How individuals make choices explains addiction’s distinctive, non-eliminable features

Gene M. Heyman

Department of Psychology & Neuroscience, Boston College, Chestnut Hill, MA 02467, United States

ARTICLE INFO

Keywords: Addiction Substance use disorder Drug abuse Matching law Choice Remission Optimal choice Kicking the habit Going cold turkey Hitting Bottom

ABSTRACT

In keeping with the goals of this Special Issue, this paper poses the following questions: What are addiction’s non-eliminable features and can they be explained by one or more general principles? I have added the qualifier “distinctive” to these goals, as in “distinctive non-eliminable features.” The result is a highly heterogeneous list, which includes features of addiction’s natural history, such as its high remission rates, its unique idioms (e.g., “kicking the habit”), and its patented interventions, such as Alcoholics Anonymous. I show that each of these distinctive features reflects how individuals make choices. In particular, they reflect the competing claims of two basic choice processes: global maximizing of the sort assumed in introductory economics textbooks and Herrnstein’s matching law, which has empirical rather than theoretical roots. These are basic choice processes, which apply to all decision making, not just drugs and not just addicts. Nevertheless, they can result in addiction when one of the options has the capacity to undermine the value of competing interests and undermine global maximizing. Conversely, the analyses also show that the two basic choice processes combine so as to predict that addiction is a semi-stable state that is biased to resolve in favor of remission. These predictions are supported by the high rates of addiction, by the high rates of remission from addiction, and by the fact that remission is often unassisted or “spontaneous.” The analyses fail to support the idea that pathological psychological processes lead to addiction. Rather they show that addiction emerges from the interactions of normal choice processes and the behaviorally toxic effects of drugs.

1. Introduction

The editors of this Special Issue, Shane Glackin and Serge Ahmed invited the contributors to provide a “comprehensive, integrated programme of inquiry” that makes sense of the “non-eliminable” features of addiction. This is a worthy challenge, which, I believe, calls for the following two-stage response: an account of the “non-eliminable” features that distinguish addiction from other psychiatric disorders and then the development of the common thread or threads that turn the distinctive features into a coherent whole. This exercise promises to identify what is essential and unique about addiction, which, in turn, will provide researchers with the most strategic targets for further scientific progress and, similarly, provide clinicians the most strategic targets for interventions. In line with the two-step strategy, I first introduce features of addiction that distinguish it from other psychiatric disorders and then introduce the principles that explain these non-eliminable features. My thesis is that addiction stands apart from other psychiatric disorders, and that it emerges from the way in which the toxic properties of addictive drugs interact with two choice

principles and their motivational and cognitive bases. According to Herrnstein’s matching law (1970) [72] choices are based on the current values of the competing options; according to global maximizing of the sort encountered in introductory economic textbooks, choices are based on the best combination of the competing options. Matching makes the next moment better; maximizing makes the overall situation better. Together, these two principles, plus the behaviorally toxic properties of addictive drugs, explain the behavioral features of addiction—from onset to remission—the words drug users typically use to describe their experience, and the emergence of addiction specific countermeasures, such as Alcoholics Anonymous (AA).

2. The non-eliminable features of addiction that distinguish it from other psychiatric disorders

My list of distinctive features are the American Psychiatric Association’s criteria for identifying addiction (e.g., Diagnostic and Statistical Manual of Mental Disorders, DSM-IV [68] and DSM-5 [69]), the natural history of addiction (for example, the transition from first days of drug
use to persistent drug use), the important role that drugs play in addiction, addiction’s high remission rates, the factors that encourage remission, the high frequency of unassisted remission, the idioms that are associated with addiction, and two interventions that are much more closely associated with addiction than with other psychiatric disorders: AA and contingency management plans. People suffering from anxiety do not mark their recovery from the point that they “hit bottom,” but addicts often do, and individuals suffering from depression have not organized themselves into enduring, world-wide, self-help organizations, such as AA.

2.1. The DSM’s approach to addiction

The authors of the DSM avoid the word “addiction” and instead have used “drug dependence” (e.g., DSM-IV) and more recently, “substance use disorder” (DSM-5). The labels notwithstanding, the manual’s account has remained very much the same over its several revisions (III, IIIR, IV, IVR, and now 5). According to the DSM committees, addiction’s “essential” or “cardinal” feature is the persistence of drug use despite significant negative consequences. Although not explicitly stated, the authors of the DSM must mean that an addict is someone who keeps using drugs even though the negative consequences of drug use outweigh the positive consequences on balance. That is, it would make no sense to call addiction a disorder if this were not the case. Put in different words, according to the DSM committees (as well as common sense), addiction is self-destructive drug use. The DSM also lists specific signs and symptoms, such as withdrawal and tolerance. This list has varied somewhat from one revision to the next, but these changes are secondary to the essential issue: persistent, self-destructive drug use.

2.2. The natural history of addictive drug use: honeymoon period, followed by the remission/relapse seesaw, followed by remission

“Addiction” refers to a pattern of heavy drug use. A weekend binge does not make one an addict; the diagnosis requires persistent drug use over time. Likewise, a weekend of sobriety does not imply that one is no longer an addict. Remission, defined as either controlled drug use or abstinence, identifies a persistent pattern of behavior, not a discrete event. There is also a characteristic etiological trajectory. Biographies and memoirs document an initial honeymoon period in which the consequences of drug use are positive on balance (e.g., [1],[2]). During this time, direct drug effects, such as intoxication, are highly positive, and many of the indirect effects of drug use are also positive. For example, the social aspects often include friendships with other drug users and a sense of pride that can come with not being afraid to take chances for the promise of a thrill. However, drug-related problems eventually emerge, and the initial positive highs become increasingly hard to re-experience: “I didn’t quit dope, it quit me” [2,3].

As the negative consequences increase, the pattern of drug use becomes increasingly unstable. Periods of abstinence seesaw back and forth with periods of resumed use. Or to paraphrase an overused joke, “It’s easy to quit drinking, I have done it many times.” But eventually, most addicts no longer use drugs at clinically significant levels, as is discussed next.

Addiction is often referred to as a “chronic, relapsing disease.” Consequently, it is surprising that every American, nation-wide, mental health survey reported that addiction is the disorder with the highest or near highest remission rate. Given the discrepancy between what the research says and what many experts say about the nature of addiction, readers should have the opportunity to see the data for themselves. Fig. 1 shows remission rates for the three surveys that distinguished between abuse and dependence. These are the most rigorous, comprehensive epidemiological studies of psychiatric disorders ever carried out in the United States [4–8]. The researchers recruited thousands of subjects with the goal of creating study samples that reflected key demographic characteristics of the American population. Importantly, they did not limit the participants to those in treatment, as is the case for treatment follow-up studies. This is critical to establishing an accurate picture of the nature of addiction; most addicts do not make use of treatment, and those that do often suffer from additional psychiatric and medical disorders, which confound the analysis of addiction (e.g., [9,10]). I calculated the data points in the figure on the basis of tables provided by the studies. That is, the researchers did not present the remission rates directly.

On the x-axis is the study; on the y-axis is the percentage of individuals who met the DSM criteria for addiction to an illegal drug at some point in their life, but not for at least a year prior to their participation in the surveys. The percentages, reading from left to right, are 76, 83 and 81%. The average age in these studies was about 42 years old, so that the data suggest that most individuals addicted to an illicit drug had probably remitted by about age 30. More detailed accounts of age of remission support this inference [8,10]. Other key findings were that dependence on alcohol and cigarettes persisted for much longer than dependence on illicit drugs, and of the illicit drugs, dependence on marijuana persisted longest (e.g., [11]).

There are also a few smaller scale studies that recruited subjects independently of their treatment history. The participants who were addicted to illicit drugs typically remitted by about age 30 [12,75]. These reports are particularly informative because they reveal the same pattern of results as shown in Fig. 1, yet their methods were quite different. For instance, in these more targeted projects, the researchers interviewed the participants multiple times, tested them for drug use, and had secondary sources for validating the participant’s reports.

However, the natural history of addiction, like any natural history, is multifarious. According to Fig. 1, a substantial number of those who met the criteria for drug dependence did not remit by their early 40 s. Thus, the slogan “addiction is a chronic relapsing disorder” may apply to some addicts, but not most. The challenge then for researchers is an account that can explain why most but not all addicts remit.

2.3. Addictive substances and activities

Explanations of why addictions involve drugs emphasize tolerance, withdrawal, and reward. However, according to the DSM, tolerance and withdrawal are neither necessary nor sufficient conditions for addiction, and there are many substances and activities that individuals find highly rewarding that are not addictive. People love music, both listening to and playing, and to do crafts, yet there are few if any chamber music or...
knitting addicts (using DSM criteria). Note, the point here is not why a particular drug is the focus of an addiction but why addictions tend to be about drugs rather than other highly rewarding and compelling substances and activities.

2.4. Factors that inspire remission

Memoirs, interviews, historical accounts of societal level shifts in drug policy, and formal research projects provide the details as to the factors that precede and surround remitting from addiction. Despite their methodological differences, these sources tell the same story. Everyday matters having to do with family, work, health, income, legal status, romantic relationships and whatever else constitutes the business of life lead addicts to quit [13–18,74]. When addicts talk about why they quit or cut back on drugs, they bring up marital relations, children, parents, witnessing an overdose, the threat of legal sanctions, the threat of financial hardships, health risks, and new romances. The formal research projects paint the same picture (e.g., [15,16]): everyday matters are the stuff of remission. In addition, auto-biographical accounts frequently include moral judgments. Ex-addicts often say they quit drugs because they hoped to be a better parent or to earn back the respect of their own parents. This aspect of the self-report literature is concisely put by an ex-addict: “God did not put me on this earth to be using heroin” [19].

The historical record tells a similar story but on a larger scale. From 1898–1914 heroin was a legal drug in the United States, distributed by the Bayer Company of Germany, and sold in pharmacies and mail-order catalogs. In 1914 the US Congress passed a bill that restricted access to heroin (the Harrison Narcotics Tax Act), which, after a series of court cases, had the effect of eliminating legal access to all opiates. According to the few drug researchers of the time, this shift in policy (and attitudes) reduced addiction rates by about 50 % [20] and see [21]. About 50 years later this “natural experiment” was repeated. Enlistees in Vietnam had easy access to cheap, high quality heroin. About 46 % reported experimenting with heroin and other opiates, and about half of those who gave these drugs a try went on to become regular users, who after a few months identified themselves as addicted [12]. However, upon return to the US, 88 % of the enlistees who met the criteria for dependence when in Vietnam remitted. This was not simply a question of access. According to interviews, many did use heroin from time to time after returning home [12], but the “sordid” conditions that surrounded heroin in the States dissuaded them from regular use. Thus, in the late teens and the early 70 s of the last century, the possibility of legal sanctions and anti-drug social mores decreased opiate use in regular users, who suffered withdrawal symptoms, and described themselves as addicted.

The memoirs and historical episodes show that the persistence of addiction varies systematically as a function of societal attitudes and the pressures and demands of everyday life. Clinics and drug courts have put these facts to use. Treatment interventions based on the idea that addicts want to quit were conceived from the outset to be beyond their early 40 s. But, also recall that Fig. 1 shows that about 20 % of those who met the DSM criteria for addiction had yet to remit. Some had to be beyond their early 40 s—the average age of the analytic sample. Hence, a successful account of addiction should explain why not all addicts remit.

2.6. Addiction’s idioms

The experience of addiction has produced a handful of well-known idioms, such as “kicking the habit” and “going cold turkey.” These are unique to addiction and necessarily reflect the experience of addicts, as opposed to those who study addicts. Thus, they offer a chance to understand addiction as it is understood by those who experience it.

2.7. Addiction specific interventions

Self-help groups, such as AA, and contingency management interventions that offer rewards in exchange for symptom free tests have played much more of a role in addiction than in other psychiatric disorders. Importantly, alcoholics, not clinicians, created AA. Thus, like the idioms, AA is particularly useful for understanding what distinguishes addiction from other psychiatric disorders.

2.8. A summary of what needs to be explained

The list of non-eliminable features includes the following features:
(1) The persistence of drug use despite mounting negative consequences.
(2) An etiologic trajectory that often begins with an initial “honey-moon period,” followed by a period of ambivalent drug use, followed by lifelong controlled use or abstinence.
(3) High remission rates, with remission often taking place without professional assistance.
(4) Symptoms that succumb to everyday pressures and moral concerns.
(5) A unique set of idioms, which reflects how addicts talk about their experience.
(6) Interventions, which, like the idioms, owe more to the experience of addiction than to the professional concerns of those who study and/or treat addiction.

Addiction is the only psychiatric disorder that has this profile. Depression, anxiety, and schizophrenia do not start off with a honey-moon period, nor does the threat of jail bring about a cure. Yet, just as those suffering from depression do not want to be depressed, at some point most addicts do not want to be addicts—that is, they usually remit. Thus, the challenge is to explain voluntary yet self-destructive behavior, to explain actions that persist despite their long-term costs, and to also explain how a pattern of behavior that appears compulsive, nevertheless voluntarily ends. But, also recall that Fig. 1 shows that about 20 % of those who met the DSM criteria for addiction had yet to remit. Some had to be beyond their early 40 s—the average age of the analytic sample. Hence, a successful account of addiction should explain why not all addicts remit.

3. Choice and the matching law

The point of Section 2 was to provide an overview of addiction’s distinctive features. In Section 4, I make the case that these features reflect how individuals make choices and the behaviorally toxic effects of drugs. In this section, I introduce the discussion of choice, focusing on
Herrnstein’s matching law and its relationship to overall reward maximizing. Both principles are necessary for describing the essential role of choice in addiction, and, put more generally, both are necessary for understanding human voluntary action. That is, one of the corollaries of the analyses presented in this essay is that ordinary, non-pathological choice processes can yield seriously suboptimal, self-destructive behavior, particularly when one of the options is a behaviorally toxic drug.

According to the matching law, the overall distribution of choices approximates the overall distribution of rewards. For instance, if there are just two options, as is the case in many experimental studies, then:

\[ B_1/(B_1 + B_2) = R_1/(R_1 + R_2), \]

where \( B_i \) is the frequency of choices 1 and 2, and \( R_i \) is the frequency of the rewards associated with choices 1 and 2. Importantly, Eq. 1 is an empirical generalization, based on observation, not theory. In contrast, until relatively recently, economists who were interested in individual choice typically began with the assumption that consumers choose rationally and then derived the functions that maximized the utility of the available options (e.g., [27,71]). First, I provide the empirical facts and logic that connect the matching law to self-destructive yet voluntary behavior; then I address the distinctive “non-eliminable” features listed above.

3.1. The generality of the matching law

The first reference to Eq. 1 as a “law” that I could find appears in a paper by Baum and Rachlin in 1969 [70]. However, in 1969, Eq. 1 had been tested extensively only with pigeons in Skinner boxes. Nevertheless, Baum and Rachlin’s claim proved true [70]. Eq. 1’s domain now extends well beyond pigeons in experimental chambers. The species that have served as subjects included rats, humans, monkeys, and even cows and coyotes (e.g., [28–31]). The settings included, but are not restricted to, foraging in natural habitats [32], football stadiums (the decision to go for a one-point or two-point conversion, [33]), baseball diamonds (pitch selection, [34]), basketball courts (two- vs. three-point shots, [35]), and chess tournaments (the decision to use the Queen’s gambit opening, [36]). The rewards included, but are not restricted to, food, money, brain stimulation, drugs, alcohol, and social approval (e.g., [28, 37, 38]). In the experiments just listed, the options were invariably substances and activities that had positive short- and long-term effects. There are also experiments in which one of the options was an activity that had harmful effects. In studies in which institutionalized patients served as subjects, the matching law described the relative frequencies of disruptive behavior [39,40], and in a study in which “at-risk” teenage girls served as subjects, a version of Eq. 1 described the absolute frequencies with which the girls practiced safe and risky sex [41]. That the matching law’s domain also includes activities that are desirable in the short-run but self-destructive in the long run suggests that it may also help explain self-destructive drug use. Finally, a recent study showed that the matching law also described covert choice processes in a visual attention task [42]. We arranged conditions so that the subject’s gaze remained fixed but attention shifted back and forth between two small stimuli. The matching law described the correlation between the relative amount of attention paid to each stimulus and the relative likelihood that the stimulus provided the information necessary for a correct response.

Fig. 2 provides a selective summary of these results. On the x-axis is the relative frequency of “reward”; on the y-axis is the relative frequency of choice. The rewards include socially mediated approval (e.g., nods of agreement), food, brain stimulation, money, and in the case of the attention allocation experiment, the subject’s judgement as to whether he or she made a correct response (there was no overt feedback).

Under most, if not all, conditions, choice and the values that guide choice interact in complex ways. Typically, an item or activity’s reward value decreases as a function of how frequently it is chosen; conversely, these same items and activities increase in reward value as time spent doing something else increases. There are, of course, exceptions: up to a point, some activities and commodities, such as potato chips, increase in value with each “bite,” and for many individuals, skilled activities increase in value as time spent doing them increases (e.g., devoted amateur and professional musicians). Moreover, since value does not increase without limit, these relations are necessarily non-linear. These complexities challenge, if not preclude, the capacity to learn the optimal allocation of choices, even for activities that are repeated frequently. Into this breach steps the matching law. It provides a simple choice rule, which, given that choice is proportional to reward, has some chance of approximating the optimal choice allocation. In support of this point, highly reliable mathematical models of choice experiments which arranged rewards according to basic schedules of reinforcement, such as variable-interval and variable-ratio, showed that matching approximated but did not equal the optimal distribution of choices over a moderately wide range of conditions [43,44]. Thus, the matching law is a simple “rule of thumb” for choice, which does a reasonably good job navigating the complexities of making the right choice in settings in which choice and its consequences interact dynamically.

However, as is inherent in rules of thumb, the matching law is blind to conditions that violate typical conditions. Figs. 3, 4, and 5, below, highlight the matching law’s shortcomings. Fig. 3 is for a non-drug case; Figs. 4 and 5 model choice when one of the options is an addictive drug. Together, the three figures make several general points: (1) the limitations of matching law guided choice are not unique to drugs and addiction, (2) when one of the options is an addictive drug or shares many of the properties of addictive drugs, the matching law’s blind spots can yield disastrous results, as measured by the overall rates of return.
(3) and, to varying degrees, voluntary behavior is suboptimal. In other words, addiction, is an extreme example of the limitations inherent to all voluntary action.

3.3. Chinese meals vs. Italian meals

Imagine two people with identical food preferences who eat out each night at either of two restaurants: Chinese or Italian. Assume that at each restaurant the value of a meal declines linearly as a function of how often it is frequented, that there is an initial bias in favor of Chinese food, and that for both of these individuals the value of a meal reflects its frequency as calculated over the most recent ten meals. However, our two hypothetical subjects differ in how they frame their eating-out options. Subject A considers eating out in “either or” terms. He chooses which restaurant to go to a day at a time, choosing the one which at the moment of choice has the higher value. The left panel shows this approach; it is referred to as “local bookkeeping.” Subject B frames her choices in terms of what is the best, overall eating-out policy. She asks, “which combination of meals provides the most enjoyable dining experience?” For example, which would be more pleasing: 6 Chinese and 4 Italian meals or 4 Chinese and 6 Italian meals? The right panel shows this approach and identifies it as “global bookkeeping.”

The sloping straight lines of the left panel plot a possible version of these assumptions for the local bookkeeper. The crossing point identifies the overall distribution of choices that the local bookkeeping strategy produces. For instance, notice that if the number of Chinese meals exceeds those identified by the crossing point, the current value of an Italian is greater than the current value of a Chinese meal, so that the local bookkeeper would go to the Italian restaurant next. Similarly, if the number of Italian meals exceeds those identified by the crossing point, the local bookkeeper will choose to eat at the Chinese restaurant next. Thus, the overall distribution of choices gravitates to the crossing point, which in this example is 6.6 Chinese meals for every 3.4 Italian meals, on average. Herrnstein and Vaughan [45] labeled this approach to choice “melioration.” The term emphasizes the motivational dimension of choice: to make things better.

The crossing point is also the choice proportion that satisfies the matching law. For example, if choices are allocated so that the average value of each option is approximately the same, which is what happens at the crossing point, it is necessarily the case that the overall amount of reward obtained from each option is proportional to the number of times each option is chosen, which is exactly what Eq. 1 says should happen. Thus, matching and local bookkeeping go hand in hand. If you are a local bookkeeper, and you always choose the option which at that moment has the higher value, your choices necessarily obey the matching law. Or put another way, local bookkeeping yields the choice...
equilibrium identified by the crossing point, which is also the distribution of choices specified by the matching law.

The curve in the right panel of Fig. 3 traces out the value of every possible ten-meal combination of Italian and Chinese meals: 10 Italian and 0 Chinese meals, 9 Italian and 1 Chinese meal, and so on (using the same value functions as in the left panel). The highest point on this curve is the combination of Italian and Chinese meals which offers the most pleasurable dining experience, which for the value functions in this example, is four Chinese meals for every six Italian meals, on average. Notice that this yields more overall enjoyment than the matching law, local bookkeeping equilibrium, even though local bookkeeping means you always choose the option that at the moment has the highest value. Also notice that not only do global and local bookkeeping produce different choice equilibriums, they are also opposite. The global bookkeeper goes to the Italian restaurant more frequently than the Chinese restaurant (although she has an initial bias for Chinese food), whereas the local bookkeeper goes to the Chinese restaurant more frequently (thereby realizing his initial bias).

The two graphs provide the following insights into voluntary behavior. First, there is more than one way to frame the available options. Individuals can choose between competing items or between competing bundles, composed of different combinations of each of the items. The global bookkeeping approach takes into account the manner in which the current choice affects the future value of each option, whereas local, item-by-item, bookkeeping ignores this critical dimension of choice. Second, local bookkeeping implies the matching law. Third, local and global bookkeeping can yield different overall choice frequencies holding all else constant. Fourth, the matching law choice equilibrium produces a lower overall rate of return. Fifth, the matching law inefficiency involved taking too many meals at the Chinese restaurant. Put more generally, the matching law choice equilibrium implies consuming too much of the favored item.

An important take-away of these five observations is that to varying degrees voluntary behavior is often suboptimal. Put in psychological terms, the restaurant graph says that the motivational state that leads to sub-optimal behavior is the desire to make the next day, the next meal, the next moment, the best possible one.

4. Explaining the non-eliminable features of addiction

In what follows, local bookkeeping, global bookkeeping, and the toxic properties of drugs combine to generate the distinctive features of addiction. Notice, that in this account, principles that apply to all choices can lead to excessive drug use and addiction. In local bookkeeping, choices are motivated by the desire to make the next moment better, which yields the matching law. In global bookkeeping, choices are motivated by the desire to make one’s overall condition, say, one’s lifestyle, better. My analyses build on Herrnstein, Prelec, and Vaughan’s account of melioration [46–48], my earlier work on addiction (2009, 2013), and the relationship between the matching law and optimal choice (e.g., [43, 44]).

4.1. The persistence of drug use despite negative consequences

Fig. 4 provides a local bookkeeping/matching law analysis of how drug use leads to addiction. The x-axes show the number of days out of the most recent thirty days that the drug was chosen. The filled triangles show the values of the non-drug activity, assuming no drug choices. The
titled open triangles show the impact of drug use on the non-drug activities.

Initially, and in correspondence with the honeymoon period that often characterizes the initial period of drug use, modest levels of drug use make everything better. However, in this example, the choice proportions associated with modest levels of drug use differ markedly from the choice equilibrium called for by the matching law, hence, drug use increases. This increase drives down the value of the non-drug activities and also drives down the value of the drug (for example, tolerance to its positive hedonic effects). In Fig. 4, drug use persists despite mounting negative consequences. In other words, the graph is a “literal” translation of (1) the essential feature of addiction according to the DSM and of (2) the memoirs and interviews that recount how drug use eventually made everything worse (e.g., [2]). Thus, the matching law predicts that when one of the options is a drug that drives down the value of non-drug activities, individuals who frame their choices in terms of their current values will overconsume the drug and persist in drug use despite significant drug-triggered negative consequences. In short, local bookkeeping (which implies the motivation to make the next moment better), a rewarding behaviorally toxic substance, and the matching law choice equilibrium combine to produce addiction as defined by the DSM.

4.2. Why are drugs the most likely focus of an addiction?

Only a small number of the many highly rewarding substances and activities become the focus of an addiction, and these few are usually drugs (and some would argue are only drugs). Fig. 5 helps explain why. The axes are the same as in Fig. 4. The two top panels show a highly rewarding but behaviorally toxic commodity (one that undermines the value of the competing alternatives). For the toxic, rewarding commodity, the matching law equilibrium and the global, optimal equilibrium are far apart. In the bottom two panels, the highly rewarding commodity is not toxic. Indeed, as its consumption increases, the value of the initially less favored good increases. This may be a direct effect of the highly rewarding activity, for example, professional success can produce “fringe” benefits that enrich one’s personal life, and also, reduce the amount of time spent in the field that brought success, thereby functioning as a kind of brake. Under these conditions, the matching law equilibrium is not so distant from the optimal, global equilibrium. This is the usual state of affairs. For instance, we are often eager to get back to work after a vacation, or, after a few hours with a good novel, we are ready to go outside and take care of overdue chores. Stated more generally, for non-toxic highly rewarding commodities and activities, local bookkeeping does reasonably well, as measured by the overall rates of return relative to the global equilibrium. In contrast, drug use does not win honors and privileges that end up curtailing drug use.

According to Fig. 5, then, the reason drugs are more likely than other substance and activities to become the focus of an addiction is that they are particularly good at undermining the value of competing activities and interfering with global bookkeeping. The following observations support this prediction.

With the exception of cigarettes, which are discussed below, a distinguishing feature of all addictive drugs is that they are intoxicating and behaviorally toxic. Intoxication undermines the cognitive capacities necessary for global bookkeeping, and it interferes with the performance of most, if not all, non-drug activities. That is, intoxication is one of the ways in which addictive drugs reduce the value of nondrug activities. Other toxic effects include withdrawal symptoms, drug related diseases, such as hepatitis and AIDS, and the socially mediated liabilities of heavy drug use, such as the threat of job loss, legal sanctions, and disapproval. As shown in Figs. 4 and 5, a common feature of addictive drugs is that they reduce the value of competing activities, which, from a local bookkeeping perspective, invites further drug use.

The graphs, however, do not tell the whole story as to what makes a substance or activity addictive. First, unlike most other highly rewarding activities and commodities, drugs do not elicit satiation. With the exception of alcohol, milligram levels produce intoxication. Thus, there is no built in restraints on consumption other than incapacitation. In contrast, consumable non-drug rewards fill you up, and motoric non-drug rewards eventually yield to exhaustion. Second, and related to these points, as a function of their potency, addictive drugs produce psychological effects that have few, if any, non-drug substitutes. For instance, accounts of the first experiences with heroin often focus on its unique effects: “filling me up with a sensation that was like nothing I ever felt before” [1], and “[i]t was the most intense nothingness there ever was” [see 49 for reference]. Thus, intoxication can be a highly desirable state, and drugs have cornered the market on intoxication.

The graphs identify the properties that make a commodity addictive. Observation reveals that drugs are the commodities that best illustrate these properties.

4.3. Why cigarettes are addictive even though they are not intoxicating and have only recently become behaviorally toxic

Cigarettes challenge this account. They are addictive but not intoxicating, and prior to the prohibitions on smoking that emerged in the last quarter of the last century, they were not obviously behaviorally toxic. However, there is more than one way to interfere with global bookkeeping, and, likewise, there is more than one way to win the preference battle.

First, consider whether it is even possible to calculate the costs and benefits of smoking. The beneficial effects of smoking are nearly instantaneous, whereas smoking-induced diseases take years to emerge and do so with a probability well below 1.0. Moreover, individuals heavily discount future and probabilistic penalties [50]. Under these circumstances, it is difficult to tell if the subjective value of smoking is a winning or losing proposition.

Second, in the past, smoking did not have to undermine competing non-smoking activities in order to be preferred. As described elsewhere [49], cigarettes had no competition. You could smoke anywhere, at any time: on horseback, in bed, at your desk, in the car. This was possible because smoking did not displace any activities—there was nothing to displace. Smoking created its own unique niche. Thus, to win the preference battle, smoking only had to be better than doing nothing. In addition, smoking cessation induces a withdrawal state for which cigarettes are the best cure, and, according to interesting animal research (e.g., [51]), smoking, like stimulants (e.g., [52]), can enhance the reward value of correlated, salient activities and stimuli. Thus, smoking is addictive for the same reasons as other drugs: it challenges global bookkeeping and, prior to the imposition of anti-smoking legislation, there were no competing activities that were of higher value at the moment of choice.

4.4. Why do most addicts remit?

Most individuals who meet the DSM criteria for addiction remit, with those who are addicted to illegal drugs typically doing so by about age 30 [11,53]. These well-established (but not well known) findings distinguish addiction from other psychiatric disorders. They also suggest that the factors that predict remission will also distinguish addiction from other psychiatric disorders. This is tested next.

The graphs show that when one of the options is behaviorally toxic, the matching law choice equilibrium yields a much lower overall rate of returns than does the global, optimal choice equilibrium. This is in part what makes addiction a self-destructive pattern of voluntary behavior; and it is also what makes addiction an unstable pattern of behavior. Any change in the drug user’s situation which encourages global bookkeeping will simultaneously increase overall benefits and decrease drug use. The increase in benefits will, in turn, increase the likelihood of global bookkeeping, which will decrease drug use, and so on. Thus, the lower rate of reward associated with matching law choice equilibrium is
an abstract index of the self-destructive nature of addiction, but, simultaneously, the gap between it and the global optimum choice equilibrium provides the motivational incentive for reducing drug use.

However, an unstable choice equilibrium is not in itself a sufficient condition for change. If remission is common then the agents that bring about change should be common as well. As described in Section 2, one of the common themes of the literature on “natural recovery” is that remission often comes about because of familial relations, employment, health, and the everyday pressures that influence everyday choices (e.g., [15,16]). Thus, Figs. 4 and 5 plus the research on natural recovery combine to predict that remission rates for addiction should be high. As shown in Fig. 1, this is exactly what the major epidemiological surveys found.

The logic that explains addiction’s high remission rates also explains why remission so often takes place without professional assistance. If the matching law choice equilibrium can be trumped by the global choice equilibrium, and the sort of factors that encourage global bookkeeping are not uncommon, then remission should not necessarily require a specialized intervention. This doesn’t mean that clinical interventions are without value. Figs. 4 and 5 say that clinical interventions that help enrich non-drug activities and promote global bookkeeping will catalyze remission.

4.5. Why don’t all addicts remit?

Factors that are high on the list of preventing remission are the same as those that promote remission, just opposite in sign. The correlates of dependence not ending are lower job skills, lower levels of educational attainment, lower income, younger age, and other well-known signs and symptoms of psychological and social dysfunction. This has been documented extensively elsewhere (e.g., [8,9,53,54]).

4.6. Addiction’s unique idioms reflect the matching law/local-global bookkeeping analysis

The terms “hitting rock bottom,” “kicking the habit,” “going cold turkey” are specific to addiction. No one goes “cold turkey” from depression or diabetes. There are also addiction-specific excuses, such as “one last time” or “this is a special occasion.” These unique features promise to help identify how addiction differs from other psychiatric disorders.

The idioms “kicking the habit” and “going cold turkey” are literal references to heroin withdrawal and figurative allusions to the experience of quitting heroin all at once and on one’s own. Paralleling the idioms, Figs. 4 and 5 show that a shift in the frame of reference when making drug choices has the potential to terminate excessive drug use all at once. Closely related to the heroin idioms is the expression, “hitting bottom.” Recovered addicts, particularly alcoholics, often date the change in their relationship with drugs to having “hit bottom.” What they mean is that they foresaw that the next—and perhaps the last—step in their downward spiral was (1) the loss of everything, and that (2) if they did not change now, there would never be another opportunity to change. In effect, “hitting bottom” is taking stock of one’s life, which is a way of talking about global bookkeeping.

Relapse and other forms of backsliding are often precipitated by the verbal formulas “this is a special occasion,” or “this is the last time,” or “tomorrow I turn over a new leaf.” According to Figs. 3, 4 and 5, these are perfect excuses. On the last choice in a series of choices, the conflicting dictates of local and global choice disappear. The global perspective assumes future choices, whose values are affected by the present choice. When the choice at hand is the last of a series of choices, the only possible framework is the present, which is to say, the local perspective. Thus, if a cigarette makes the next minute better and there is but one minute left (imagine a Chicxulub-size meteorite hurtling toward earth), then both reason and motivation say you should smoke that cigarette. Of course, there may not really be a meteorite in the immediate future, but that is another matter. From the perspective of the conflicting demands set by local and global bookkeeping, “one last time” is the perfect excuse.

These verbal formulas emerged spontaneously from the experience of addiction. They nicely capture many of addiction’s unique properties, and they nicely reflect the conclusions that follow from the bookkeeping/matching law analyses of Figs. 3, 4, and 5.

4.7. Interventions

Contingency management programs are a preferred form of treatment for addiction, but have not become a dominant form of treatment for other psychiatric disorders. Their basis is the idea that addicts retain the capacity to say no to drugs and will do so if offered a better alternative than continued drug use. That such programs have flourished is strong evidence that addicts remain voluntary drug users. However, AA provides a more interesting example of the role that choice plays in addiction. Many of its principles are great examples of global bookkeeping, and its practices are great examples of social, non-drug rewards.

Like the idioms reviewed above, AA spontaneously emerged from the experience of alcoholism; it is not the off-shoot of a government program or school of psychological thought. Rather, it is the quintessential grassroots organization. Historians trace its conception to a late night kitchen table conversation in Akron, Ohio, between two dramatically self-destructive drinkers, Bob Smith, a physician, and Bill Wilson, a financial speculator. If the analyses presented in these pages are correct, AA’s doctrines and practices will address the conflicting claims of local and global bookkeeping.

Chapter 5 of Alcoholics Anonymous’s Big Book [67] provides an account of how the program works. The centerpiece is the 12-steps to recovery. Seven of the 12 explicitly state that the road to remission is to initiate a personal inventory of the various harms that drinking has caused. These include the damage the alcoholic has brought upon himself or herself and the harms that the alcoholic has inflicted on others. The other steps implicitly call for the same sort of global bookkeeping exercise. The AA message is that in order to stop drinking, the alcoholic has to (1) fully recognize that the costs of drinking have greatly outweighed its benefits, and (2) to take to heart AA’s promise that by following the AA rules, it is possible to create a new, alcohol-free life. In short, the 12-steps are global bookkeeping and the promise of a better life.

The 12-steps is what AA says. What AA does is help its members realize their goals by providing positive role models, by creating opportunities to socialize without alcohol, and by creating opportunities to help others through the sponsorship program. In effect, AA has institutionalized global bookkeeping and created social roles that encourage its practice. Although AA has long resisted scrutiny, two well-documented research reviews attest to its effectiveness [55,56]. The matching law/bookkeeping analysis of addiction leads to the prediction that a program that follows the essential lines of AA should work—evidence says that it does.

There is, of course, much more to AA than presented here. Many find the program’s emphasis on spirituality, God, and the claim that alcoholism is a disease off-putting. My focus has been on what has made AA so viable. According to its two founders, its principles and practices reflect the experiences of “hundreds of alcoholics.” According to the ideas presented in this paper, this means that AA helps its members shift from a local to a global bookkeeping perspective in regards to alcohol. As just reviewed, the 12-steps, the drug-free meetings, and the sponsorship programs do just this.
5. Objections

5.1. Is there any evidence for this account?

Two lines of evidence support the analyses. First, the analyses began with a list of seven factors that distinguish addiction from other psychiatric disorders: The bookkeeping/matching law framework successfully explained each distinctive factor. In effect, this is analogous to six replications. Second, there is experimental support from human studies. In these experiments, the subjects had two options, and the choice/reward contingencies supported two possible equilibriums: the matching law equilibrium and the global equilibrium, just as in Figs. 3-5 (e.g., [46,57]). In contrast to analogous animal studies (e.g., [58]), there was a good deal of individual variation. Some subjects gravitated toward the local bookkeeping, matching law equilibrium, some toward the optimal, global bookkeeping equilibrium. A colleague and I investigated one possible source of this variation [59]. We arranged a series of simple computer games in which global bookkeeping earned more money than local bookkeeping overall, but on any given trial, local bookkeeping earned more. Many of the participants were long-time heroin and cocaine users. Years of drug use predicted the degree to which the subjects shifted from global to local bookkeeping (toward the matching law choice equilibrium), just as predicted by Figs. 3-5 and the accompanying analyses.

However, in my view the idioms and AA’s approach to treatment provide the strongest support for the bookkeeping/matching law analyses presented in these pages. The idioms and AA are unique to addiction and emerged spontaneously from the experience of drug use. To my knowledge no other analysis of addiction even begins to predict how addicts talk about addiction or how addicts have organized themselves to deal with addiction.

5.2. But valued non-drug opportunities may be hard to access or not even exist

Implicit in the idea that addicts quit drugs voluntarily is that there is relatively easy access to non-drug activities that are more rewarding than the drugs (from the perspective of global bookkeeping, if not local bookkeeping). However, this may seem a highly unrealistic assumption given the poor economic prospects of many Americans (e.g., [60]), and the overall decline in measures of well-being over the last 40 years or so (e.g., [61]). Two points resolve the apparent contradictions. First, as heavy drug use continues, its negative consequences increase, particularly for illegal drugs. Eventually, the alternatives to drugs will provide more value than do drugs, at least from a global perspective. Second, remission only requires that life at sub-clinically significant levels of drug use is less aversive than life at clinically significant levels. The mounting negative consequences of drug use make this increasingly likely. In support of these points, alcohol and cigarettes do not carry the burdens that come with illegal activities; accordingly, dependence on these two drugs persists much longer than dependence on any of the illegal addictive drugs (e.g., [11]). Thus, this combination of findings suggests that the aversive consequences of drug use, which the DSM defines as addiction’s essential feature, is also an essential feature of remission.

5.3. But drugs change the brain

Drugs change the brain (e.g., [62,63]). Given that the brain is the organ of choice, drug consumption could change brains in ways that increase preference for the drugs relative to non-drugs. Up to a point, this would be no different than any other factor that influenced preferences. However, it is imaginable that drugs change the brain so that voluntary drug use becomes involuntary drug use.

Alan Leshner (the previous director of NIDA) made just this point. Writing in Science, he speculates (199773):

“A metaphorical switch in the brain seems to be thrown as a result of prolonged drug use. Initially, drug use is a voluntary behavior, but when that switch in thrown, the individual moves into the state of addiction characterized by compulsive drug seeking and use.”

This is the fundamental assumption of the brain disease model that has been so widely promoted by spokespersons for the federal health agencies, clinicians, researchers and the media (e.g., [64–66]). Leshner explains his point as follows: “That addiction is tied to changes in brain structure and function is what makes it fundamentally a disease.” This is a remarkably narrow understanding of the relationships between experience, the brain and behavior. Not just drugs change the brain, but any experience that alters our actions, our thoughts, and our feelings must do so by changing our brain. Obviously, the question is not whether drugs change brains, but whether they do so in ways that render drug use no longer voluntary. This is an empirical question, one that the data on remission answer. Fig. 1, the success of AA, the success of contingency management programs, the prevalence of natural recovery are all testament to the fact that individuals who meet the DSM criteria for addiction voluntarily remit from addiction, and, accordingly, are not compulsive drug users, suffering from a disease.

6. Limitations

6.1. The emotional aspects of local and global bookkeeping are missing

This account leaves much out. My claim is that a few simple principles explain the natural history of addiction and, in particular, the features that distinguish addiction from other patterns of behavior. But behavior involves the push and pull of desire, aspirations, regrets, hopes, and memory; behavior is embodied. In contrast, the graphs and their supporting descriptions are not. Imagine, a topographic map of one of the poles, it might help you form a picture of the terrain but not of how cold it felt. The missing emotions invite questions. For instance, when looking at the graphs, it is natural to wonder: “Why isn’t everyone a global bookkeeper?” However, we can similarly ask: Why isn’t everyone always virtuous? Why does anyone ever engage in behaviors that they will regret? And how can someone risk his family, career, and, in one famous case, his presidency, for a few moments of pleasure? What this essay has not conveyed is the motivational power of local bookkeeping. With little effort, a drug can make you feel better right now, right this instant. For many, nothing else has this power. Nevertheless, most people do not become drug addicts, and of those that do, most remit. The graphs show that this is possible, but the actual steps, considerations, tricks, and emotions that make up the experience of remission are absent.

6.2. Individual differences

One of the themes of research on drug use and addiction is individual differences. Most of those who have used an addictive drug do not go on to become regular users or addicted, and among those who do become addicted, how long they stay so varies widely. Many of the demographic predictors of these differences are well established, such as the inverse relationship between years of education and the duration of dependence (e.g., [4]). However, the individual differences that are correlated with local and global bookkeeping have been little investigated. For instance, assuming that such differences exist, do they include aspects of personality, such as the “Big Five” traits, measures of cognition, such as IQ, or both? These questions have yet to be explicitly pursued.

7. Concluding remarks

The guidelines for this Special Issue invited the contributors to consider addiction’s “non-eliminable” features. This is a useful exercise. It will spearhead greater understanding of addiction, which should, in
turn, help promote more strategic interventions. To the guidelines, I added the qualifier “distinctive.” Although the resulting list of features was heterogeneous, the graphs and discussion revealed that they emerged from common sources: the behaviorally toxic effects of drugs plus processes that apply to all choices, not just drug choices. To quote Pogo in Walt Kelly’s cartoon strip of the name, “We have met the enemy and he is us.”

The data and logic reviewed in this essay provide a very different view of addiction than the claim that it is a “chronic, relapsing disease.” The analyses predict that addiction is a semi-stable state that is more likely to resolve in favor of remission than of continued heavy drug use. In support of this point, most addicts remit. The logic of local and global bookkeeping, as illustrated by Figs. 3–5, and the facts of drug use, as documented in scores of behavioral studies, could not be in better agreement. The correlates of remission are the correlates of choice. This, in turn, suggests that remission without professional assistance is possible. In line with this logic, self-help groups and unassisted recovery are typical of addiction, but not of any other psychiatric disorder. Moreover, and as emphasized, the matching law and local and global bookkeeping not only explain addiction’s natural history, they also explain addiction’s unique idioms and unique interventions. To my knowledge, no other approach provides as comprehensive or as detailed an account of addiction.

Acknowledgements

I would like to thank Verna Mims, Phoebe Pott-Heyman, and Jennifer For their thoughtful suggestions and help with the preparation of this manuscript. The research was funded in part by grants from Boston College.

References


