A Theory of Natural Addiction*

Trenton G. Smith†
University of California, Los Angeles

Attila Tasnádi‡
Budapest University
of
Economic Sciences and Public Administration

December 18, 2003

Abstract

Economic theories of rational addiction aim to describe consumer behavior in the presence of habit-forming goods. We provide a biological foundation for this body of work by formally specifying conditions under which it is *optimal* to form a habit. We demonstrate the empirical validity of our thesis with an in-depth review and synthesis of the biomedical literature concerning the action of opiates in the mammalian brain and their effects on behavior. Our results lend credence to many of the unconventional behavioral assumptions employed by theories of rational addiction, including adjacent complementarity and the importance of cues, attention, and self-control in determining the behavior of addicts. Our approach suggests, however, that addiction is "harmful" only when the addict fails to implement the optimal solution. We offer evidence for the special case of the opiates that harmful addiction is the manifestation of a mismatch between behavioral algorithms encoded in the human genome and the expanded menu of choices–generated for example, by advances in drug delivery technology–faced by consumers in the modern world.

Key Words: endogenous opioids, sugar addiction, behavioral ecology, neuroendocrinology, autism

*The authors would like to thank Ted Bergstrom, Dan Fessler, Ted Frech, Mark Kleiman, Albrecht Morgenstern, Georg Nöldeke, Athanasios Orphanides, Bob Schillberg, Chris Stoddard; and seminar participants at the Bonn Graduate School of Economics and the UCLA Center for Behavior, Evolution, and Culture for helpful comments and discussions. Financial support from the Deutsche Forschungsgemeinschaft, Graduiertenkolleg (DFG) 629 at the University of Bonn and from the Global Fellows Program of the International Institute at the University of California, Los Angeles is gratefully acknowledged.

†UCLA International Institute, 11343 Bunche Hall, Box 951487, Los Angeles, CA 90095-1487 (e-mail: trent@international.ucla.edu)

‡Department of Mathematics, Budapest University of Economic Sciences and Public Administration, Fővám tér 8, 1093 Budapest, Hungary (e-mail: attila.tasnadi@math.bke.hu)

1 Introduction

The immature seed pod of *papaver somniferum* (i.e., the opium poppy) contains a bitter, milky sap. Even in this, its most natural form, opium is a powerful drug, a stimulant narcotic poison that can induce hallucinations, profound sleep, or death. Reduced to its most sought-after chemical constituent, morphine, or further processed into heroin, opium is highly addictive and can have dramatic effects on the behavior, health, and well-being of its users. Opium's natural and synthetic derivatives (collectively known as the *opiates*) have well-known effects on human physiology and behavior: once they make their way into the bloodstream, opiates reliably induce a state of euphoria and pain relief, often followed by an increase in food consumption [45], [72], [75]. Many who experience this state of mind find it pleasurable, and are inclined to try it again. But chronic use of opiates can result in severely impaired health,[1] and desperate addicts sometimes resort to theft or prostitution to obtain money to sustain the habit [78]. Given the potentially lamentable personal and social consequences of drug addiction (and the undeniable fact that legal restrictions have not been fully effective in eliminating drugs like heroin from the streets), many would agree that modern society would be much improved if our species could somehow rid itself of this particular human weakness.

Though the effects of opiates have been known to man for more than five millennia [16], only in recent decades has modern science made clear that opiate-like substances are also produced naturally in the bodies of humans and other mammals. These substances are known collectively as the *endogenous opioids*, and, like their poppy-derived counterparts, they have been shown to induce euphoria, pain relief, and appetite stimulation [15], [74], [118], [128].

The similarity of opiates and the endogenous opioids might seem something of a curiosity at first blush. Given the dramatic negative effects of opiates, what business do our bodies have producing their chemical cousins? There are, fortunately, many ways to answer this question, as the scientific literature is now replete with evidence demonstrating the circumstances under which our bodies produce endogenous opioids, the distribution of and variation in the endogenous opioid system across species, speculation about their evolutionary origins, and even confirmation that the biochemical "recipe" for endogenous opioids is firmly–and apparently universally–encoded in the human genome. This essay will attempt to identify circumstances under which a tendency to become "addicted" might serve a useful function, review supporting evidence from the biomedical literature, and ask what our findings might tell us about drug addiction. In other words, we will develop a theory of natural addiction.

[1] The medical complications of chronic heroin use, for example, can include fetal death, scarred and/or collapsed veins, bacterial infections of the blood vessels and heart valves, abscesses (boils) and other soft tissue infections, disease of the liver or kidney, pneumonia, and tuberculosis. Death from overdose is not uncommon [78].

2 Background

2.1 Rational Addiction

A major source of inspiration for this investigation, and therefore a reasonable starting point for this essay, has been the rich body of theoretical and empirical work on addiction within the economics literature. This literature of *rational addiction* employs the formal mathematical tools of the economist in modeling addiction as a well-defined decision problem to be solved by an optimizing consumer. This approach allows for–and indeed, to some extent requires–the precise statement of the properties of the decision environment that generate addiction. It also allows for the application of the standard tools of welfare analysis in developing implications for drug policy.

The essential feature of most theories of rational addiction is the concept of *adjacent complementarity*, first employed in this context by Becker and Murphy in 1988 [8]. Adjacent complementarity requires that consumption of an addictive good today generates even more consumption of that good tomorrow–or more precisely, that the marginal utility of consumption increases with experience. This property is more general than the popular conception of addiction, of course, and–as Becker and Murphy emphasize–could be used to describe any consumptive behavior in which habits are formed.

The work of Becker and Murphy is notable for its bold assertion that the decision to consume addictive substances is indeed a decision, and as such it can be viewed as a rational decision in a standard economic framework: to be sure, the argument goes, there may be negative personal consequences stemming from addiction, but the fact that many people nevertheless choose to consume addictive substances suggests that–for these people–the benefits of addiction must outweigh the costs. From this beginning, behavioral implications such as the responsiveness of addicts (or potential addicts) to drug prices and criminal penalties, or the dynamics of addiction (e.g., why some people might choose to quit "cold turkey") can be derived. Indeed, the model offered by Becker and Murphy does seem to capture many aspects of the behavior of addicts, and its main empirical prediction (that announced increases in the *future* price of addictive goods should decrease *current* consumption) has been largely borne out in subsequent analysis (see, e.g., [7], [47], and [48]).

In spite of the success of the Becker-Murphy theory of rational addiction, several authors have subsequently noted that in many respects the particulars of Becker-Murphy are not consistent with what is known about the psychology of addiction and the subjective experience of addicts. It has been suggested, for example, that rather than the world of perfect information, foresight, and self-knowledge implicit in Becker-Murphy, addicts face uncertainty regarding the future consequences of addiction [83], [84], may have problems with self-control [34], [48], [49], [82], and may be influenced by emotional or psychological states [11], [63], [67]. There is clearly some truth in each of these critiques, but all of these authors continue to take as given the primitive behavioral property responsible for

addiction: adjacent complementarity.[2] In what follows we take a step back from this descriptive approach and ask under what circumstances habit formation of the type implied by the theory of rational addiction might be optimal. In particular, given the universality of the brain chemistry that makes our species and many others susceptible to drug addiction, we ask under what *natural* conditions the quirky behavioral property known as adjacent complementarity might have arisen. It is our hope in doing so that a more parsimonious and richly descriptive theory will result.

2.2 A Few Words on Biological Foundations

Our aim, to be more explicit, is to identify circumstances in the evolutionary history of the human species in which addiction-like behavior was *optimal* in a well-defined sense.[3] Those who doubt the utility of a biological perspective in the study of human behavior would do well to heed the lessons of John Garcia, who revolutionized behavioral psychology in the 1960s with a series of laboratory experiments that upset the conventional wisdom in the field of operant learning. At the time, a major effort was underway to develop and refine a general theory of learning in the tradition of Pavlov and Skinner, and the standard method of hypothesis testing was to present a laboratory animal (the most thoroughly investigated of which was the white Norway rat) with opportunities to "learn" by pairing an arbitrary stimulus (e.g., a flashing light, or a ringing bell) with an arbitrary reward (e.g., food pellets, or a sweetened solution). What Garcia and his co-investigators showed was that the results of these experiments (i.e., the extent to which an animal learns) are very much dependent on the nature of the stimulus/reward pair employed. In the classic experiment, four groups of (genetically identical) Norway rats were given water, presented with a warning signal, then punished. The classical prediction in this situation is that the subjects will learn to associate the punishment with the ingestion of water, and eventually to avoid punishment by stopping ingestion as soon as the warning signal is observed. In the first two groups, this worked beautifully: the first group was presented with a noise-and-light signal followed by an electric shock; the second was given distinctively-flavored water followed by an artificially (X-ray) induced nausea; and both groups learned to avoid water whenever the warning signal (whether audiovisual or gustatory) was observed. The surprise came with the second two groups: in these groups, the signals and punishments were switched, with one group being presented with the noise-and-light warning followed by artificially

[2]Though we emphasize complementarity here, precise formulations vary. Gul and Pesendorfer, for example, define a good to be addictive if past consumption makes a person more prone to over-consume the good in the future [49]; Bernheim and Rangel [11] define an addictive good as one for which past consumption enables neutral cues to trigger a "hot state" that can lead a person to consume the good again even if additional consumption is not in his best interest; and Orphanides and Zervos [84] generate intertemporal complementarity by allowing addictive goods to alter the time preferences of consumers. Nevertheless, these authors all rely on strong–or at least unconventional–primitive assumptions regarding the behavior of addicts.

[3]Several authors have argued that knowledge of human evolutionary history might help to inform economic theory. See, for example, Hirshleifer [52], [53], Rogers [94], Bergstrom [10], Robson [93], Smith [105], and Samuelson [97].

induced nausea, and the other being presented with distinctively flavored water followed by an electric shock. Here is the kicker: these two groups did not learn [38]. Though Garcia's results were greeted with incredulity by his fellow practitioners–the paper was rejected by the leading journals in the field, and one scientific peer publicly declared that "Those findings are no more likely than birdshit in a cuckoo clock"[4]–they were later replicated many times and in many different species, including humans.[5] The anomaly Garcia identified may have frustrated the search for a general theory of learning, but it came as no surprise to biologists. In the evolutionary history of foraging animals (including the Norway rat), there is every reason to believe that an individual with the ability to associate unusual tastes with subsequent internal malaise–or the ability to associate loud noises with physical external pain–would have had a selective advantage. One who behaved, on the other hand, as though it were possible that loud noises could make him ill, or as though he believed strange tastes could be the cause of sharp pain in his forepaws, might not be so likely to survive. In light of the revolution inspired by Garcia's work, learning theorists now refer to "prepared learning,"–where the "preparation" for learning is implicitly written into our genes by the historical process of natural selection–and it is now accepted that no behavior can be viewed as the exclusive product of either "learning" or "instinct".

Now we would like to apply this perspective to the phenomenon of addiction. One way of going about this would be to search for examples in which addiction-like behavior is exhibited by animals in their natural habitat, and then to identify reasons why–in *natural* settings–such behavior might have given its practitioners an edge, over the ages, in the currency of survival and reproduction. Hypotheses thus arrived at would then be subject to the usual scrutiny of scientific method: variation in the relevant environmental variables would be expected to generate corresponding variation in addiction-like behavior, both within and across species, and so on. This approach suffers from at least two drawbacks: i) in spite of the rigorous debate in the behavioral and medical sciences over what exactly constitutes an addiction, no consensus has emerged[6]; and ii) it might turn out that

[4]Fortunately, Garcia persisted and eventually published his early work in relatively obscure journals. Years later, one editor of a top journal–a distinguished learning theorist himself–declared on the occasion of his retirement that his one regret was rejecting Garcia's groundbreaking paper [101].

[5]Psychologist Martin Seligman coined the term "sauce béarnaise phenomenon" in relating his own experience with nausea aversion. It is now well known that humans, like rats, have a propensity to develop aversions to foods (for Seligman, the food in question was filet mignon with béarnaise sauce) associated with nausea, even if there is a significant delay between ingestion and illness (for Seligman it was six hours) and the victim is aware that the food was not the cause of illness (Seligman later discovered that, while his wife had eaten béarnaise sauce and not become ill, a colleague with whom he'd had close recent contact experienced a concurrent bout of the flu). Seligman, who years later still found béarnaise sauce unbearable, puzzled over the fact that many aspects of his experience (including the fact that he failed to develop aversions to other experiences from the night in question, including the opera, filet mignon, or the presence of his wife) stood in contradiction to his knowledge of classical conditioning theory. Garcia's work–the first report of which was published, as it turned out, the same week Seligman fell ill–cleared up the mystery [101].

[6]Indeed, one aim of this essay is to propose a meaningful definition of addiction. This issue is discussed further

the behavioral manifestations of "addiction" in natural settings bear very little resemblance to their modern counterparts in neuroscience laboratories and urban ghettos.

An alternative approach, suggested in the opening paragraphs of this essay, is to begin with an addictive *substance*, make note of the internal biochemical and physiological changes it induces in users, and search for examples of circumstances in which these same internal changes are observed in animals in their natural habitat. These circumstances will then presumably lead, as above, to hypotheses about the natural origins of addiction. This approach is possible, of course, only when scientific knowledge of the relevant internal molecular processes is in a relatively advanced state. In what follows we will make use of the fact that heroin, the quintessential example of an addictive substance, affects its victims by mimicking the endogenous opioids, one of the most thoroughly studied molecular systems in modern neuroscience. We acknowledge at the outset that this approach, with its narrow focus on a single class of substances, runs the risk of generating conclusions with only limited generality; this issue will be discussed further in Section 4.1.

2.3 The Adaptive Function of Endogenous Opioids

2.3.1 Opiates and Opioids

It has long been known that rats, given the opportunity, will self-administer morphine to the point of addiction. Whether pushing a lever to trigger an intravenous injection or sipping from a dilute solution, opiate-using rodents exhibit all the symptoms of addiction seen in their human counterparts: active substance-seeking behavior, reinforcement, tolerance, and withdrawal (see, e.g., [51], [118], [121]). In addition to being exceedingly convenient for the purposes of conducting experimental research on addiction, the fact that we share such a complex trait with a relatively distant cousin in the animal kingdom suggests strongly that there is something deeply innate and biological about drug addiction.

The specifics of the activity of morphine within the body have become known relatively recently. One of the more useful early innovations in opioid research has been the discovery of drugs that block or counteract the effects of the opiates. These drugs, known as *opioid antagonists*, often have opiate-like chemical structures and exhibit little or no interaction with non-opiate drugs. The theoretical underpinning to the action of opioid antagonists is that they interact with *opioid receptors* and compete with the opiate ligand.[7] In other words, when an opiate molecule enters the bloodstream,

in Section 3.2.1.

[7] Receptors and ligands are the locks and keys, respectively, of biochemistry. Though the degree of specificity can vary, ligands typically serve as the body's messengers: by virtue of their unique physical and chemical properties, ligands have the ability to selectively activate their target receptors, often triggering physiological responses at the cellular level. The textbook example of a ligand/receptor system is insulin, secreted by the pancreas in response to high blood sugar and detected by receptors throughout the body, touching off a variety of compensatory processes that bring blood sugar back into the normal range. Common subcategories of ligands include hormones, peptides,

it circulates through the body until it comes into contact and binds with an opioid receptor, which is then activated. If a large number of opioid receptors are activated simultaneously (e.g., if the concentration of opiates in the bloodstream is high), this triggers the cascade of physiological and behavioral changes associated with opiate use. Opioid antagonists, on the other hand, prevent the action of the opiates, often by binding with (but *not* activating) the target receptors, thus physically blocking the opiates from taking effect [23].[8] Though the opioid receptor was for many years merely a hypothetical construct, in the early 1970s advances in biochemical assay technology enabled scientists to confirm that there was indeed an opiate-specific receptor, located in cells throughout the body (though particularly concentrated, as it turns out, in certain regions of the brain) with test-tube reactivities mirroring the pharmacological activity of opiates and their antagonists, while exhibiting no reactivity with other drugs [88], [103], [112]. Today we know that there are at least three sub-types of opioid receptor (dubbed μ-, δ-, and κ-opioid receptors), and the genes coding for them have been identified (on chromosomes 6, 1, and 8 in humans, respectively).[9]

Opioid antagonists have been invaluable tools for addiction research. Early studies showed that the opioid antagonists *naloxone* and *naltrexone* effectively attenuate the physiological and behavioral effects of morphine in rats and monkeys, and even induce symptoms of withdrawal in morphine-using subjects [50], [61], [122]. The subsequent approval of these drugs for use in humans has provided evidence of their effects on subjective experience [46]. Naltrexone, for example, is known to effectively block the feeling of euphoria associated with heroin use, and for this reason it was once viewed as a promising treatment for heroin addiction. Unfortunately, the effects of opioid antagonists on drug self-administration are not straightforward: just as rats and monkeys have been known to *increase* consumption of morphine and heroin in response to naltrexone treatment (presumably to compensate for the reduction in hedonic effect), human addicts aware of the effects of naltrexone will often voluntarily discontinue treatment in order to once again experience the hedonic pleasures of heroin. For this reason, methadone (a mildly addictive opioid *agonist*) is often used for weaning addicts from heroin, although naltrexone is sometimes an effective tool under controlled (inpatient) conditions, as a surgical implantation, or with particularly motivated patients [56], [73].

The first direct confirmation of endogenous opioids came from Terenius and Wahlström in 1974 [113], who demonstrated that both cerebrospinal fluid and the extracts of rat brains had morphine-like receptor binding properties. Further investigations showed that three classes of endogenous opioids are produced in our bodies: the enkephalins, the endorphins, and the dynorphins. Of these, one in particular–*β-endorphin*–has perhaps the most morphine-like properties[10] and has received

and neurotransmitters [81].

[8] For this reason, opiates and other drugs that bind to and activate opioid receptors are sometime referred to as *opioid agonists*.

[9] For a review, see Gavériaux-Ruff and Kieffer [39].

[10] For example, repeated adminstration of β-endorphin leads to tolerance of its pain-relieving action and to morphine-like withdrawal symptoms when blocked with naloxone [116], [123], and β-endorphin (as well as the enkephalins) are

a great deal of attention in subsequent research. As is the case with the opioid receptors, the genes coding for each of the three endogenous opioid types have been identified; indeed, genetically engineered mice (so-called "knockout mice") lacking genes for the various opioid receptors and opioid precursor molecules have now been produced, making it possible to test more extensively the role of the endogenous opioid system in regulating physiology and behavior [39], [40].

In spite of the enormous body of research into the intricacies of the workings of the endogenous opioid system in humans and other animals, very little attention has been paid to the question of natural origins we hope to address here. There are, however, some intriguing clues: opioids have been shown to mediate a number of behaviors in controlled laboratory experiments, including most prominently drug self-administration, intracranial electrical self-stimulation, pain avoidance, sexual activity, feeding, and conditioned place preference. Of these, only the last four are likely to be relevant in natural settings, and two of these (pain avoidance and sexual activity) are generally *diminished* by opiate administration [117], [118]. This leaves feeding (which is stimulated, in the short-term, by opiates) and place preference (which can be reliably generated by opiate administration), both of which play important–and related–roles in the lives of foraging animals. In the spirit of recent review articles that have emphasized the centrality of endogenous opioids in ingestive behavior, palatability, and food cravings [74], [128], we will focus our attention on the adaptive role of the endogenous opioids in guiding feeding behavior in natural settings.[11] The next section will examine more closely the role of the endogenous opioids in modulating feeding behavior, and provide a sketch of the kind of adaptive problem they seem to be designed to solve. A formal statement of this problem is provided in Section 3.

2.3.2 Endogenous Opioids and Ingestive Behavior

One of the most thoroughly studied aspects of the action of endogenous opioids on behavior is their role in the short-term regulation of food intake. Numerous studies have shown that rats, for example, will eat less after being injected with opioid antagonists (e.g., [54], [86], [98], [104]), and the reverse is true for morphine and other opioid agonists (including β-endorphin), which reliably generate an *increase* in short-term food intake [45], [70], [95].[12] Similar responses have been observed in a wide variety of other foraging animals, from slugs [59] and cockroaches [60] to cats [37], pigs [3], and

addictive when self-administered by laboratory animals [9], [119]. The parallels with morphine were apparent early on: the word "endorphin" is derived from the phrase "endogenous morphine".

[11] The endocrinology of feeding is, of course, more complicated than this: many other molecular signals have been implicated in short-term feeding behavior, including serotonin, dopamine, neuropeptide Y, and cholescystokinin. These molecules are neglected here for the sake of brevity. Readers interested in the molecular complexities of short-term ingestive behavior are referred to the review by Cooper and Higgs [24]; the molecular basis of the long-term regulation of caloric intake–a different but related adaptive problem–is reviewed in Cummings and Schwartz [26].

[12] Although opioid antagonists were once viewed as a promising tool in the treatment of obesity, most studies have shown that they have little effect on feeding or body weight in the long term [102].

humans [2], [21].

A widely held view in the scientific community posits that opioids mediate food intake by influencing the perceived palatability of foods [74], [128]. There is a surprising amount of evidence for this in studies of animal behavior: the food-modulating effects of opioid agonists and antagonists are to some extent specific to palatable foods (usually sweetened with sugar or saccharin, but the effect has also been demonstrated for salty or alcohol-containing ingesta), with opioid agonists selectively increasing preference for palatable foods and opioid antagonists selectively decreasing preference for palatable foods (e.g., [25], [55], [69], [68]). Doyle *et al* [28] even find evidence for the palatability hypothesis in the facial expressions of rats: morphine injections increase food intake and increase positive hedonic facial reactions to a bitter-sweet solution, but do not change aversive reactions. In humans, opioid antagonists reduce both the hedonic ratings of palatable foods and the pleasantness ratings of palatable food odors, but do not reduce stated hunger ratings [29], [33], [129], [127]. In general, the subjects of these experiments appear to be implying–by their behavior, by their facial expressions, and by their words–that tasty foods are *even tastier* under the influence of morphine, but their appeal is diminished (relative to less-palatable foods) under the influence of naltrexone.

If it is true, as the behavioral effects of opioid agonists and antagonists seem to suggest, that endogenous opioids in our brains cause food to taste good, then we might expect that good-tasting food causes our brains to release endogenous opioids. There is evidence that this is indeed true: for example, the consumption of sweetened foods (but not bitter foods) causes an immediate release of β-endorphin in the brains and cerebrospinal fluid of rats [31], [126], and acute exposure to sweets induces reduced pain avoidance in rats and human infants, an effect that can be reversed with naltrexone [12], [13], [14].

So we are presented with the following puzzle: When an individual eats a food containing sugar, he triggers a biochemical cascade that causes him to eat more of that particular food, irrespective of his immediate caloric needs. Why might such a system have evolved? In other words, what competitive advantage might be gained by foraging animals that exhibit such a preference for sweet foods?

The answer given by behavioral ecologists is derived from the distribution of sugar in nature. High concentrations of simple carbohydrates are found in natural settings only in ripe fruit and raw honey, both of which reliably contain a host of valuable micronutrients (and, importantly, a dearth of toxins).[13] This suggests a simple role for endogenous opioids in an optimal foraging framework: when an environmental cue (such as the presence of sugar) indicates that a particular food is likely

[13]This distribution is not accidental: in a textbook example of coevolution, fruit-bearing plants rely on the services of foraging animals to disperse their seeds. This explains why fruit nearly always contains bitter or sour compounds and a green skin prior to maturation (i.e., before the seed is viable) and–when fully ripened–comes packaged not only with a brightly colored skin and edible, nutritionally valuable flesh, but also with non-digestible seeds or pits (see, e.g., Raven *et al*, [92], pp. 546-551).

to have nutritional value, endogenous opioids are released in the brain, generating the appropriate behavioral response. As we have seen, opiates exacerbate this effect; opioid antagonists counteract it.

Of course, sugar is not the only cue omnivorous animals use in distinguishing beneficial foodstuffs from harmful or useless ingesta. The tongues of humans and other omnivorous mammals have compound-specific receptors not only for simple carbohydrates but also for sodium, glutamate[14], and essential fatty acids–all of which could serve as nutritional cues in natural environments–and also for many toxic compounds, the dangers of which are perceived as a bitter taste [41], [110]. Interestingly, a (double blind, placebo-controlled) study by Yeomans and Gray [127] of the effects of naltrexone on the palatability of various foods found that pleasantness ratings were most dramatically affected for sweetened, fatty, and high-protein foods–that is, foods that contain chemical constituents for which there exist the molecular machinery necessary for direct tongue-to-brain neurological circuits.

In social animals such as humans, sheep, and chickens, it is also known that *social* cues play an important role in dietary choice.[15] Although there has been little study of the molecular basis for this mechanism, there is at least indirect evidence that endogenous opioids are involved. Autism, a developmental disorder characterized by a specific deficit in the ability to read social cues[16], is also associated with increased incidence of food cravings, pica (the ingestion of non-food objects), and eating problems [91]. That endogenous opioids might be involved in autism is suggested not only by the many similarities between the symptoms of autism and those of opiate addiction[17], but also by the elevated levels of endogenous opioids found in the cerebrospinal fluid of autistics [42], and by the observation that treatment of autistics with naltrexone increases social behavior and decreases hyperactivitiy, aggressiveness, and self-injurious behavior [19], [65], [66]. A concise explanation for these observations is that autistic children, thanks to abnormal endogenous opioid activity, are unable to infer the value of local foodstuffs by observing the actions of their parents and peers.

Any field biologist can testify that foraging is first and foremost a spatial decision problem. Neither fruits and vegetables nor predators and prey are randomly distributed in the wild–rather, a given type of plant or animal is likely to be found in a particular place at a particular time, or in a particular situation. That foraging animals do in fact have sophisticated mechanisms for associating place with reward is evidenced by the large literature describing the role of the endogenous opioids in

[14] *Glutamate* is a form of the amino acid *glutamine*, a molecular building block of protein. Glutamate is found in many natural foods; it is also the "G" in the flavor enhancer MSG.

[15] See review in Smith [106].

[16] This hypothesis is stated most clearly by Baron-Cohen [5]. More specifically, Baron-Cohen describes autism as "mindblindness" or the inability to infer intention from the actions of others. This would presumably be a prerequisite for any moderately sophisticated social acquisition of dietary habits.

[17] Symptoms common to both autism and opiate addiction include social withdrawal, insensitivity to pain, repetitive behaviors, calmness, fearlessness, growth retardation, feeding problems, a readily elicited gag reflex, and susceptibility to epileptic seizures [58].

generating conditioned place preference. Usually conducted in laboratory settings in which subjects (typically mice or rats) have access to multiple artificial "foraging sites," studies of this type have consistently shown that opiates can induce strong place preferences, and that the process of learning place preference can be blocked by administration of opioid antagonists.[18] Genetic variation in susceptibility to conditioned place preference has also been demonstrated, with some rat strains showing greater sensitivity to the place-conditioning effects of morphine and heroin [4].

There is also indirect evidence that opioid-mediated sensitivity to environmental cues varies with nutritional requirements. During pregnancy, for example, it has been hypothesized that vitamin C requirements rise (possibly because vitamin C enhances iron absorption) [35], a notion supported by the finding that vitamin C intake during pregnancy is the only micronutrient that predicts birth weight [71]. This increased physiological demand for vitamin C is accompanied by intensified food cravings, primarily for sweet foods [32]. Though the advent of artificially sweetened foods may have diminished the utility of such cravings, a meta-analysis of studies of gestational cravings found that fruit and fruit juice are nevertheless the most frequently reported target of gestational cravings [36]; and moreover a correlation between fruit consumption during pregnancy and serum vitamin C levels has been demonstrated [85]. Though the molecular basis for gestational cravings is not well-understood, endogenous opioids are implicated both by the specificity of cravings and by the many disturbances in the endogenous opioid system known to occur during gestation.[19] It does not seem farfetched, given the apparent coincidence of sweet- and fruit-specific cravings with both increased micronutrient requirements and disturbances in the endogenous opioid system, to suggest that in a pre-industrial environment such cravings might have served an adaptive function by increasing fruit intake during gestation.

Another way of thinking about the role of endogenous opioids in guiding behavior is that they serve as regulators of attention. This could explain why pain relief and loss of sexual appetite might be associated with acute opiate use: if a more urgent matter is at hand (e.g., this tasty treat needs to be eaten before someone else gets it!) it would make sense to temporarily set aside other concerns–such as a sore foot, or a sexual liaison. This view also seems to fit well with our knowledge of other (i.e., not food-related) instances in which endogenous opioids provide pain relief. The "runner's high" (generated by β-endorphin) associated with extended physical activity could make sense in natural environments, in which such activity would only take place for good reason (e.g., fleeing danger, or pursuing game); perhaps in such situations it would be better to ignore minor

[18] Beach [6] observed in 1957 that morphine-dependent rats developed strong preferences for the side of their cages in which the morphine was administered. Subsequent studies have successfully used morphine and other opiates to induce place preference in a variety of species, as well as place *aversion* when–in place of opiates–the opioid antagonist naloxone is used [77], [76].

[19] For example, endogenous opioid levels are known to rise during pregnancy in rats, monkeys, and humans, and all three classes of endogenous opioids (i.e., endorphins, enkephalins, and dynorphins) have been shown to vary dramatically during pregnancy [43], [44], [89], [115].

aches and pains until after the main event.[20] And interestingly, the "placebo effect" is apparently governed by endogenous opioids: several careful studies of postoperative pain have found that the pain relief induced by placebo can be blocked by the administration of opioid antagonists.[21] Here, the effect (to stop worrying about a persistent pain) would presumably make sense in situations in which the subject is treated and reassured (e.g., "Not to worry–I've seen cases like yours and they always heal quickly. Just take this medicine.") by an authority figure with experience in such matters.[22] In other words, endogenous opioids appear to regulate pain in a way that is consistent with an adaptive system of attention management.

Although the endogenous opioids no doubt have functions other than the regulation of short-term ingestive behavior, and short-term ingestive behavior is no doubt regulated by numerous other neuroendocrine processes in addition to the endogenous opioids, our aim here is to effect a simple demonstration of the way in which the solution to one particular adaptive problem might generate addiction-like behavior. In the next section we offer a formal model in which an environmental cue (the representative example of which is sugar) serves as an aid in the solution of a basic foraging problem: avoidance of micronutrient deficiency.

3 A Model, with Supporting Evidence

3.1 Chemical Cues and The Diet Problem

3.1.1 A Balanced Diet

In what follows we present a stylized model of nutritional ecology with informative cues. A foraging animal ("agent") is faced with a menu of two foods, x and a, and must choose how much of each to consume, given the limited capacity m of his gut and food densities 1 and $1/p$, respectively. There is a single limiting micronutrient for which there is a critical threshold: if the agent does not consume k units of nutrient, he will die. Unfortunately, this implies that survival is by no means certain, as the nutrient concentrations in foods x and a are independent random variables, denoted C_x and C_a, with distribution functions F_x and F_a, respectively. Our agent can, under these circumstances, do no better than to minimize the odds of death by malnutrition.[23] Formally, the *balanced diet problem* can be stated as follows:

[20] Indeed, there is substantial evidence of opioid-mediated analgesia in the presence of predation risk (see, e.g., [62], [96]).

[21] For a review, see ter Riet, *et al* [111].

[22] This interpretation finds additional support in the observation that–while the placebo effect appears to be mediated by opioids in instances of *postoperative* pain–the effect is much weaker when the pain is inflicted directly by the experimenter [111].

[23] Or equivalently, to maximize his probability of survival. In accordance with the principles of evolution by natural selection, agents able to calculate, intuit, or otherwise implement the correct solution to this problem would presumably out-compete their more death-prone brethren, and come to dominate the population in the long run.

$$\begin{aligned} \max_{x,a} \quad & P\left(C_x x + C_a a \geq k\right) \\ \text{s.t.} \quad & x + pa \leq m \\ & x, a \geq 0 \end{aligned} \tag{1}$$

Given the inherent uncertainty in this decision problem, it is clear that a cue providing new information about the nutritional value of one of the foods might alter the outcome. We will consider such a signal by positing two (informational) states of the world: one in which no cue is present, as above, and one in which a "positive" cue is observed, implying that the concentration of the limiting micronutrient in good a (the "addictive" good) is given by the random variable $\widehat{C}_a$, with distribution function $\widehat{F}_a$. To distinguish between these two information states, we will refer to the *no-cue balanced diet problem* and the *positive-cue balanced diet problem.* For purposes of illustration, we restrict our attention initially to distribution functions having the form $F\left(c;\gamma\right) = c^{\gamma}$, where $\gamma > 0$, and in particular, to the following parameterization:

Case 1 *"Cobb-Douglas Cue":* $\begin{aligned} F_x\left(c_x\right) &= c_x \\ F_a\left(c_a\right) &= c_a^{\beta} \\ \widehat{F}_a\left(c_a\right) &= c_a^{\widehat{\beta}} \end{aligned}$ *, where β and $\widehat{\beta}$ are parameters of the distribution functions such that $\widehat{\beta} > \beta > 0$*

We are now ready to state the following proposition:

Proposition 1 *If the agent faces concentration distributions described by Case 1 and the solution (x^*, a^*) to the no-cue balanced diet problem is such that $x^*, a^* > k$, then in the solution $(\hat{x}, \hat{a})$ to the positive-cue balanced diet problem, his consumption of good a will be strictly greater, $\hat{a} > a^*$.*

All proofs are provided in the Appendix. [24,25]

It is also possible to make a more general statement about the conditions necessary for the cue to result in an increase in consumption of good a (relative to the no-cue optimum a^*), and doing so will provide a useful illustration of the intuition behind the problem. In Proposition 2, we consider

[24]The assumption that x^* and a^* are greater than k in Proposition 1, made for analytical convenience, deserves comment. In the naturalistic interpretation given here, this assumption is equivalent to assuming that in the absence of a positive cue, the agent will choose to consume enough of each good that *if the nutrient concentration in either of these goods were 100%*, then his consumption of that good alone would be enough to ensure his survival. Given the typically miniscule amounts of micronutrients in natural foods (relative to the total weight of the food), and the large amounts of food typically ingested by foraging animals (relative to the required weight of micronutrients), we expect that this condition would rarely be violated in natural settings.

[25]In the Appendix we show that in Case 1, for $x, a > k$, the agent's behavior will be observationally equivalent to that of an agent maximizing a Cobb-Douglas utility function (Corollary 2). The properties of this class of utility functions are well known.

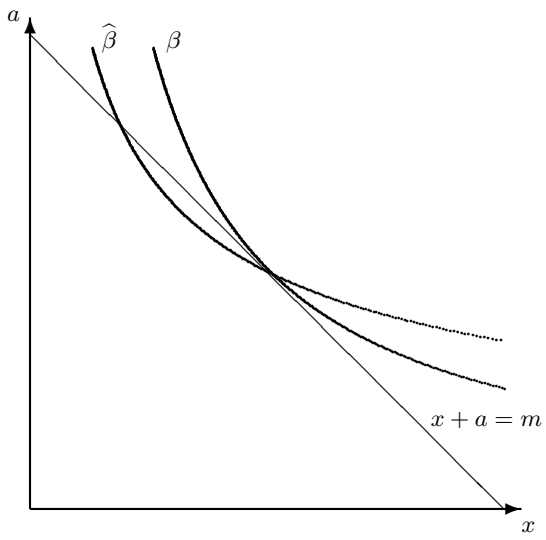

Figure 1: Effect of a Positive Cue ($\beta = 1$ and $\widehat{\beta} = 2$)

the effect on the probability of survival of making a small (ε) movement along the budget line in the direction of increasing a. Such a movement, of course, simultaneously decreases the amount of food x consumed, so there are two distinct effects on the probability of survival, which we denote the *ε-benefit* (i.e., the increase in probability of survival attributable to the increase in a) and the *ε-loss* (i.e., the decrease in probability of survival attributable to the decrease in x).[26] To formulate our more general condition we impose the following assumption.

Assumption 1 *The distribution functions F_x, F_a and $\widehat{F}_a$ have density functions f_x, f_a and $\widehat{f}_a$, respectively.*[27] *Moreover, the probability mass is distributed over the entire unit interval, i.e., $f_x(c_x) > 0$ whenever $c_x \in (0,1)$ and $f_x(c_x) = 0$ whenever $c_x \notin (0,1)$. The corresponding conditions are also satisfied by f_a and $\widehat{f}_a$.*[28]

Now we are ready to state our necessary and sufficient condition.

Proposition 2 *Let Assumption 1 be fulfilled. In addition, assume that there exist unique solutions $x^*, a^* > k$ and $\widehat{x}, \widehat{a} > k$ to the no-cue balanced diet problem and the positive-cue balanced diet problem, respectively, and that the indifference curves associated with the objective functions of the no-cue balanced diet problem and of the positive-cue balanced diet problem going through point (x^*, a^*) cross only at point (x^*, a^*). Then a cue increases the consumption of a if and only if there exists an $\varepsilon > 0$ such that the ε-benefit exceeds the ε-loss.*

[26] Explicit definitions of ε-benefit and ε-loss are provided in the Appendix.

[27] We assume the existence of density functions to simplify the statement of Proposition 2.

[28] Restricting the support of these density functions to the unit interval is motivated by basic laws of physics: nutrient content cannot be less than zero or greater than 100%.

The intuition behind Proposition 2 is perhaps best illustrated by considering how various realizations of the random variables C_x and $\widehat{C}_a$ translate into survival. The probability of survival for a given allocation (x, a) can be represented graphically in (c_x, c_a) space. In Figure 2, the line intersecting the vertical axis at $\frac{k}{a}$ (i.e., $c_x x + c_a a = k$) represents for our agent the ultimate threshold: any realization (c_x, c_a) above this line and the agent survives; any realization below and he dies. With an ε-movement along the budget line (i.e., to the allocation $\left(x - \varepsilon, a + \frac{\varepsilon}{p}\right)$), the threshold pivots to intersect the vertical axis at $\frac{k}{a+\frac{\varepsilon}{p}}$. This change in consumption increases the survival area by B while it decreases the survival area by L. Thus the probability that (c_x, c_a) falls in B is the ε-benefit, while the probability that (c_x, c_a) falls in L is the ε-loss.

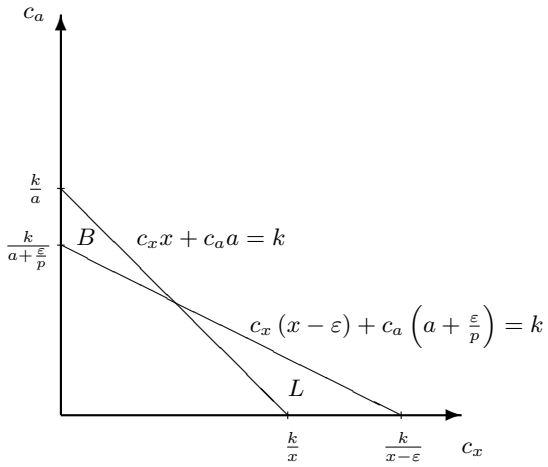

Figure 2: ε-benefit and ε-loss

Thus far we have addressed only the effect of an informationally valuable cue in a static choice environment. But if our aim is to say something about the relationship between this environment and addiction-like behavior, we will need to consider the dynamics of the balanced diet problem. In the next section we examine the simple dynamics of choice when learning is possible.

3.1.2 The Simple Dynamics of Learning

In order to capture the notion of learning within our formal framework, it is necessary to re-formulate the decision problem slightly. In Section 3.1.1, consumption decisions were always made with complete knowledge of the probability distributions underlying the foods of choice. To accommodate learning, we will now impose a modicum of foresight on our agent, requiring him to choose a diet *before* confirming the presence or absence of a cue. In so doing, we will show how the simple dynamics of learning can, under very general conditions, generate adjacent complementarity, the behavioral property driving theories of rational addiction.

Again, we distinguish between two states: the no-cue state in which the agent's objective function is given by $\overline{u}(x, a) = P(C_x x + C_a a \geq k)$ and the positive-cue state in which his objective function is

given by $u(x,a) = P\left(C_x x + \widehat{C}_a a \geq k\right)$. The following assumption imposed on these two "utility" functions is consistent with the framework developed in Section 3.1.1.

Assumption 2 *The functions u and $\overline{u}$ are twice continuously differentiable and strictly concave in the area satisfying $x, a > k$. In addition, the no-cue balanced diet problem and the positive-cue balanced diet problem have unique solutions $\left(x^{\overline{c}}, a^{\overline{c}}\right)$ and (x^c, a^c)respectively, with $a^c > a^{\overline{c}}$; $x^c, a^c > k$; and $x^{\overline{c}}, a^{\overline{c}} > k$.*

We shall denote by Π_t the random variable describing the agent's prior beliefs in period t concerning the possible probabilities with which a positive cue might arise. Hence, Π_t maps into the space of probabilities and thus takes values on $[0,1]$. We shall denote by g_t the density function describing the distribution of Π_t.[29]

A bundle (x,a) is a survival maximizing bundle of period t if it solves

$$\begin{aligned} \max_{x,a} \quad & E\left[\Pi_t u(x,a) + (1-\Pi_t)\,\overline{u}(x,a)\right] \\ \text{s.t.} \quad & x + pa \leq m, \\ & x, a \geq 0 \end{aligned} \tag{2}$$

In what follows we write v_t for the objective function of problem (2). We can re-write problem (2) as $\max_{a \geq 0} v_t(m - pa, a)$ by the monotonicity of v_t. After choosing a bundle $(m - pa, a)$ the agent observes the period t outcome (i.e., presence or absence of a cue) and updates his beliefs for period $t+1$ in a Bayesian manner. In order to isolate the effect of the positive cue, we will assume the agent's budget constraint remains the same in every period. If he observes a positive cue in period t, his posterior beliefs are given by

$$g_{t+1}(\pi) = \frac{\pi g_t(\pi)}{\int_0^1 \pi g_t(\pi)\, d\pi} \tag{3}$$

for all $\pi \in [0,1]$. We shall denote by Π_{t+1} the random variable corresponding to density function g_{t+1}.

As noted in Section 2.1, the dynamic property known as adjacent complementarity has been identified as essential to a behavioral theory of addiction. In the present framework, an analogous property can be concisely defined:

Definition 1 *The agent's behavior meets the conditions for* adjacent complementarity *at period t and at point $(m - pa, a) \in \mathbb{R}^2_+$ on the budget line if $\frac{d}{da} v_t(m - pa, a) < \frac{d}{da} v_{t+1}(m - pa, a)$.*

Our next proposition states sufficient conditions for adjacent complementarity.

[29] It is convenient but not necessary to assume that the distribution of Π_t possesses a density function; our results can be derived by assuming a general distribution function G_t.

Proposition 3 *If Assumption 2 is satisfied, then a positive cue generates adjacent complementarity at any time t and at any point (x, a) lying on the budget line such that $a \in \left[a^{\overline{c}}, a^{c}\right]$. Moreover, if the agent chooses a bundle in period t such that $x^*, a^* > k$, then $a^* \in \left(a^{\overline{c}}, a^{c}\right)$ and $x^* = m - pa^*$.*

Corollary 1 *Under the assumptions of Proposition 3 the agent's behavior exhibits adjacent complementarity at the optimal solution (x^*, a^*) of problem (2).*

3.2 Addiction as Learning Gone Awry

3.2.1 What is Addiction?

In the theory of rational addiction, a good is addictive, roughly speaking, if its *marginal* (instantaneous) utility increases with experience. Rational addictions, however, may be either beneficial or harmful depending on how experience affects *total* (instantaneous) utility: in a beneficial addiction total utility increases over time, while in a harmful addiction it decreases over time. In the popular lexicon, the word addiction generally excludes the former case: one might have "good" and "bad" habits, but addiction is generally taken to imply a regrettable behavior. The careful reader will have noticed that thus far the model we have presented seems to imply that all "addiction" is beneficial: in the learning dynamic we have proposed, utility (or its proxy in our framework, the expectation of survival probability) is always increasing over time, as the agent learns more about the world in which he lives. Now we turn our attention to the subject of harmful addiction.

The circumstances we propose as being conducive to harmful addictions are perhaps best illustrated by returning to our representative example of a behavioral cue. Although sugar is conveniently associated with valuable micronutrients in natural settings, the advent of commercially viable sugar refining technology early in the twentieth century changed this association dramatically. Today, foods with the highest sugar content often contain no micronutrients whatsoever; in fact, one of the most consistent messages of modern health advocates has been a simple admonition: *eat less sugar*. But the biochemical system upon which we rely in choosing our foods has not changed: being encoded in our genes in a way that leaves it mostly immune to conscious manipulation, the endogenous opioid system still reacts to sweet foods *as if* they remained a rare and valuable commodity. Within our framework, this implies a discrepancy between the behavior of the agent and maximization of the objective function v_t. We will find it useful, therefore, to specify a *subjective function*, $\widetilde{v}_t$, that reflects the probability distributions $\widehat{F}_a$ and F_a and density function g_t prevalent in what might be called the agent's "ancestral environment"–that is, the conditions to which the agent's behavior is adapted.[30] This suggests the following definitions:

[30] Decision theorists have long held that beliefs can be thought of as *subjective*, or implicit in one's behavior–i.e., that beliefs can be inferred by observation, with no knowledge of the actor's internal cognitive processes. Leonard Savage's classic treatise [99] provides a formal but accessible exposition of this notion.

Definition 2 *The agent's behavior meets the conditions for* subjective adjacent complementarity *at period t and at point $(m-pa,a)\in\mathbb{R}^2_+$ on the budget line if $\frac{d}{da}\widetilde{v}_t\left(m-pa,a\right)<\frac{d}{da}\widetilde{v}_{t+1}\left(m-pa,a\right)$.*

Definition 3 *The agent's behavior meets the conditions for a* harmful addiction *at the decision sequence $(x_1,a_1),(x_2,a_2),\ldots$ on the budget line if it satisfies subjective adjacent complementarity at (x_t,a_t) and $v_t\left(x_t,a_t\right)>v_{t+1}\left(x_{t+1},a_{t+1}\right)$ for all periods t.*

As our example suggests, we will consider the special case in which the cue does not convey any kind of information, so that $\widehat{F}_a$ is identical to F_a. This implies–for the *objective* function v–that $v_t=v_{t+1}=\overline{u}$ in equilibrium for all periods t independently of the agent's beliefs about the probabilities of the arrival of a positive cue. The agent's subjective function, however, still specifies that $\widehat{C}_a$ and C_a have different distributions. We shall denote by $(\widetilde{x}_t,\widetilde{a}_t)$ the solution of the problem that we obtain from (2) by replacing v_t with $\widetilde{v}_t$. Now we are ready to formulate our proposition concerning harmful addiction.

Proposition 4 *Let Assumption 2 and $a^c>\widetilde{a}_t>a^{\overline{c}}$ for all t be fulfilled. Moreover, suppose that although a dissociation of the cue from the limiting micronutrient results in an uninterrupted series of positive cues that provide no information, the agent continues to maximize the subjective function $\widetilde{v}_t$. Then a harmful addiction will occur at the sequence $(\widetilde{x}_1,\widetilde{a}_1),(\widetilde{x}_2,\widetilde{a}_2),\ldots$.*

In effect, Proposition 4 says–for the special case of technological change considered here–that harmful addictions are the product of a mismatch between the modern world and the "beliefs" about the world implicit in our behavior.[31] This mismatch is generated, in the present example, by the rapidity with which food processing technology has advanced while the human genome (in which, it bears repeating, the "recipe" for the endogenous opioid system is literally written) has remained effectively unchanged. In other words, our approach suggests that a harmful addiction is a habit acquired under false pretenses.

In spite of the evidence presented thus far, some readers may nevertheless remain uncomfortable with the notion of "sugar addiction". Recent studies by Hoebel and his colleagues have suggested that sugar does indeed share more properties with drugs of addiction than had previously been thought: feeding rats excessive amounts of sugar, for example, and then either depriving them of food or injecting naloxone induces symptoms typical of opiate withdrawal such as teeth chattering, forepaw tremor, and head shakes [22]. Remaining skeptics may find solace in the fact that replacing "sugar" with the word "alcohol" in this story will not change its character or consistency with available scientific knowledge.[32] The distribution of alcohol in nature mirrors that of sugar (i.e., it

[31] The "evolutionary mismatch" theory of substance abuse represents the conventional wisdom among students of human evolution. See, for example, the work of Nesse and others [79], [80], [125] (but see also footnote 36).

[32] It is the experience of the authors that although few will acknowledge an overly zealous propensity for alcohol consumption in themselves, many are able to identify alcoholism in others.

is found only in ripe fruit), it is subject to opioid-mediated self-administration[33], and only industrial fermentation and distillation technologies have made it readily available in the modern world [30]. The recent identification of human genes that confer a higher risk of alcoholism provides further support for the notion that alcohol consumption might have had adaptive significance in human evolutionary history [100].

It does not seem inappropriate to suggest that heroin addiction also fits well with the "mismatch" model of addiction. When a user seeks out and injects heroin in order to experience once again the sudden activation of the opioid receptors in his brain, he is following an ancient algorithm: *when the opportunity arises, devote your energies to activities that make you feel like this.* The algorithm is, of course, more complicated than this, and vestiges of the original function of the behavior can be seen in some of the particulars of the experience of addicts: heroin addicts, for example, reportedly experience overpowering cravings for sugar during withdrawal [124], and relapses among reformed addicts are often triggered by place-specific contextual cues [20]. Heroin, like refined sugar and distilled alcohol, is a relatively recent innovation in our collective history, first synthesized from morphine in 1874.[34]

That the neurologically active substances such as morphine, cocaine, and nicotine found in plant tissues might be harmful to our health is not surprising when the origins of these compounds are considered. Once thought to be merely metabolic by-products, plant ecologists now believe that compounds like these (which are energetically costly to produce but have no apparent function in plant physiology) arose in the course of plant evolution as defensive mechanisms designed to deter herbivorous animals.[35] These substances effectively deter herbivory because they are highly potent neurotoxins: for instance, the oral ingestion of as little as 2 grams (0.07 oz.) of raw opium (or 0.2 grams refined morphine, or 0.04 grams nicotine), can be fatal [87]. Coevolutionary forces work both ways, of course, and animals have in turn evolved methods of avoiding plant toxins, most notably by detecting them and ejecting them, either by tasting them directly with receptors on the tongue and spitting them out (bitter aversion) or by vomiting upon the onset of illness (nausea aversion) [106]. This helps to explain why drugs of addiction are commonly smoked, snorted, or injected but rarely chewed up and eaten: our bodies have natural mechanisms that prevent ingestion of toxins. That some of these toxins, taken in moderation, selectively activate specific "reward" centers in the brain that govern addiction appears to be an accident of plant-herbivore coevolution. This "accident" nevertheless displays all the hallmarks of an adaptation in natural settings.[36] Raw opium, for

[33]For a review of the large scientific literature implicating opioids in alcohol's addictive properties see Van Ree *et al* [117], pp. 375-378.

[34]In one of the more spectacular blunders in the annals of the pharmaceutical industry, heroin was originally developed and marketed by The Bayer Company as a less-addictive form of morphine [16].

[35]See, e.g., Raven *et al*, pp. 546-551 [92].

[36]This hypothesis is not completely uncontroversial–some argue that substance abuse may have been around long enough for the human genome to have developed defensive mechanisms (see, e.g., Sullivan and Hagen [109]). The

instance, contains a host of alkaloids in addition to morphine, many of which (e.g., papaverine, codeine, narcotine, and thebaine) have little or no narcotic effect but act as stimulants of the medulla and spinal cord [87]. Taken together in their natural form, these compounds constitute a dangerous drug cocktail, and one important function of drug delivery technologies is to isolate–and thus detoxify–the target compound.

It has been conjectured that such a mismatch between objective reality and the "beliefs" about the world implicit in our genes could explain the subjective difficulties with "self-control" many people report when describing their experience with drugs of abuse or sweetened treats [105].[37] In the present context, the "belief" implicit in our genes (and also implicit–when viewed in the broad context provided here–in our behavior) is that foods containing simple carbohydrates are nearly always nutritionally valuable, while the objective reality is that in today's world such foods are more often than not lacking in such value. A literal interpretation of our theory of harmful addiction implies an agent who perpetually expects a large benefit to accrue from a particular activity, but–when the expected benefits are not realized–finds himself constantly regretting his past actions.[38] We contend that this interpretation has meaningful parallels with economic theories of self-control or "time inconsistent" behavior: in these theories, the agent typically overweights (i.e., assigns higher utilities to) current consumption to the detriment of his long-term well-being (i.e., his future utility stream) [1], [64]. Within our framework, time inconsistency is the manifestation of emotional mechanisms maladapted to certain aspects of the modern world and underlain by molecular processes that operate below the level of consciousness.

3.2.2 Dopamine and the Neurobiology of Learning

Even the most devout hedonist would admit that habit formation (including drug addiction) can be viewed as a case of learning. If the hedonist finds that consumption of a particular good gives him pleasure, for example, it would make sense to take this information into account when making subsequent consumption decisions. If subsequent consumption provides additional information (presumably also of a positive nature) about the hedonic properties of the good (e.g., higher levels of consumption correspond to more a intense pleasure), then the resulting behavioral dynamic could closely approximate that predicted by a theory of rational addiction. But this explanation quickly stretches thin where harmful addictions are involved: the typical drug addict quickly becomes aware of the hedonic properties of his drug of choice, and his behavior often becomes increasingly

debate, however, is mostly one of degree: no one would argue, for example, that hypodermic needles have been around long enough for humanity to develop an innate aversion to heroin.

[37] Indeed, in some situations admitting reasonable levels of subjective uncertainty can transform an apparent self-control problem into an optimal behavioral strategy. Sozou, for example, shows that hyperbolic discounting of future rewards can be optimal where default is possible and the hazard rate is uncertain [107].

[38] This interpretation is, of course, *overly* literal, because our hypothetical "expectation" need not be conscious or even subject to conscious control.

pathological (i.e., less informed by the dictates of informed rationality) over time. Is it really appropriate, as we have suggested, to view harmful addiction as some kind of learning disorder?

One way of answering this question would be to identify the neurological basis of learning: if our hypothesis is correct, drugs of abuse would be expected to act on the same physical substrates employed in healthy, natural learning processes. While modern science is far from providing a definitive picture of how the internal workings of the mammalian brain translate into sophisticated learning abilities, several authors have noted that one process in particular conforms well with the predictions of classical learning theory: dopamine transmission in the limbic system.[39,40,41] Although dopaminergic neurons are present in many areas of the mammalian brain, many of those located in limbic system appear to have the intriguing property of being subject to activation by natural rewards only when associative learning is taking place: in classic Pavlovian conditioning experiments, these neurons are activated *in the presence of novel, but not conditioned stimuli.* In other words, these neurons seem to indicate that *learning* is taking place: when an animal is first presented with a novel visual or auditory stimulus while being fed a tasty treat, dopaminergic neurons in his limbic system light up; but with experience, the stimulus/treat pairing loses the ability to activate these neurons. Once a subject is conditioned, only "surprises"–such as the pairing of food reward with a stimulus not previously associated with the reward–will re-activate the system.

So what are the effects of opiates on dopamine transmission in the limbic system? Interestingly, stimulation of dopamine transmission in this part of the brain is one of the few properties shared by virtually all addictive drugs–not only opiates, but also alcohol, nicotine, cocaine, amphetamines, and Δ^9-tetrahydrocannabinol.[42] The most powerfully addictive drugs, however, differ from natural rewards in that their effects on dopamine transmission are not diminished by repeated administration.[43] This would seem to suggest that if dopamine in the limbic system does in fact represent a

[39]Though precise definitions vary, the limbic system in the mammalian brain is comprised of several interconnected structures, generally including the amygdala, hippocampus, hypothalamus, septum, nucleus accumbens, cingulated gyrus, and parts of the cortex. The limbic system is thought to play a central role in the regulation of emotions [17].

[40]Although neurons (nerve cells) can employ more than one neurotransmitter at a given synapse (a *synapse* is the gap between cells across which neurotransmitter ligands carry information), much intercellular communication in the mammalian brain is mediated by neurotransmitters such as dopamine, norepinephrine, epinephrine, or serotonin in distinct cells. Hence the terms "dopaminergic neuron," "serotinergic neuron," etc. See, for example, [130].

[41]Several authors have focused particular attention on the nucleus accumbens shell and the ventral tegmental area as the putative loci of associative learning [27], [108]. Unfortunately, brain mapping is far from an exact science, and similar observations have been made in other brain structures. The advanced state of current technology, however, is evident in the recent report of Waeltl *et al* [120] (complete with simultaneous measurement of eye position and the activity of individual neurons in the subjects' brains), in which monkeys were trained to associate the delivery of fruit juice with distinctive visual stimuli.

[42]Δ^9-tetrahydrocannabinol, or THC, is the pharmacologically active constituent of marijuana.

[43]The difference presumably stems from the fact that the endogenous neurochemical signals generated by natural rewards are subject to adaptive regulation; exogenous ligands (i.e., drugs) are not subject to such limitations. It is

physical substrate of learning, then the sort of learning that takes place in the presence of drugs of addiction is properly viewed as pathological.

4 Discussion

4.1 Beyond Opioids

From the discussion of the previous section, it is clear that the logic of our analysis might well apply to substances that act on receptor systems having nothing to do with the opioids. We urge caution on those who would extend our analysis in this way; and though for the remainder of this essay we will turn to a more general discussion of substance abuse, we do so with some degree of trepidation. That some substances target other receptor systems suggests that they are disrupting a different adaptive response. In principle, before drawing conclusions about the adaptive function of the receptor system targeted by a given drug of abuse, systematic study of the role of the target system in natural settings should be undertaken. Once the adaptive function of the target receptors are identified, specific implications (including, for example, circumstances under which it is likely to be used, or the sense in which it meets the criteria for "harmful" addiction) for the relevant addictive substance would presumably follow. The prospect of such an undertaking for all behavior-altering substances is daunting, given the unfortunate fact that–for most drugs–our understanding of the myriad effects of such substances on human behavior, cognition, and physiology remains poor.

More potential pitfalls to be encountered on the road to a more general theory of addiction are to be found in the shortcomings of our own analysis, of which we are all too aware. For starters, the foraging problem we've specified is overly simplistic: although the stylized trade-off between two goods we examine in Section 3 above is a useful framework for analysis and not inconsistent with standard practice in the economics literature, foraging animals are actually faced with a wide variety of potential foods of uncertain quality. This suggests a model of search in which the trade-offs among many choices need to be considered, perhaps with explicit consideration of time costs, risk of poisoning, risk of starvation, etc.[44] We have neglected these aspects for the sake of clarity. Another important simplification we have made is to treat the endogenous opioid system as if it had only one function. Although there is indeed an abundance of evidence pointing to a role for endogenous opioids in the foraging problem we've specified (avoidance of micronutrient deficiency), and all of the endogenous opioids share similar molecular structures, the differences among them are important: these small differences determine the affinity of a given opioid for a given receptor type or sub-type. For example, two recently discovered endogenous ligands–dubbed *endomorphin-1* and

important to note that the distinction is not absolute: drugs of addiction *are* subject to habituation, but to a much lesser degree than natural rewards.

[44] See Smith [106] for a more complete exposition of this problem. Mercer and Holder [74] also implicitly advocate consideration of the search problem with their emphasis of the role of the endogenous opioids in food cravings.

endomorphin-2–have distinct effects on behavior: the former induces conditioned place preference, while the latter induces conditioned place *aversion*. Studies with selective antagonists and knockout mice have suggested that both ligands act on the μ_2-subtype of the μ-opioid receptors, though only endomorphin-2 appears to also interact with δ- and κ-opioid receptors.[45] Yet another problem arises from the fact that opioid receptors are found throughout the brain and the body, and as a result the effects of opioids on behavior and physiology can vary dramatically depending on where they find their targets. This creates a sticky measurement problem for scientists studying opioids: for example, there is evidence that natural levels of endogenous opioids in the body's periphery may vary inversely with endogenous opioid levels in the brain, making misinterpretation of data all too easy [74].

4.2 Foreseeing Addiction

As noted in Section 2.1, the question of the degree to which consumers foresee the consequences of drug addiction is the subject of much debate in the rational addiction literature. Though we have intentionally suppressed the possibility of foresight in our analysis (in order to emphasize that even myopic decision-making can generate the dynamic properties necessary for habit formation), there can be no doubt that consumers are, to some extent, aware of the future (social, health, and economic) consequences of drug abuse. Not that people "know" the consequences of addiction in the same way people "know" it's good to eat ripe fruit, or the way people "know" the diet that sustained them in childhood is unlikely to harm them later in life. But there's every reason to expect that people could learn from health that worsens with use, from watching relatives die or suffer, from reading about health consequences, or from warnings on labels. Indeed, such learning could well be interpreted within our framework as constituting "informative cues" that influence the agent's beliefs (i.e., $\widehat{F}_a$, F_a, and g_t) and behavior accordingly. Or, extending our framework a bit, an agent who becomes aware of the harmful dynamic associated with drug use might well choose to quit "cold turkey" as the only way to stop the arrival of the hedonic "false cues" associated with use.

That consumers choose with foresight is the driving force behind the main empirical prediction of the rational addiction literature: that *future* price increases will generate a decrease in *current* consumption. Several studies have borne out this prediction by measuring, for example, the effect of announced (but not yet effective) increases in cigarette taxes.[46] We acknowledge this empirical phenomenon, but would also suggest an alternative explanation: it could well be that, coincident with the announcement of tax increases on cigarettes, there is an increase in public awareness (due,

[45] For a review of these findings, see Bodnar [15], p. 2324.

[46] The most compelling support for this phenomenon is provided by Gruber and Köszegi [48]; see also references therein.

perhaps, to increased news coverage or increased funding of public health campaigns) of the health consequences of smoking–indeed, such public awareness might well precede and precipitate legislative action. We leave the question of the relative importance of these competing hypotheses (foresight vs. awareness) for future research.

4.3 Implications for Public Policy

When the time comes to translate a carefully crafted economic theory of addiction into recommendations for public policy, it quickly becomes clear that core assumptions about information and personal responsibility drive everything. After all, the fact that drugs (both legal and illicit) are bought and sold in the marketplace is what motivates the use of economic analysis in the first place, and the efficiency of the market mechanism in allocating resources is beyond dispute. If it were true that consumers choose to consume addictive substances with complete foresight, without uncertainty or self-control problems, there would seem to be little justification for the government to interfere with market transactions.[47] But with uncertainty, incomplete information, and time inconsistency, a role for policy is introduced for "paternalistic" reasons that don't necessarily apply to ordinary consumer goods.[48] This aspect of drug policy finds support within our framework. While it is hard to deny that the lives of heroin junkies might be improved by restrictions on drug availability, perhaps the most surprising aspect of our findings is that this same line of reasoning could be applied to the consumption of refined sugar.

We expect that our findings will come as good news to advocates of public health. We want to be quick to note, however, that we do not mean to imply that individuals who choose to smoke, or drink, or eat sweets are necessarily *wrong*, in the personal sense, to do so. First of all, the discrepancy we noted between the objective function v_t and the subjective function $\widetilde{v}_t$ in Section 3.2.1 is to some extent irrelevant to the considerations of human welfare implicit in economic analysis of public policy. Conventional economic analysis is right to focus on "as if" utility maximization–the equivalent of our subjective function–because it represents our best proxy for the real pleasure and pain experienced by consumers in everyday life.[49] It is true that–given the historical novelty of

[47] This is strictly true, of course, only under idealized market conditions. If, for example, the additional healthcare costs incurred by a smoker are covered by insurance, considerations of economic efficiency would dictate that the smoker incur an equivalent cost (in the form of, say, a cigarette tax or increased insurance premiums) contingent on his decision to smoke. The possibility of external effects (e.g., crime, or second-hand smoke) imposed by addicts on others could also provide justification for market intervention. Such considerations are not unique to addictive substances (and therefore will be largely neglected in the present analysis), but they would certainly deserve attention in a more complete analysis of drug policy.

[48] Such policies might include, for example, taxes on drugs [34], [48], public education campaigns [11], [83], restrictions on advertising [11], [63], [83], regulation of drug dispensation [11], [18], restrictions on public consumption [11], [63], rehabilitation programs [83], and criminalization [11].

[49] It has been noted (by, for example, Kahneman *et al* [57]) that consumer behavior often appears to be inconsistent

modern drug technology–that long-term consequences are likely, in the absence of concerted efforts to educate the populace, to be systematically underestimated by the average user. But the trade-off between immediate pleasure and future costs is just that: a trade-off; and it is not hard to imagine circumstances in which indulging in addiction might, on balance, make an individual better off. Indeed, such circumstances are suggested by changes in the incidence of smoking in the U.S. in the past four decades: as increasing regulation and aggressive public education campaigns have sharply reduced the prevalence of smoking[50], the incidence of smoking has become increasingly concentrated in those with lower socioeconomic status and in individuals suffering from such behavioral or affective disorders as depression, adult attention deficit hyperactivity disorder, anxiety disorders, and bulimia. In an insightful review of the medical literature, Pomerleau [90] has argued that in each of these cases nicotine dependence appears to ameliorate the symptoms, making life more livable for the afflicted.[51] In other words, though smoking might play no beneficial role in a perfect world in which all live healthy, happy lives, that is not the world in which we live.

5 Conclusion

That human behavior can usefully be thought of as "rational" is a central tenet of economic theory. But what do we mean by rationality? In the popular lexicon, rationality is coolly deliberate, conscious decision-making. This standard of rationality is clearly not met for many users of illicit drugs. A less rigorous standard is "as if" rationality–satisfied if observed behavior is consistent with the solution to an optimization problem. Is this definition of rationality met for drug addiction? It depends very much on how the problem is specified: if one is willing to be flexible with the domain of preferences, with time inconsistency, and with uncertainty and prior beliefs, then surely *any* pattern of behavior can be justified as the solution to an optimization problem. This is not to say that the rigorous scientific debate over the essential properties of a positive economic theory of addiction has not produced useful insights–on the contrary, this process is the lifeblood of scientific knowledge. Nevertheless, it is our hope that by providing the beginnings of a biological foundation for the theory of rational addiction, we will have helped in some small way to better inform the debate. The framework developed here provides support for the notion that adjacent complementarity can be expected to be important in some consumptive behaviors, and that addiction might indeed be related to problems with self-control, to emotional mechanisms, and to false prior beliefs. But making sense of these disparate behavioral phenomena is easier when we acknowledge directly that

with the intertemporal maximization of hedonic experience. Biased or false subjective "beliefs" of the sort implied by our subjective function $\widetilde{v}_t$ may provide one source of such inconsistency.

[50] Between 1965 and 1990, for example, smoking among U.S. adults declined from 40% to 29% [114].

[51] Though the "self-medication" role for psychotropic substances remains controversial, few would argue that the pain-relieving function of drugs such as morphine has not proved beneficial in the practice of modern medicine.

what we observe is the manifestation of a sophisticated biological system in which environmental cues trigger predictable internal neurological and physiological responses; that this system shows all the signs of being adapted to a pre-industrial environment; and that drugs of abuse, largely developed in the modern era, have the demonstrable ability to disrupt this system.

Appendix

Proof of Proposition 1. If the agent chooses bundle (x, a), then the nutritional content is $C_x x + C_a a$, which has density function

$$\begin{aligned} h(t) &= \int_{-\infty}^{\infty} \frac{1}{ax} f_x\left(\frac{y}{x}\right) f_a\left(\frac{t-y}{a}\right) dy = \\ &= \int_{\max\{0,t-a\}}^{\min\{x,t\}} \frac{\beta}{ax}\left(\frac{t-y}{a}\right)^{\beta-1} dy = \\ &= \left[-\frac{1}{x}\left(\frac{t-y}{a}\right)^{\beta}\right]_{\max\{0,t-a\}}^{\min\{x,t\}} \end{aligned}$$

Carrying out the substitutions we obtain

$$h(t) = \begin{cases} \frac{1}{x}\left(\frac{t}{a}\right)^{\beta} & \text{if} \quad x \le a, 0 \le t < x, \\ \frac{1}{x}\left(\frac{t}{a}\right)^{\beta} - \frac{1}{x}\left(\frac{t-x}{a}\right)^{\beta} & \text{if} \quad x \le a, x \le t < a, \\ \frac{1}{x} - \frac{1}{x}\left(\frac{t-x}{a}\right)^{\beta} & \text{if} \quad x \le a, a \le t \le x+a, \\ \frac{1}{x}\left(\frac{t}{a}\right)^{\beta} & \text{if} \quad a < x, 0 \le t < a, \\ \frac{1}{x} & \text{if} \quad a < x, a \le t < x, \\ \frac{1}{x} - \frac{1}{x}\left(\frac{t-x}{a}\right)^{\beta} & \text{if} \quad a < x, x \le t \le x+a, \\ 0 & \text{if} \quad t < 0 \text{ or } x+a < t. \end{cases}$$

Now in order to derive the agent's objective function we must determine $P(C_x x + C_a a \ge k) = \int_k^{\infty} h(t)\, dt$. In particular,

$$P(C_x x + C_a a \ge k) = \begin{cases} 0 & \text{if} \quad a + x \le k \\ 1 - \frac{k-a}{x} - \frac{1}{\beta+1}\frac{a}{x} + \frac{1}{\beta+1}\frac{a}{x}\left(\frac{k-x}{a}\right)^{\beta+1} & \text{if} \quad a \le k < a + x \\ 1 - \frac{1}{\beta+1}\frac{a}{x}\left(\frac{k}{a}\right)^{\beta+1} + \frac{1}{\beta+1}\frac{a}{x}\left(\frac{k-x}{a}\right)^{\beta+1} & \text{if} \quad x \le k < a \\ 1 - \frac{1}{\beta+1}\frac{a}{x}\left(\frac{k}{a}\right)^{\beta+1} & \text{if} \quad 0 \le k < x \\ 1 & \text{if} \quad k < 0, \end{cases}$$

whenever $x \le a$, and

$$P(C_x x + C_a a \ge k) = \begin{cases} 0 & \text{if} \quad a + x \le k \\ 1 - \frac{k-a}{x} - \frac{1}{\beta+1}\frac{a}{x} + \frac{1}{\beta+1}\frac{a}{x}\left(\frac{k-x}{a}\right)^{\beta+1} & \text{if} \quad x \le k < a + x \\ 1 - \frac{k-a}{x} - \frac{1}{\beta+1}\frac{a}{x} & \text{if} \quad a \le k < x \\ 1 - \frac{1}{\beta+1}\frac{a}{x}\left(\frac{k}{a}\right)^{\beta+1} & \text{if} \quad 0 \le k < a \\ 1 & \text{if} \quad k < 0, \end{cases}$$

whenever $x > a$. The indifference curves can be divided into five distinct regions, which we illustrate in Figure 3. The death zone $A^0 = \left\{(x,a) \in \mathbb{R}^2_+ \mid a + x \leq k\right\}$ in which the probability of survival

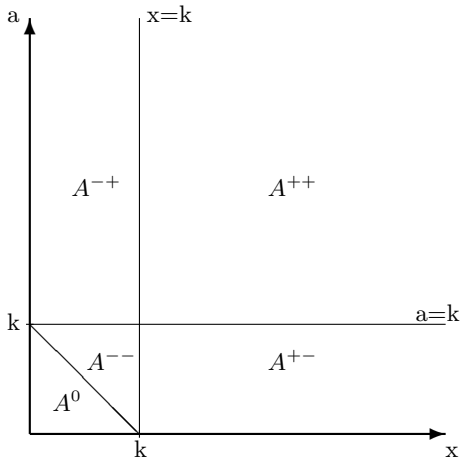

Figure 3: Five Regions

equals zero, the low-survival region $A^{--} = \left\{(x,a) \in \mathbb{R}^2_+ \mid k < a + x, a \leq k, x \leq k\right\}$ in which survival probability is positive but the consumption levels of both goods are insignificant (i.e., $a, x \leq k$), the region $A^{-+} = \left\{(x,a) \in \mathbb{R}^2_+ \mid k < a, x \leq k\right\}$ in which the consumption level of a is significant while that of x is not, the region $A^{+-} = \left\{(x,a) \in \mathbb{R}^2_+ \mid a \leq k, k < x\right\}$ in which the consumption level of x is significant while that of a is not, and the region $A^{++} = \left\{(x,a) \in \mathbb{R}^2_+ \mid k < a, k < x\right\}$ in which the consumption levels of both x and a are significant.[52] With the exception of area A^0, where the region of indifference consists of the entire area of A^0, the level curves going through regions A^{--}, A^{-+}, A^{+-} and A^{++} are, as can be verified, strictly decreasing and continuously differentiable.

For our purposes, region A^{++} will play the major role, and therefore we describe the indifference curves passing through this area in detail. In particular, within A^{++} we have hyperbolic indifference curves given by

$$x = \frac{1}{(1-q)(\beta+1)} \frac{k^{\beta+1}}{a^{\beta}}, \tag{4}$$

where the probability of survival q must lie in $\left(\frac{\beta}{\beta+1}, 1\right)$. As can be easily checked, the indifference curves are convex within areas A^{+-} and A^{++}. We show in Figure 4 the indifference curves associated with survival probabilities 0.5, 0.6, 0.7, 0.8, 0.9, 0.95 and 0.975, with parameter values $\beta = 1$ and $k = 1$.

An important property of the points lying in area A^{++} is that the slope of the indifference curves going through a fixed point $(x,a) \in A^{++}$ increases (or, stated differently, decreases in absolute value)

[52] We want to emphasize here that the regions depicted in Figure 3 do not only arise for the specific distribution functions specified by Case 1, but arise in many other cases–for instance, whenever Assumption 1 is satisfied.

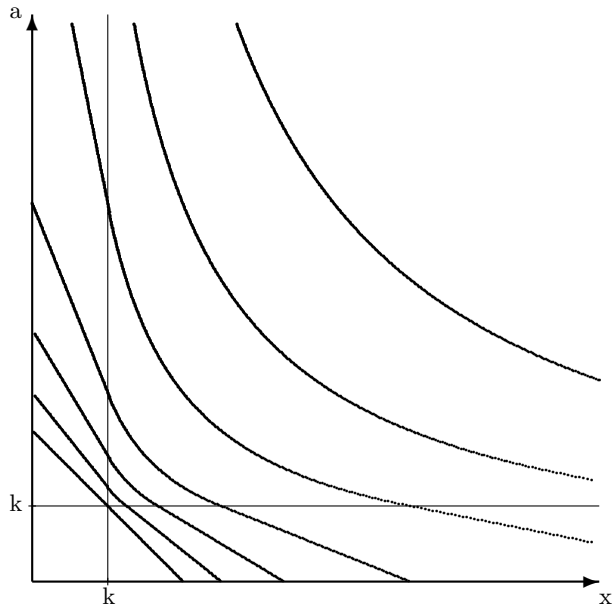

Figure 4: Indifference Curves ($\beta = 1$, $k = 1$)

as β increases. This can be verified by first calculating the following derivative in implicit form

$$\frac{da}{dx} = -\frac{\frac{\partial}{\partial x}P\left(C_x x + C_a a \geq k\right)}{\frac{\partial}{\partial a}P\left(C_x x + C_a a \geq k\right)} = -\frac{\frac{1}{1+\beta}k^{1+\beta}\frac{1}{x^2 a^{\beta}}}{\frac{1}{1+\beta}k^{1+\beta}\frac{\beta}{x a^{1+\beta}}} = -\frac{a}{\beta x}$$

and then calculating

$$\frac{\partial}{\partial \beta}\left(-\frac{a}{\beta x}\right) = \frac{a}{x\beta^2} > 0. \tag{5}$$

We can conclude by (5) that a positive cue increases the consumed amount of good a, since as we already know the indifference curves are decreasing, continuous and moreover convex within A^{++}. Thus, a positive cue causes the agent to move in the direction of increasing a along the budget line. The agent might even move into region A^{-+} (indeed, the possibility of a corner solution arises here, as indifference curves in this region can be concave for parameter values $\beta \in (0, 1)$), but this causes no problem because in this case we would have an even greater increase in the consumption of good a. ■

From (4) one can see that in A^{++} the indifference curves are Cobb-Douglas indifference curves.[53] Hence, we can formulate the following corollary:

Corollary 2 *Under the assumptions of Proposition 1, the objective function $P\left(C_x x + C_a a \geq k\right)$ is observationally equivalent to a Cobb-Douglas utility function within area A^{++}.*

[53]The indifference curves of a Cobb-Douglas utility function $U(x, a)$ take the form $x^{\gamma}a^{\delta} = q$, where $\gamma > 0$ and $\delta > 0$ are parameters and q is a constant corresponding to the level. Any function with level curves that can be represented in this way is observationally equivalent to a Cobb-Douglas utility function.

Before we can proceed with the proof of Proposition 2 we need to define the notions of ε-benefit and ε-loss formally. If the agent switches from bundle (x,a) to bundle $\left(x-\varepsilon, a+\frac{\varepsilon}{p}\right)$, where $\varepsilon>0$, he increases the survival area in (c_x,c_a)-space by

$$B:=\left\{(c_x,c_a)\in\mathbb{R}^2_+ \mid 0\leq c_x\leq \frac{k}{pa+x}, \frac{pk}{pa+\varepsilon}-c_x\frac{p\left(x-\varepsilon\right)}{pa+\varepsilon}\leq c_a\leq \frac{k}{a}-c_x\frac{x}{a}\right\},$$

while he decreases the survival area by

$$L:=\left\{(c_x,c_a)\in\mathbb{R}^2_+ \mid 0\leq c_a\leq \frac{k}{pa+x}, \frac{k}{x}-c_a\frac{a}{x}\leq c_x\leq \frac{k}{x-\varepsilon}-c_a\frac{pa+\varepsilon}{p\left(x-\varepsilon\right)}\right\}$$

(see Figure 2). We refer to the increased probability of survival attributable to the additional area B as ε-*benefit* and to the decreased probability of survival attributable to the lost area L as ε-*loss*.

Remark 1 *For any given $\varepsilon\in(0,x)$ we can calculate the ε-benefit by*

$$\int_0^{\frac{k}{pa+x}}\int_{\frac{pk}{pa+\varepsilon}-c_x\frac{p(x-\varepsilon)}{pa+\varepsilon}}^{\frac{k}{a}-c_x\frac{x}{a}} f_x\left(c_x\right)\widehat{f}_a\left(c_a\right)dc_a dc_x$$

and the ε-loss by

$$\int_0^{\frac{kp}{pa+\varepsilon}}\int_{\frac{k}{x}-c_a\frac{a}{x}}^{\frac{k}{x-\varepsilon}-c_a\frac{pa+\varepsilon}{p(x-\varepsilon)}} f_x\left(c_x\right)\widehat{f}_a\left(c_a\right)dc_x dc_a.$$

Proof of Proposition 2. The monotonicity of the objective functions $P\left(C_x x+C_a a\geq k\right)$ and $P\left(C_x x+\widehat{C}_a a\geq k\right)$ in (x,a), outside the area A^0, imply that the indifference curves associated with a positive probability of survival are strictly decreasing curves and that the agent will select in both cases a bundle lying on his budget line $x+pa=m$.

Since $P\left(C_x x+C_a a\geq k\right)=\int_k^{x+a}\int_{\max\{0,t-a\}}^{\min\{x,t\}}\frac{1}{ax}f_x\left(\frac{y}{x}\right)f_a\left(\frac{t-y}{a}\right)dydt$ is a continuous function in (x,a), the upper contour sets of the objective function are closed. This, and knowing that the indifference curves associated with the objective function $P\left(C_x x+C_a a\geq k\right)$ are strictly decreasing curves in case of positive survival probabilities, implies that these indifference curves must be continuous. For the same reasons $P\left(C_x x+\widehat{C}_a a\geq k\right)$ must also have continuous indifference curves.

If there exists an $\varepsilon\in(0,x^*)$ such that the ε-benefit exceeds the ε-loss, then

$$\begin{aligned} P\left(C_x x^*+\widehat{C}_a a^*\geq k\right) &= P\left(L\cup C\right)<P\left(B\cup C\right)= \\ &= P\left(C_x\left(x^*-\varepsilon\right)+\widehat{C}_a\left(a^*+\frac{\varepsilon}{p}\right)\geq k\right), \end{aligned} \tag{6}$$

where

$$C=\left\{(c_x,c_a)\in\mathbb{R}^2_+ \mid c_x x^*+c_a a^*\geq k \text{ and } c_x\left(x^*-\varepsilon\right)+c_a\left(a^*+\frac{\varepsilon}{p}\right)\geq k\right\},$$

and therefore, the agent increases his probability of survival by exchanging $\left(x^*-\varepsilon, a^*+\frac{\varepsilon}{p}\right)$ for (x^*,a^*). The single-crossing property imposed on the two indifference curves associated with the

no-cue balanced diet problem and the positive-cue balanced diet problem that pass through (x^*, a^*), together with equation (6) implies that the former indifference curve must lie below the latter whenever $x > x^*$. Hence, consuming less than a^* in the positive-cue balanced diet problem will be dominated by the allocation $a^* + \frac{\varepsilon}{p}$., and the solution to the positive-cue balanced diet problem necessarily occurs where $\widehat{a} > a^*$.

It remains to be shown that if there does *not* exist an $\varepsilon \in (0, x^*)$ such that the ε-benefit exceeds the ε-loss, then the solution to the positive-cue balanced diet problem will occur where $\widehat{a} \leq a^*$. Noting that this implies that expression (6) cannot be satisfied, $\widehat{a} \leq a^*$ is implied by the uniqueness of the solution $\widehat{a}$. ■

Proof of Proposition 3. First, we verify that a positive cue increases the expected probability of the occurrence of the cue, i.e., $E\Pi_t < E\Pi_{t+1}$. To see this multiply both sides of equation (3) by π and thereafter integrate both sides with respect to π to get

$$E\Pi_{t+1} = \int_0^1 \pi g_{t+1}(\pi) = \frac{\int_0^1 \pi^2 g_t(\pi)\, d\pi}{\int_0^1 \pi g_t(\pi)\, d\pi} = \frac{E\Pi_t^2}{E\Pi_t}. \tag{7}$$

Now, employing a Jensen-type inequality for strictly convex functions and for non-degenerate as well as non-negative random variables we obtain $E\Pi_t < E\Pi_{t+1}$ by (7).

For any point $(m - pa, a)$ lying on the budget line we obtain from the definitions of v_t and v_{t+1} (suppressing the arguments $(m - pa, a)$) that

$$\frac{d}{da}v_{t+1} - \frac{d}{da}v_t = (E\Pi_{t+1} - E\Pi_t)\left(\frac{\partial u}{\partial a} - p\frac{\partial u}{\partial x} - \left(\frac{\partial \overline{u}}{\partial a} - p\frac{\partial \overline{u}}{\partial x}\right)\right). \tag{8}$$

Noting that the first factor of the right-hand side of (8) is positive, we now need to show that the second factor is positive. This is equivalent to demonstrating the inequality

$$\left(\frac{\frac{\partial u}{\partial a}}{\frac{\partial u}{\partial x}} - p\right)\frac{\partial u}{\partial x} > \left(\frac{\frac{\partial \overline{u}}{\partial a}}{\frac{\partial \overline{u}}{\partial x}} - p\right)\frac{\partial \overline{u}}{\partial x} \tag{9}$$

at $(m - pa, a)$. We have $\frac{\partial u}{\partial x}(m - pa, a) > 0$ and $\frac{\partial \overline{u}}{\partial x}(m - pa, a) > 0$ by the monotonicity of the utility functions outside the death zone A^0.

Since we assumed that the optimal solutions associated with the no-cue balanced diet problem and the positive-cue balanced diet problem have unique solutions in A^{++}, we get $\left(x^{\overline{c}}, a^{\overline{c}}\right)$ and (x^c, a^c) by solving the respective first-order conditions. Hence,

$$\frac{\frac{\partial u}{\partial a}(x^c, a^c)}{\frac{\partial u}{\partial x}(x^c, a^c)} = p \quad \text{and} \quad \frac{\frac{\partial \overline{u}}{\partial a}\left(x^{\overline{c}}, a^{\overline{c}}\right)}{\frac{\partial \overline{u}}{\partial x}(x^{\overline{c}}, a^{\overline{c}})} = p, \tag{10}$$

where $x^c = m - pa^c$ and $x^{\overline{c}} = m - pa^{\overline{c}}$ by the monotonicity of the objective functions. Therefore, as illustrated in Figure 5, it follows from strict concavity of the objective functions that $\frac{\partial u}{\partial a}/\frac{\partial u}{\partial x} > p$ if $k < a < a^c$, $\frac{\partial u}{\partial a}/\frac{\partial u}{\partial x} < p$ if $a > a^c$, $\frac{\partial \overline{u}}{\partial a}/\frac{\partial \overline{u}}{\partial x} > p$ if $k < a < a^{\overline{c}}$ and $\frac{\partial \overline{u}}{\partial a}/\frac{\partial \overline{u}}{\partial x} < p$ if $a > a^{\overline{c}}$ at a point $(m - pa, a)$. Thus, (9) is satisfied for all $a \in \left[a^{\overline{c}}, a^c\right]$.

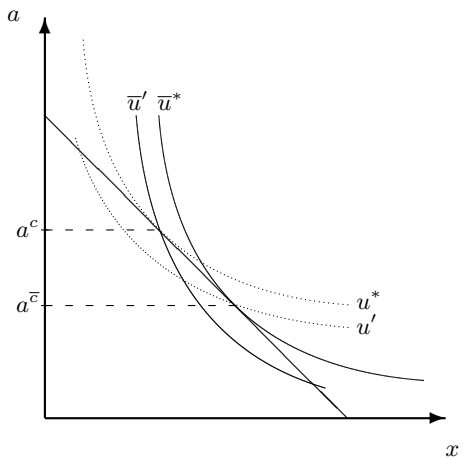

Figure 5: Intermediate Optimum

Finally, we need to verify our statement on the solution (x^*, a^*) of problem (2). Observe that the optimal solution (x^*, a^*) at time t must lie on the budget line by the monotonicity of the objective function of (2). Hence, $x^* = m - pa^*$. Since u and $\overline{u}$ are strictly concave in A^{++}, it follows that v_t is also strictly concave in A^{++}. Therefore, and by the assumption that the optimal solution associated with problem (2) lies also in A^{++}, the first-order condition

$$E\Pi_t \left(\frac{\frac{\partial u}{\partial a}}{\frac{\partial u}{\partial x}} - p \right) \frac{\partial u}{\partial x} + (1 - E\Pi_t) \left(\frac{\frac{\partial \overline{u}}{\partial a}}{\frac{\partial \overline{u}}{\partial x}} - p \right) \frac{\partial \overline{u}}{\partial x} = 0$$

of problem (2) determines $(x^*, a^*) = (m - pa^*, a^*)$ and we see that

$$\frac{\frac{\partial u}{\partial a}}{\frac{\partial u}{\partial x}} - p \quad \text{and} \quad \frac{\frac{\partial \overline{u}}{\partial a}}{\frac{\partial \overline{u}}{\partial x}} - p \tag{11}$$

cannot have the same sign.[54] However, this implies–as illustrated in Figure 5–that $a^* \in \left(a^{\overline{c}}, a^c\right)$, because otherwise the two expressions in (11) would have identical signs. ■

Proof of Proposition 4. Suppose that the agent maximizes his subjective function $\widetilde{v}_t$ instead of v_t in problem (2). Observe that for the subjective function $\widetilde{v}_t$ we can apply Proposition 3. Hence, subjective adjacent complementarity is satisfied at $(\widetilde{x}_t, \widetilde{a}_t)$ and we must have $0 = \frac{d}{da}\widetilde{v}_t\left(m - p\widetilde{a}_t, \widetilde{a}_t\right) < \frac{d}{da}\widetilde{v}_{t+1}\left(m - p\widetilde{a}_t, \widetilde{a}_t\right)$. Now since $\widetilde{a}_{t+1}$ is the unique solution of $\frac{d}{da}\widetilde{v}_{t+1}\left(m - pa, a\right) = 0$ within the region $x, a > k$, it follows that $\widetilde{a}_t < \widetilde{a}_{t+1}$. It remains to be shown that the agent's true expected survival probability given by $v_t\left(m - p\widetilde{a}_t, \widetilde{a}_t\right)$ decreases strictly in time. Clearly, the agent is moving along the budget line farther and farther away from his true expected survival maximizing bundle $(x^{\overline{c}}, a^{\overline{c}})$ because $\widetilde{a}_t$ increases. This implies by the strict concavity of $v_t = \overline{u}$ that $v_t\left(\widetilde{x}_t, \widetilde{a}_t\right) > v_{t+1}\left(\widetilde{x}_{t+1}, \widetilde{a}_{t+1}\right)$. ■

[54] Observe that the two terms in (11) cannot equal zero at the same point on the budget line because of $a^{\overline{c}} < a^c$ and (10).

References

[1] Ainslie, George "Derivation of 'Rational' Economic Behavior from Hyperbolic Discount Curves," *American Economic Review*, May 1991, *81(2)*, 334-340.

[2] Atkinson, R.L., "Naloxone Decreases Food Intake in Obese Humans," *Journal of Clinical Endocrinology and Metabolism*, 1982, *55*, 196-198.

[3] Baldwin, B.A., and R.F. Parrot, "Effects of Intracerebroventricular Injection of Naloxone on Operant Feeding and Drinking in Pigs," *Pharmacology Biochemistry and Behavior*, 1985, *22*, 37-40.

[4] Bardo, M.T., J.K. Rowlett, and M.J. Harris, "Conditioned Place Preference Using Opiate and Stimulant Drugs: A Meta-Analysis," *Neuroscience and Biobehavioral Reviews*, 1995, *19*, 39-51.

[5] Baron-Cohen, Simon, *Mindblindness: An Essay on Autism and Theory of Mind*, Cambridge MA: MIT Press, 1995.

[6] Beach, H.D., "Morphine Addiction in Rats," *Canadian Journal of Psychology*, 1957, *11*, 104-112.

[7] Becker, Gary S., Michael Grossman, and Kevin M. Murphy, "An Empirical Analysis of Cigarette Addiction," *American Economic Review*, June 1994, *84*, 396-418.

[8] Becker, Gary S., and Kevin M. Murphy, "A Theory of Rational Addiction," *Journal of Political Economy*, 1988, *96(4)*, 675-700.

[9] Belluzzi, J.D., and L. Stein, "Enkephalin May Mediate Euphoria and Drive-Reduction Reward," *Nature*, 1977, *266*, 556-558.

[10] Bergstrom, Ted, "Economics in a Family Way," *Journal of Economic Literature*, December 1996, *34(4)*, 1903-1934.

[11] Bernheim, B. Douglas, and Antonio Rangel, "Addiction and Cue-Conditioned Cognitive Processes," *NBER Working Paper No. w9329*, November 2002.

[12] Blass, E.M., "Functional Interaction Between the Positive Effect of Sweet and the Negative Effect of Pain and Distress," *Appetite*, 1986, *7*, 243.

[13] Blass, E.M., E. Fitzgerald, and P. Kehoe, "Interactions Between Sucrose, Pain and Isolation Distress," *Pharmacology Biochemistry and Behavior*, 1987, *26*, 483-489.

[14] Blass, E.M., and L.B. Hoffmeyer, "Sucrose as an Analgesic for Newborn Infants," *Pediatrics*, 1991, *87*, 215-218.

[15] Bodnar, Richard J., and Maria M. Hadjimarkou, "Endogenous Opiates and Behavior: 2001," *Peptides*, 2002, *23*, 2307-2365.

[16] Booth, Martin, *Opium: A History*, London: Simon & Schuster, Ltd., 1996.

[17] Brick, John, and Carlton Erickson, *Drugs, the Brain, and Behavior: The Pharmacology of Abuse and Dependence*, Binghamton, New York: Haworth Medical Press, 1998.

[18] Camerer, Colin, Samuel Issacharoff, George Loewenstein, Ted O'Donoghue, and Matthew Rabin, "Regulation for Conservatives: Behavioral Economics and the Case for 'Asymmetric Paternalism,"' *University of Pennsylvania Law Review*, 2003, *151*, 1211-1254.

[19] Campbell, M., P. Adams, A.M. Small, L.M. Yesch, and E.L. Curren, "Naltrexone in Infantile Autism," *Psychopharmacological Bulletin*, 1988, *24*, 135-139.

[20] Carson-DeWitt, Rosalyn, ed., *Encyclopedia of Drugs, Alcohol, and Addictive Behavior*, 2nd edition, New York: Macmillan Reference USA, 2001.

[21] Cohen, M.R., R.M. Cohen, D. Pickar, and L. Murphy, "Naloxone Reduces Food Intake in Humans," *Psychosomatic Medicine*, 1985, *47*, 132-138.

[22] Colantuoni, Carlo, Pedro Rada, Joseph McCarthy, Caroline Patten, Nicole M. Avena, Andrew Chadeayne, and Bartley G. Hoebel, "Evidence that Intermittent, Excessive Sugar Intake Causes Endogenous Opioid Dependence," *Obesity Research*, 2002, *10*, 478-488.

[23] Cooper, Jack R., Floyd E. Bloom, and Robert H. Roth, *The Biochemical Basis of Neuropharmacology*, 8th edition, Oxford University Press, 2002.

[24] Cooper, S.J., and S. Higgs, "Neuropharmacology of Appetite and Taste Preferences," in C.K. Legg and D.A. Booth, eds., *Appetite, Neural and Behavioral Bases*, New York: Oxford University Press, 1994, 212-242.

[25] Cooper, S.J., and S. Turkish, "Effects of Naltrexone on Food Preference and Concurrent Behavioral Responses in Food-Deprived Rats," *Pharmacology Biochemistry and Behavior*, 1989, *33*, 17-20.

[26] Cummings, David E., and Michael W. Schwartz, "Genetics and Pathophysiology of Human Obesity," *Annual Review of Medicine*, 2003, *54*, 453-471.

[27] Di Chiara, Gaetano, "Drug Addiction as a Dopamine-Dependent Associative Learning Disorder," *European Journal of Pharmacology*, 1999, *375*, 13-30.

[28] Doyle, T.G., K.C. Berridge, and B.A. Gosnell, "Morphine Enhances Taste Palatability in Rats," *Pharmacology Biochemistry and Behavior*, 1993, *46*, 745-749.

[29] Drewnowski, A., D.D. Krahn, M.A. Demitrack, K. Nairn, and B.A. Gosnell, "Taste Responses and Preferences for Sweet High-Fat Foods: Evidence for Opioid Involvement," *Physiology and Behavior*, 1992, *51*, 371-379.

[30] Dudley, Robert, "Evolutionary Origins of Human Alcoholism in Primate Frugivory," *The Quarterly Review of Biology*, March 2000, *75(1)*, 3-15.

[31] Dum, J., C. Gramsch, and A. Herz, "Activation of Hypothalamic β-Endorphin Pools by Reward Induced by Highly Palatable Food," *Pharmacology Biochemistry and Behavior*, 1983, *18*, 443-447.

[32] Fairburn, C.G., A. Stein, and B. Jones, "Eating Habits and Eating Disorders During Pregnancy," *Psychosomatic Medicine*, 1992, *54*, 665-672.

[33] Fantino, M.J., J. Hosotte, and M. Apfelbaum, "An Opioid Antagonist, Naltrexone, Reduces Preference for Sucrose in Humans," *American Journal of Physiology*, 1986, *252*, R91-R96.

[34] Fehr, Ernst, and Peter K. Zych, "Do Addicts Behave Rationally?" *Scandinavian Journal of Economics*, 1998, *100(3)*, 643-662.

[35] Fessler, Daniel M.T., "Reproductive Immunosuppression and Diet: An Evolutionary Perspective on Pregnancy Sickness and Meat Consumption," *Current Anthropology*, February 2002, *43(1)*, 19-61.

[36] Flaxman, S.M., and P.W. Sherman, "Morning Sickness: A Mechanism for Protecting Mother and Embryo," *Quarterly Review of Biology*, 2000, *75*, 113-148.

[37] Foster, J.A., M. Morrison, S.J. Dean, M. Hill, and H. Frenk, "Naloxone Suppresses Food/Water Consumption in the Deprived Cat," *Pharmacology Biochemistry and Behavior*, 1981, *14*, 419-421.

[38] Garcia, John, "Relation of Cue to Consequence in Avoidance Learning," *Psychonomic Science*, 1966, *4*, 123-124.

[39] Gavériaux-Ruff, Claire, and Brigitte Kieffer, "Opioid Receptors: Gene Structure and Function," in C. Stein, ed., *Opioids in Pain Control: Basic and Clinical Aspects*, Cambridge: Cambridge University Press, 1999, 1-20.

[40] Gavériaux-Ruff, C., and B.L. Kieffer, "Opioid Receptor Genes Inactivated in Mice: The Highlights," *Neuropeptides*, 2002, *36(2-3)*, 62-71.

[41] Gilbertson, T.A., and I. Kim, "Chemoreception of Fat," *Chemistry of Taste: Mechanisms, Behaviors, and Mimics,* ACS Symposium Series, 2002, *825*, 180-191.

[42] Gillberg, C. L. Terenius, and G. Lonnerholm, "Endorphin Activity in Childhood Psychosis," *Archives of General Psychiatry*, 1985, *42*, 780-783.

[43] Goland, R.S., S.L. Wardlow, R.I. Stark, and A.G. Frantz, "Human Plasma Beta-Endorphin During Pregnancy, Labor, and Delivery," *Journal of Clinical Endocrinology and Metabolism*, 1981, *52*, 74.

[44] Golub, M.S., J.H. Eisele, F.Y. Hwang, and H. Arbabzadeh, "Late-Pregnancy Changes in Peripheral Plasma Beta-Endorphin in Rhesus Monkeys," *Gynecology and Obstetrics Investigations*, 1989, *27*, 113-117.

[45] Gosnell, B., and A.S. Levine, "Stimulation of Ingestive Behavior by Preferential and Selective Opioid Agonists," in S.J. Cooper and P.G. Clifton, eds., *Drug Receptor Subtypes and Ingestive Behaviour*, San Diego: Academic Press, 1996, 147-166.

[46] Griffiths, R.R. and R.L. Balster, "Opioids: Similarity Between Evaluations of Subjective Effects and Animal Self-Administration Results," *Clinical Pharmacology and Therapeutics*, 1979, *25*, 611-617.

[47] Grossman, Michael, Frank J. Chaloupka, and Ismail Sirtalan, "An Empirical Analysis of Alcohol Addiction: Results from Monitoring the Future Panels," *Economic Inquiry*, January 1998, *36(1)*, 39-48.

[48] Gruber, Jonathan, and Botond Köszegi, "Is Addiction 'Rational'? Theory and Evidence," *Quarterly Journal of Economics*, November 2001, *116(4)*, 1261-1303.

[49] Gul, Faruk, and Wolfgang Pesendorfer, "A Theory of Addiction," mimeo, January 2001.

[50] Harrigan, S.E., and D.A. Downs, "Continuous Intravenous Naltrexone Effects on Morphine Self-Administration in Rhesus Monkeys," *Journal of Pharmacology and Experimental Therapeutics*, 1978, *204*, 481-486.

[51] Headlee, C.P., H.W. Coppock, and J.R. Nichols, "Apparatus and Technique Involved in a Laboratory Method of Detecting the Addictiveness of Drugs," *Journal of the American Pharmaceutical Association*, 1955, *44*, 229-231.

[52] Hirshleifer, Jack, "Economics from a Biological Viewpoint," *Journal of Law and Economics*, 1977, *20*, 1-52.

[53] Hirshleifer, Jack, "The Expanding Domain of Economics," *American Economic Review*, 1985, *75(6)*, 53-68.

[54] Holtzman, S.G., "Behavioral Effects of Separate and Combined Administration of Naloxone and d-Amphetamine," *Journal of Pharmacology and Experimental Therapy*, 1974, *189*, 51-60.

[55] Hubbell, C.L., S.A. Czirr, G.A. Hunter, C.M. Beaman, N.C. LeCann, and L.D. Reid, "Consumption of Ethanol Solution is Potentiated by Morphine and Attenuated by Naloxone Persistently Across Repeated Daily Administrations," *Alcohol*, 1986, *3*, 39-54.

[56] Jaffe, J.H., and W.R. Martin, "Opioid Analgesics and Antagonists," in A.G. Gillmina, T.W. Rall, A.S. Nies, and P.Taylor, eds., *The Pharmacological Basis of Therapeutics*, New York: Pergamon Press, 1990, 485-521.

[57] Kahneman, Daniel, Peter P. Wakker, and Rakesh Sarin, "Back to Bentham? Explorations of Experienced Utility," *Quarterly Journal of Economics*, May 1997, *112(2)*, 375-405.

[58] Kalat, James W., "Letter to the Editor: Speculations on Similarities Between Autism and Opiate Addiction," *Journal of Autism and Childhood Schizophrenia*, 1978, *8(4)*, 477-479.

[59] Kavaliers, M., M. Hirst, and G.C. Teskey, "Opioid Systems and Feeding in the Slug: Similarities to and Implications for Mammalian Feeding," *Brain Research Bulletin*, 1985, *14*, 681-685.

[60] Kavaliers, M., M.A. Guglick, and M. Hirst, "Opioid Involvement in the Control of Feeding of an Insect, the American Cockroach," *Life Sciences*, 1987, *49*, 665-672.

[61] Killian, A.K., K. Bonese, and C.R. Schuster, "The Effects of Naloxone on Behavior Maintained by Cocaine and Heroin Injections in the Rhesus Monkey," *Drug and Alcohol Dependence*, 1978, *3*, 243-251.

[62] Kavaliers, M., and D.D. Colwell, "Sex Differences in Opioid and Non-Opioid Mediated Predator-Induced Analgesia," *Brain Research*, 1991, *568*, 173-177.

[63] Laibson, David, "A Cue Theory of Consumption," *Quarterly Journal of Economics*, February 2001, *116(1)*, 81-119.

[64] Laibson, David, "Golden Eggs and Hyperbolic Discounting," *Quarterly Journal of Economics*, May 1997, *112(2)*, 443-477.

[65] Leboyer, M., M.P. Bouvard, and M. Dugas, "Effects of Naltrexone on Infantile Autism," *Lancet*, 1988, *1*, 715.

[66] Leboyer, M., M.P. Bouvard, P. Lensing, and J.M. Launay, "Opioid Excess Hypothesis of Autism: A Double-Blind Study of Naltrexone," *Brain Dysfunction*, 1990, *3*, 285-298.

[67] Loewenstein, George, "Out of Control: Visceral Influences on Behavior," *Organizational Behavior and Human Decision Processes*, March 1996, *65(3)*, 272-292.

[68] Lynch, W.C., "Opiate Blockade Inhibits Saccharin Intake and Blocks Normal Preference Acquisition," *Pharmacology Biochemistry and Behavior*, 1986, *24*, 833-836.

[69] Lynch, W.C., and L. Libby, "Naloxone Suppresses Intake of Highly Preferred Saccharin Solutions in Food Deprived and Sated Rats," *Life Sciences*, 1983, *33*, 1909-1914.

[70] Martin, W.R., C.G. Wickler, C.G. Eades, and F.T. Pescor, "Tolerance to and Physical Dependence on Morphine in Rats," *Psychopharmacologia*, 1963, *4*, 247-260.

[71] Mathews, F., P. Yudkin, and A. Neil, "Influence of Maternal Nutrition on Outcome of Pregnancy: Prospective Cohort Study," *British Medical Journal*, 1999, *319*, 339-343.

[72] McKim, William A., *Drugs and Behavior: An Introduction to Behavioral Pharmacology*, Prentice Hall, 2002, 5th edition.

[73] Mello, Nancy K., and S. Stevens Negus, "Preclinical Evaluation of Pharmacotherapies for Treatment of Cocaine and Opioid Abuse Using Drug Self-Administration Procedures," *Neuropsychopharmacology*, 1996, *14(6)*, 357-424.

[74] Mercer, Michele E., and Mark D. Holder, "Food Cravings, Endogenous Opioid Peptides, and Food Intake: A Review," *Appetite*, 1997, *29*, 325-352.

[75] Morley, J.E., S. Parker, and A.S. Levine, "The Effect of Butorphanol Tartrate on Food and Water Consumption in Humans," *American Journal of Clinical Nutrition*, 1985, *42*, 1175-1178.

[76] Mucha, R.F., and S.D. Iversen, "Reinforcing Properties of Morphine and Naloxone Revealed by Conditioned Place Preferences: A Procedural Examination," *Psychopharmacology*, 1984, *82*, 241-247.

[77] Mucha, R.F., D. Van der Kooy, M. O'Shaughnessy, and P. Bucenieks, "Drug Reinforcement Studied by the Use of Place Conditioning in Rat," *Brain Research*, 1982, *243*, 91-105.

[78] National Institute on Drug Abuse, "Heroin: Abuse and Addiction," *NIDA Research Report Series, NIH Publication No. 00-4165*, September 2000.

[79] Nesse, R.M., "An Evolutionary Perspective on Substance Abuse," *Ethology and Sociobiology*, 1994, *15*, 339-348.

[80] Nesse, Randolph M., and Kent C. Berridge, "Psychoactive Drug Use in Evolutionary Perspective," *Science*, 3 October 1997, *278*, 63-66.

[81] Nelson, Randy J., *An Introduction to Behavioral Endocrinology, 2nd edition*, Sunderland, MA: Sinauer, 2000.

[82] O'Donoghue, Ted, and Matthew Rabin, "Addiction and Present-Biased Preferences," mimeo, July 9, 2002.

[83] Orphanides, Athanasios, and David Zervos, "Rational Addiction with Learning and Regret," *Journal of Political Economy*, 1995, *103(4)*, 739-758.

[84] Orphanides, Athanasios, and David Zervos, "Myopia and Addictive Behaviour," *Economic Journal*, January 1998, *108*, 75-91.

[85] Ortega, R.M, M.E. Quintas, P. Andres, R.M. Martinez, and A.M. Lopez-Sobaler, "Ascorbic Acid Levels in Maternal Milk: Differences With Respect to Ascorbic Acid Status During the Third Trimester of Pregnancy," *British Journal of Nutrition*, 1998, *79*, 431-437.

[86] Ostrowski, N.L., N. Rowland, T.L. Foley, T.L. Nelson, and L.D. Reid, "Morphine Antagonists and Consummatory Behavior," *Pharmacology Biochemistry and Behavior*, 1981, *14*, 549-559.

[87] Parfitt, Kathleen, ed., *Martindale: The Complete Drug Reference, 32nd Edition*, Taunton, Mass.: Pharmaceutical Press, 1999.

[88] Pert, Candace B., and Solomon H. Snyder, "Opiate Receptor: Demonstration in Nervous Tissue," *Science*, 1973, *179*, 1011-1014.

[89] Petraglia, F., M. Baraldi, G. Giarre, F. Facchinetti, M. Santi, A. Volpe, and A.R. Genazzani, "Opioid Peptides of the Pituitary and Hypoththalamus: Changes in Pregnant and Lactating Rats," *Journal of Endocrinology*, 1985, *105*, 239-245.

[90] Pomerleau, Cynthia S., "Co-Factors for Smoking and Evolutionary Psychobiology," *Addiction*, 1997, *92(4)*, 397-408.

[91] Raiten, Daniel J., and Thomas Massaro, "Perspectives on the Nutritional Ecology of Autistic Children," *Journal of Autism and Developmental Disorders*, 1986, *16(2)*, 133-142.

[92] Raven, Peter H., Ray F. Evert, and Susan E. Eichhorn, *Biology of Plants*, 6th edition, W.H. Freeman and Company, New York, 1999.

[93] Robson, Arthur J., "The Biological Basis of Economic Behavior," *Journal of Economic Literature*, March 2001, *39(1)*, 11-33.

[94] Rogers, Alan R., "Evolution of Time Preference by Natural Selection," *American Economic Review*, June 1994, *84(3)*, 460-481.

[95] Rudski, J.M., M.K. Grace, C.J. Billington, and A.S. Levine, "Increased Food Intake Following Morphine or MPT Administration Into the Ventral Tegmental Area of Rats," *Society for Neuroscience Abstracts*, 1992, *18*, 1427.

[96] Saksida, L.M., L.A.M. Galea, and N. Kavaliers, "Predator-Induced Opioid and Non-Opioid Mediated Analgesia in Young Meadow Voles: Sex Differences and Developmental Changes," *Brain Research*, 1993, *617*, 214-219.

[97] Samuelson, Larry, "Information-Based Relative Consumption Effects," forthcoming in *Econometrica*.

[98] Sanger, D.J., P.S. McCarthy, J.A. Lord, and C.F. Smith, "The Anorectic Properties of Opiate Antagonists," *Drugs and Developmental Research*, 1983, *3*, 137-142.

[99] Savage, Leonard J., *The Foundations of Statistics*, John Wiley & Sons, 1954.

[100] Schuckit, Marc A., "New Findings in the Genetics of Alcoholism," *Journal of the American Medical Association*, 1999, *281(20)*, 1875-1876.

[101] Seligman, Martin E.P., and Joanne L. Hager, *Biological Boundaries of Learning*, New York: Appleton-Century-Crofts, 1972.

[102] Si, E.C.C., H.U. Bryant, and G.K.W. Yim, "Opioid and Non-Opioid Components of Insulin-Induced Feeding," *Pharmacology Biochemistry and Behavior*, 1986, *24*, 899-903.

[103] Simon, E.J., J.M. Hiller, and I. Edelman, "Stereospecific Binding of the Potent Narcotic Analgesic [^{3}H] Etorphine to Rat Brain Homogenate," *Proceedings of the National Academy of Sciences*, 1973, *70*, 1947-1949.

[104] Simpkins, J.W., M.E. Michel, and R. Tuttle, "Effects of the Narcotic Antagonist, Nalmefene, on Spontaneous and Insulin-Induced Food Intake and Body Weight Gain in Male Rats," *Society for Neuroscience Abstracts*, 1985, *11*, 59.

[105] Smith, Trenton G., "Obesity and Nature's Thumbprint: How Modern Waistlines Can Inform Economic Theory," *UCSB Department of Economics, Departmental Working Paper No. 18-02*, August 23, 2002.

[106] Smith, Trenton G., "The McDonald's Equilibrium: Advertising, Empty Calories, and the Endogenous Determination of Dietary Preferences," forthcoming in *Social Choice and Welfare*.

[107] Sozou, Peter D., "On Hyperbolic Discounting and Uncertain Hazard Rates," *Proceedings of the Royal Society of London B: Biological Sciences*, October 22, 1998, *265(1409)*, 2015-2020.

[108] Spanagel, Rainier, and Friedbert Weiss, "The Dopamine Hypothesis of Reward: Past and Current Status," *Trends in Neuroscience*, 1999, *22(11)*, 521-527.

[109] Sullivan, R.J., and E.H. Hagen, "Psychotropic Substance-Seeking: Evolutionary Pathology or Adaptation?" *Addiction*, 2002, *97*, 389-400.

[110] Sullivan, Susan L., "Mammalian Chemosensory Receptors," *NeuroReport*, 21 January 2002, *13(1)*, A9-A17.

[111] ter Riet, Gerben, Anton J.M. de Craen, Anthonius de Boer, and Alphons G.H. Kessels, "Is Placebo Analgesia Mediated by Endogenous Opioids? A Systematic Review," *Pain*, 1998, *76*, 273-275.

[112] Terenius, L., "Stereospecific Interaction Between Narcotic Analgesics and a Synaptic Plasma Membrane Fraction of Rat Cerebral Cortex," *Acta Pharmacologica et Toxicologica*, 1974, *32*, 317-319.

[113] Terenius, L., and A. Wahlström, "Inhibitor(s) of Narcotic Receptor Binding in Brain Extracts and Cerebrospinal Fluid," *Acta Pharmacologica et Toxicologica*, 1974, *35 (Suppl. 1)*, 15.

[114] U.S. Department of Health and Human Services, *Reducing the Health Consequences of Smoking: 25 Years of Progress. A Report of the Surgeon General*, 1989, DHHS Publication No. (CDC) 89-8411, Rockville, Maryland.

[115] Valette, A.R., R. Desprat, J. Cros, G. Pontonniea, S. Belisle, and S. Lemaire, "Immunoreactive Dynorphin in Maternal Blood, Umbilical Vein and Amniotic Fluid," *Neuropeptides*, 1986, *7*, 124-151.

[116] Van Ree, Jan M., D. De Wied, A.F. Bradbury, E.C. Hume, D.C. Smyth, and C.R. Snell, "Induction of Tolerance to the Analgesic Action of Lipotropin C-fragment," *Nature*, 1976, *264*, 792-794.

[117] Van Ree, Jan M., Mirjam A.F.M. Gerrits, and Louk J.M.J. Vanderschuren, "Opioids, Reward, and Addiction: An Encounter of Biology, Psychology, and Medicine," *Pharmacological Reviews*, 1999, *51(2)*, 341-396.

[118] Van Ree, Jan M., Raymond J.M. Niesink, Leo Van Wolfswinkel, Nick F. Ramsey, Marleen (L.) M.W. Kornet, Wouter R. Van Furth, Louk J.M.J. Vanderschuren, Mirjam A.F.M. Gerrits, and Caroline L. Van den Berg, "Endogenous Opioids and Reward," *European Journal of Pharmacology*, 2000, *405*, 89-101.

[119] Van Ree, Jan M., D.G. Smyth, F.C. Colpaert, "Dependence Creating Properties of Lipotropin C-fragment (β-endorphin): Evidence for its Internal Control of Behavior," *Life Science*, 1979, *24*, 495-502.

[120] Waeltl, Pascale, Anthony Dickinson, and Wolfram Schultz, "Dopamine Responses Comply with Basic Assumptions of Formal Learning Theory," *Nature*, 5 July 2001, *412*, 43-48.

[121] Weeks, J.R., "Experimental Morphine Addiction: Method for Automatic Intravenous Injection in Unrestrained Rats," *Science*, 1962, *138*, 143-144.

[122] Weeks, J.R., and R.J. Collins, "Changes in Morphine Self-Administration in Rats Induced by Prostaglandin E1 and Naloxone," *Prostaglandins*, 1976, *12*, 11-19.

[123] Wei, E., and H. Loh, "Physical Dependence of Opiate-Like Peptides," *Science*, 1976, *193*, 1262-1263.

[124] Weiss, G., "Food Fantasies of Incarcerated Drug Users," *International Journal of Addictions*, 1982, *17*, 905-912.

[125] Williams, George C., and Randolph M. Nesse, "The Dawn of Darwinian Medicine," *The Quarterly Review of Biology*, March 1991, *66(1)*, 1-22.

[126] Yamamoto, T., N. Sako, and S. Maeda, "Effects of Taste Stimulation on β-Endorphin Levels in Rat Cerebrospinal Fluid and Plasma," *Physiology and Behavior*, 2000, *69*, 345-350.

[127] Yeomans, Martin R., and Richard W. Gray, "Selective Effects of Naltrexone on Food Pleasantness and Intake," *Physiology and Behavior*, 1996, *60(2)*, 439-446.

[128] Yeomans, Martin R., and Richard W. Gray, "Opioid Peptides and the Control of Human Ingestive Behavior," *Neuroscience and Biobehavioral Reviews*, 2002, *26*, 713-728.

[129] Yeomans, Martin R., and P. Wright, "Lower Pleasantness of Palatable Foods in Nalmefene-Treated Human Volunteers," *Appetite*, 1991, *16*, 249-259.

[130] Zigmond, Michael J., Floyd E. Bloom, Story C. Landis, James L. Roberts, and Larry R. Squire, eds., *Fundamental Neuroscience*, San Diego: Academic Press, 1999.