiative', in L. A. Mead (ed.), *The New Paternalism* (Washington, D.C.: Brookings Institute, 1997).

Koob, G. F., Stinus, L., Le Moal, M. and Bloom, F. E., 'Opponent Process Theory of Motivation: Neurobiological Evidence from Studies of Opiate Dependence', *Neuroscience and Biobehavioral Reviews* 13 (1989): 135–140.

Lansing, A., *Endurance: Shackleton's Incredible Voyage* (New York: Carroll & Graf, 1959).

Loewenstein, G., 'Out of Control: Visceral Influences on Behavior', *Organizational Behavior and Human Decision Processes* 65 (1996): 272–292.

Loewenstein, G., 'A Visceral Account of Addiction', in J. Elster and Ole J. Skog (eds.), *Getting Hooked: Rationality and Addiction* (Cambridge: Cambridge University Press) (1999).

Loewenstein, G., 'Because it is There: The Challenge of Mountaineering . . . For Utility Theory', *Kyklos* 52 (1999): 315–344.

Loewenstein, G., Nagin, D. and Paternoster, R., 'The Effect of Sexual Arousal on Predictions of Sexual Forcefulness', *Journal of Research in Crime and Delinquence* 3–4 (1997): 443–473.

Loewenstein, G., O'Donoghue, T. and Rabin, M., 'Projection Bias in Predicting Future Utility' (Department of Social and Decision Sciences, Carnegie Mellon University, 1998) (working paper).

Loewenstein, G., Prelec, D. and Shatto, C., 'Hot/Cold Intrapersonal Empathy Gaps and the Prediction of Curiosity.'

Loewenstein, G. and D. Schkade, 'Wouldn't it be Nice? Predicting Future Feelings', in D. Kahneman, E. Diener and N. Schwartz (eds.), *Well-being: The Foundations of Hedonic Psychology* (New York: Russell Sage Foundation Press) (1999).

Niaura, R. S., Rohsenow, D. J., Binkoff, J. A., Monti, P. M. Pedraza, M. and Abrams, D. B., 'Relevance of Cue Reactivity to Understanding Alcohol and Smoking Relapse', *Journal of Abnormal Psychology* 97 (1988): 133–152.

O'Brien, C. P., Childress, A. R., Arndt, A. T. and McLellan, G. E., Woody and Maany, 'Pharmacological and Behavioral Treatments of Cocaine Dependence: Controlled Studies', *Journal of Clinical Psychiatry* 49 (1988): 17–22.

O'Donoghue, T. and Rabin, M., 'Addiction and Self Control' (Department of Economics, Cornell University, 1997) (working paper).

Orphanides, A. and Zervos, D., 'Rational Addiction With Learning and Regret', *Journal of Political Economy* 103 (1995): 739–758.

Pendery, M. L., Maltzman, I. M. and West, L. J., 'Controlled Drinking by Alcoholics? New Findings and a Reevaluation of a Major Affirmative Study', *Science* 217 (1982): 169–174.

Read, D. and Loewenstein, G., 'Enduring Pain for Money: Decisions Based on the Perception of Memory of Pain', *Journal of Behavioral Decision Making* 12(1) (1999): 1–17.

Ruderman, A. T., 'Dietary Restraint: A Theoretical and Empirical Review', *Psychological Bulletin* 99 (1986): 247–262.

Shiffman, S., 'Relapse Following Smoking Cessation: A Situational Analysis', *Journal of Consulting and Clinical Psychology* 50 (1982): 71–86.

Siegel, R. K., 'Cocaine Free Base Abuse: A New Smoking Disorder', *Journal of Psychoactive Drugs* 14 (1982): 321–337.

Siegel, S., Krank, M. D and Hinson, R. E., 'Anticipation of Pharmacological and Nonpharmacological Events: Classical Conditioning and Addictive Behavior', in S. Peele (ed.), *Visions of Addiction* (Lexington, Mass.: Lexington Books, 1988).

Sobell, M. B. and Sobell, L. C., 'Individualized Behavior Therapy For Alcoholics', *Behavior Therapy* 4 (1973): 49–72.
Sobell, M. B. and Sobell, L. C., 'Moratorium on Maltzman: An Appeal to Reason', *Journal of Studies on Alcohol* 50 (1989): 473–480.
Solomon, R. L. and Corbit, J. D., 'An Opponent-Process Theory of Motivation', *Psychological Review* 81 (1974): 158–171.
Steele, C. M. and Southwick, L., 'Alcohol and Social Behavior: The Psychology of Drunken Excess', *Journal of Personality and Social Psychology* 48 (1985): 18–34.
Striegel-Moore, R. H., Silberstein, L. R., Frensch, P. I. and Rodin, J., 'A Prospective Study of Disordered Eating Among College Students', *International Journal of Eating Disorders* 8 (1989): 499–509.
Thaxton, N., *Pathways to Fitness* (New York: Harper and Row, 1988).
Tiffany, S. T. and Carter, B. L., 'Is Craving the Source of Compulsive Drug Use?' *Journal of Psychopharmacology* 12 (1998): 23–30.
Tomarken, A. J. and Kirschenbaum, D. S., 'Effects of Plans for Future Meals on Counterregulatory Eating: Where Have All the Unrestrained Eaters Gone?' *Journal of Abnormal Psychology* 93 (1984): 458–472.
Washton, A. M., 'Preventing Relapse to Cocaine', *Journal of Clinical Psychiatry* 49 (1988): 34–38.
Wise, R. A. and Bozarth, M. A., 'Psychomotor Stimulant Theory of Addiction', *Psychological Review* 94 (1987): 469–492.

Department of Social and Decision Sciences
Carnegie Mellon University
Pittsburgh, PA 15213
U.S.A.
(*E-mail: gl20@andrew.cmu.edu*)