Addiction is socially engineered exploitation of natural biological vulnerability

Don Ross

School of Society, Politics, and Ethics
University College Cork

School of Economics
University of Cape Town

Center for Economic Analysis of Risk
Georgia State University

December 2019

Prepared for special issue of *Behavioral Brain Research* on interdisciplinary integration in addiction research.

Abstract

Interdisciplinary study of addiction is facilitated by relative unification of the concept. What should be sought is not formal unification through literal analytic definition, which would undermine practical flexibility within disciplines and intervention practices. However, leading controversies around whether addiction should be conceived as a 'disease', and over whether addiction is 'chosen' behavior, are made more difficult to resolve by failure to apply philosophical reflection on these general concepts. Such reflection should be sensitive to two kinds of constraint: coherence in description of empirical, including neuroscientific, observation, and utility in framing normative goals in treatment and policy design. Following review of various interpretations of addiction, disease, and choice across contributing disciplines, it is concluded that addiction is most plausibly viewed as a disease at the scale of public health research and policy, but not personal (e.g. clinical) management and intervention. Addicts must make choices to recover, and in that respect addiction is a 'disorder of choice'. However, it is concluded that the most relevant sense of 'disorder' arises at the social rather than the personal scale.

Keywords: addiction; addiction as disease; addiction as chosen; neuroscience of addiction; public health; epidemiological models of disease; engineered addictive environments

1. Introduction

The concept of addiction is typical of elements of conceptual fields from applied, practically motivated sciences (e.g., engineering, economics, medical sciences) in being simultaneously constrained by considerations of coherence in description of empirical phenomena, and utility in framing normative goals. Such concepts tend to resist formal theoretical regimentation. But there is value in trying to make them sufficiently coherent across contexts to avoid rampant confusion.

I will argue that two main barriers have impeded convergence on an adequately coherent general understanding of addiction. The first has been a debate over whether addiction is or is not a 'disease', under conditions where there is even less clarity on what diseases are taken to be than there is about addiction. The second has been a well motivated shared concern to minimize social stigmatization and moralized blaming of addicts. Such stigmatization is widely agreed to be counterproductive where social welfare is concerned, and to increase the frequency with which people are treated cruelly. But this debate has also been muddied by careless semantics, in this instance around the concept of 'choice'.

My aim, from my perspective as a philosopher of science and a behavioral economist, is to make progress in stating more clearly what we should see ourselves as aiming to fix when we set out to design policies and strategies to reduce the severity and frequency of harmful consequences of addiction. I will proceed as follows. In Section 2 I will identify and criticize widespread ambiguities around what it means to affirm or deny that addiction is a disease. In Section 3 I will focus on similar, and related, confusions surrounding arguments over whether addictions are 'disorders of choice' (Heyman 2009). Section 4 will derive implications from these discussions for a unified scientific, clinical, and public health perspective on addiction.

2. Addiction as a disease

The currently widespread conception of addiction as a disease, by both a large part of the (Western-influenced) public and by dominant clinical and scientific research institutions,[1] has been criticized by a range of commentators (Fingarette 1989; Heyman 2009; Alexander 2010; Lewis 2015).[2] In particular, Lewis (2015) combines polite but intense polemical force with sound understanding of the neuroscience and genetics that provide the primary scientific evidence cited by disease model proponents. Lewis is not a revisionist about this evidence, and what I will say about it here broadly accords with his account. However, as with Heyman (2009), Lewis's

[1] Lewis (2015) surveys the extent of this dominance in the American context. Unambiguous assertion that addiction is a disease of the brain is attributed to the National Institute on Drug Abuse (NIDA), the American Medical Association (AMA), the American Society of Addiction Medicine (ASAM), and the National Institute of Mental Health (NIMH). See also https://www.centeronaddiction.org/what-addiction/addiction-disease . Outside the USA the front is less monolithic. The World Health Organisation (WHO) avoids both the words "disease" and "disorder", while otherwise endorsing the elements of the NIDA description; see https://www.who.int/substance_abuse/about/en/dependence_myths&facts.pdf .

[2] I cite critics who focus specifically on addiction, setting aside more radical perspectives according to which the entire idea of psychiatric *disease* is problematic (Watters 2010; Borsboom, Cramer, and Kalis 2018). In calling this perspective 'radical' I intend no skepticism about its legitimacy. Indeed, properly contextualized I have much sympathy with it. I set it aside here purely in consequence of space limitation.

rejection of the disease model is based mainly on what he sees as its pernicious clinical consequences, which are that it impairs many addicts' prospects for recovery by undermining their agency and sense of empowerment, while trading off moral stigmatization for stigmatization associated with allegedly chronic disability. Thus his argument and conclusion, which I regard as sound, are fundamentally normative. However, conceptual practice in science, even in practically motivated sciences such as medical research, is not driven *exclusively* by reference to implications for human welfare; there are also empirical constraints to be respected. Furthermore, as I will argue, welfare considerations extend more broadly than the domain of clinical practice.

Relationships between the concepts of addiction and disease are historically complex, inevitably so in light of the fact that the two concepts have separately tangled histories on their own. 'Addiction' in English originally meant 'enslavement'. Thus its gradual application to the condition of a person who is behaviorally and emotionally preoccupied with a narcotic substance encoded the idea of capture and control by an exogenous force. Similarly, many diseases, and especially diseases affecting behavioral and emotional stability, were historically regarded in all human cultures of which there is a relevant record as involving, at least in part, control by demons, offended ancestors, or enemies practising dark arts. There is thus a sense in which a 'disease model of addiction' predates, across cultures, the contemporary, scientifically inflected, concept of disease.

On the other hand, models of disease as exogenously caused by *natural* agents have been less associated with addiction. In cultures directly or indirectly influenced by Galen, many diseases were long regarded as consequences of natural miasma. There is no salient record of association between drug dependence or its associated consequences and miasma. When miasma were later supplanted in clinical and general understanding by germs, these new natural exogenous carriers of disease were likewise not hypothesized to be factors for addiction.

In the twentieth century, as the previously black box of endogenous causal dynamics of biological systems and anatomical change was opened by geneticists, immunologists, and neuroscientists, the prevailing concept of disease was widened. The founders of Alcoholics Anonymous (AA) in the 1930s became convinced that alcohol addiction was associated with allergy (White 1998), and this mistaken conviction is still reflected in recent editions of the *Big Book* circulated as an inspirational manual by some AA chapters. Subsequently, of course, vulnerability to addiction came more plausibly to be hypothesized to involve a genetic contribution, identification of which is a currently active research program (MacKillop and Munafó 2013). As will be discussed below, there is no necessary connection between such research and conceptions of addiction as a disease.

If we take some critical distance from entrenched conceptual associations, we might see it as puzzling that scientifically inflected disease models of addiction have been preoccupied almost exclusively with hypothesized *endogenous* sources. After all, alcohol addiction, opiate addiction, nicotine addiction, etc., involve patently obvious

exogenous causal factors. But equally obviously, alcohol, opiates, nicotine, etc., are not *sufficient* causal factors, since almost all adult humans are actually or potentially exposed to them, while relatively few become addicts. The hypothesis of a genetic basis, like the AA founders' conjecture of an immunological basis, posits an endogenous source of vulnerability to develop an 'addictive disease' conditional on exposure to an exogenous trigger. Still, we might wonder about the source of the assumption that vulnerability must have an endogenous basis. No one would regard it as a *conceptual* problem for the 'disease model' of skin cancer if it were, counterfactually, the case that all people were equally statistically likely to develop it conditional on extended exposure to strong sunlight. It would be regarded as uncontroversial that vulnerability to the disease was a function of variation on exogenous factors such as working outdoors and living in hot climates.

A key disanalogy between an environmentally driven cancer vulnerability and vulnerability to addiction arises in the fact that cancer symptoms, once triggered, persist unless removed by an intervention, whereas those symptoms caused *directly* by an addictive target tend to disappear unless the target is continuously administered. So, drunkenness is a symptom of alcohol addiction, but most alcohol addicts are only drunk some of the time, and would cease to be drunk at all if they stopped inflicting the cause on themselves. But addiction is taken to be a standing condition that persists during periods of sobriety; and where there are symptoms that don't fade between episodes - in some cases, for example, tremors – these are not part of the direct etiology of the other symptoms. Thus it is supposed that 'the disease itself' must reside in standing conditions that cause drinking, which in turn cause most of the salient but intermittent symptoms. Furthermore, the conditions themselves must be *non-standard* or *abnormal* states, if they are to account for the difference between an addict and a non-addict.

As Heyman (2009) emphasizes, it is crucial to the contemporary disease model of addiction as a *distinctive* model that addiction is widely held to be a *chronic* condition that persists even while an addict successfully maintains abstintence, as this seems to imply the existence of a standing endogenous cause of potential relapse. The view of addiction as chronic is a central feature of the AA ideology, acceptance of which is promoted as itself an essential part of treatment intervention: the addict, and family and friends, are to acknowledge that addiction can be managed but never cured. The immunological conjecture of AA's founders fit comfortably with this ideology, since all allergies were once chronic conditions and many still are. Furthermore, allergies are uncontroversially abnormal, in the sense of being departures from the functional contributions to survival and reproduction that immune systems were selected by natural selection to make.[3]

[3] The *non*-functional meaning of 'abnormal', statistical scarcity, is *not* the relevant concept here. An allergy would remain a dysfunction even if the majority of the population suffered from it. Conversely, many statistically rare conditions, such as mathematical or musical genius, are not conceptualized as dysfunctions.

Genetic profiles are standing structures, though their functional and biochemical expression varies with an organism's life stages and circumstances. But research on genetic correlates of addiction is not expected by most researchers to ultimately isolate a single responsible gene or cluster that represents an evolutionary 'mistake' in the sense of unequivocal dysfunctionality. The relationship between genes and complex phenotypical patterns like addictive behavior are simply not that simple (Keller 2000); and, in general, diseases can only be *identified with* genes that directly cause debilitating developmental outcomes before adulthood, since most other genes that reduce expected fitness in some circumstances promote it in others. Questions about how many different genetic configurations contribute to vulnerability to addiction, and to what extent in each case, are generating useful and potentially practically significant discoveries, but are independent of the question as to whether addiction is best conceptualized as a disease.

The most common basis for contemporary disease models of addiction among scientists and treatment professionals is appeal to characteristic neuroadaptations observed in addicts. For compactness of discussion I will discuss these as they have been revealed in neuropsychological research on gambling addicts. Ross et al (2008) argue that where the functional neuroscience of addiction is concerned, gambling addiction, the only purely behavioral dependence currently acknowledged as addictive in the *DSM 5* (American Psychiatric Association 2013), should be regarded as exemplifying the basic mechanism, because it involves no compounding exogenous neurochemical effects specific to particular ingested drugs. After reviewing the high-level neuroscience of gambling addiction, I will comment briefly on how this same general pattern is triggered by some substances. Crucial for the relevance of this discussion to the disease conception of addiction will be distinguishing between the neurofunctional structures that render people susceptible to addiction, and the neurofunctional *effects* of addictive consumption (that is, neuroadaptations).

What makes some forms of gambling – slot machines, roulette, video poker, but *not* poker played against a table of other people – potentially addictive is that the games in question involve sequences of statistically independent events that frequently yield reinforcing rewards and can be generated by simple, stereotyped actions on the part of the player. Three general facts about the architecture of the brain's reward learning mechanism turn this into a potential behavioral trap.

First, the dopamine-based neural learning circuit projecting from ventral striatum directs allocation of cognitive and emotional attention *and* makes and tests predictions of reward as a single, integrated process. Thus any actions that deliver information pertinent to adjustment of reward predictions are hard not to attend to and be aroused by.

Second, the ventral striatal circuit learns according to an algorithm, a specific form of Rescorla-Wagner learning, which cannot settle on a model of genuine randomness (Sutton and Barto 2018). Thus it cannot draw the conclusion that

further testing of reward predictions by additional gambles is not delivering any new information about estimated priors.

Third, the wider neural environment in which the striatal learning mechanism is embedded is built to help the organism efficiently harvest reward when its attention is strongly focused, its motivation is aroused, and easily executed probes of the external environment seem to be delivering steady learning opportunities. The signal for this is phasic dopamine concentration in nucleus accumbens. The brain facilitates reward harvesting by cueing motor preparation for the activity that delivers the information that is the focus of attention. Somatic cravings are the subjective experience of such motor preparation in the absence of opportunity for action. The cravings cue thoughts about gambling, and the thoughts sustain attention and motor preparation, so there is self-sustaining feedback. This generates one of the most aversive aspects of addiction, the crowding out of alternative attentional foci by thoughts about gambling. The gambling addict who is trying to concentrate on work, family, friends, or recreation is distracted by intrusive thoughts about gambling and by somatic discomfort. Hence she experiences relief when she succumbs and goes to the casino or the video terminal. This relief will tend to automatically be interpreted as escape from work or family problems – in typical cases the very problems caused in the first place by preoccupation with gambling. In extreme cases she "enters the zone" (Schüll 2012), the American gambling addicts' phrase for the phenomenology of having her behavior completely and effortlessly controlled by the dopaminergic mechanism.

Addictive drugs share the basic features of the syndrome just described (Potenza 2008). That is, all such drugs 'convince' the striatal dopaminergic system that an opportunity for efficient reward learning is at hand, that attention should therefore be focused on the opportunity and on cues that become associated with it (including thoughts about it), and all thereby 'convince' the brain as a whole to condition the motor system to prepare to harvest the reward, thus inducing cravings and preoccupation. In all cases, the basic signal triggering the syndrome is concentration of phasic dopamine in nucleus accumbens (Beninger 2018). This common pattern can be obscured if one focuses on differences among the specific pathways by which addictive drugs generate this concentration. Stimulant drugs such as cocaine and amphetamines inhibit reuptake of dopamine after it has facilitated synaptic transfer. Opiates recruit additional mechanisms, mu-opioid receptors, in the ventral tegmental area and nucleus accumbens. Alcohol inhibits the neurotransmitter GABA throughout the brain. Since GABA neurons inhibit dopaminergic neurons in ventral tegmental area, the second-order result of GABAergic inhibition is a flood of dopamine into nucleus accumbens. An alternative causal path to this inhibition promoted by some addictive drugs is reduction of tonic serotonin 5-HT in pre-frontal areas. Alcohol additionally promotes mu-opioid receptors. One way of understanding what all of these drugs are doing, in their different ways, is introducing noise into the signals by which the overall system registers the timing of reward delivery. Since neural learning involves prediction of, and adjustment in response to misprediction of, both the magnitude and the timing of rewards,

disruption of the clock against which timing is measured results in attempted learning that never resolves by identifying a stable model.

According to this account of the neural basis of addiction, we find no antecedent condition in the brain of the addict that could be regarded as a disorder. The striatal reward learning circuit was selected to associate the organism's actions with opportunities for learning about the magnitudes and timing of rewards. Confronted with a slot machine, the system responds in the way that has contributed to the success of the organism's ancestors. Of course the response is not optimal given *current* conditions, since it fails to learn that there is no regular relationship between any one play of the machine and another. But the explanation for this is straightforward: genuinely random, as opposed to merely stochastic, processes are vanishingly rare in nature, so ability to predict, and hence to detect, such processes was not selected for. Suboptimality *per se* is not indicative of disorder.

Perhaps, however, we might locate disorder in the neuroadaptation that characterizes the passage from learning to gamble to 'true' addiction. The brightest line in this learning continuum is crossed when pre-frontal signals that normally inhibit direct triggering of motor preparation cells from ventral striatum begin to attenuate. As noted, this is the basis of craving and preoccupation, which in turn make it much more difficult for the addicted person to implement deliberate decisions to abstain or cut down. Furthermore, this is the functional change that makes it possible to identify addicts by non-behavioral means, in neuroimaging probes. This is the only state of the individual addict's brain that could reasonably be regarded as disordered. Though conceptualizing the neuroadaptation as a disorder might have sound practical purposes and is arguably semantically natural, let us first note two restrictions on regarding it as the basis for attributing a disease.

First, the condition results from the brain undergoing a normal learning process, adaptive under circumstances encountered in the ancestral past, to an engineered novelty, the slot machine, designed deliberately to make the addictive neuroadaptation more likely (Schüll 2012). In this respect, the disorder closely resembles the obesity of a person whose unhealthy state results from her eating sugar and carbohydrates when they are available, just as her ancestors were selected to do, but who lives in an environment where the most inexpensive calories on offer are found in soda and junk food.

Second, we have no empirical grounds for hypothesizing a basis in genetic variation for the addict's *brain* to have been statistically unusually predisposed to undergo the neuroadaptation. If we are persuaded, following the authors in MacKillop and Munafó (2013), that the observed correlation between some genetic markers and addiction involves an intermediate phenotype, we do not know whether the risk that is genetically promoted is for gambling conditional on opportunities, or for relatively rapid and entrenched neuroadaptation conditional on gambling. For that matter, the relevant risk might be for some conceptually remote disposition that makes the person more likely to gamble, if she gambles, with slot machines than with some less dangerous instrument such as cards.

The point above does not constitute criticism of the intermediate phenotype approach to addiction vulnerability. For many public health purposes, fine details of mechanism specification are unimportant, depending on which elements of causal relationships are most amenable to effective policy intervention. The recent history of the relationship between problem gambling research and policy response is exemplary where this issue is concerned, and offers important insight into what is at stake when we decide whether to conceptualize addiction as a disease, and when we deliberate about what we might best specifically intend by such conceptualization.

A large proportion of the first generation of problem gambling research was funded by national 'responsible gambling' agencies in various countries that in turn were financed through voluntary or mandatory contributions from casinos and (sometimes) lottery operators. Relationships between these agencies and researchers have usually been effectively arm's-length, as required and overseen by university research ethics committees and institutional review boards. However, many in the research community have become convinced that this structure of research support has contributed to a specific systemic bias (Hancock and Smith 2017; Abbott 2017; Hancock 2011). 'Responsible gambling' public information campaigns promote the idea that gambling is harmless recreation for most participants, but that for a small, ex ante vulnerable, minority it is dangerous. Researchers drawn into policy conversations structured around responsible gambling thus tended to concentrate on trying to quantity and identify this at risk sub-population. Much early work consisted of national and provincial /state prevalence studies, using screens designed for clinical diagnosis such as, initially, the South Oaks Gambling Screen (SOGS) (Lesieur and Blume 1987), and, later, the Canadian Problem Gambling Severity Index (PGSI) (Ferris and Wynne 2001). As reviewed by Williams, Volberg, and Stevens (2012), these studies reported problem gambling prevalence rates in the range of 0.5% to 7.6%, but strongly skewed toward the lower end and averaging 2.3%. Such findings comported comfortably with the 'responsible gambling' framework.

Significant limitations of the framework from the public health perspective have emerged, however. Where the more rigorously calibrated PGSI is concerned, the operationalization of 'at risk' refers to the risk that a survey respondent, based on status *at the time of the survey*, would be classified as a problem gambler on the basis of a structured clinical interview. This differs from the standard public health understanding of risk, which aims to identify the distribution and frequency of *prospective* risk of development of a condition in a population. Most prevalence studies have used diagnostic-stem ('trigger') questions that allow surveyors to avoid annoying respondents who have said they do not gamble by then asking them batteries of questions about specific aspects of their gambling. However, as Harrison, Lau, and Ross (2019) (HLR) show in re-analyses of selected large prevalence studies using advanced statistical estimation methods, this apparently sensible practice creates sample selection bias that results in massive underestimation of problem gambling prevalence on the PGSI operationalization of risk, and even greater underestimation of prospective risk. The largest problem gambling prevalence

study conducted to date, incorporated in one wave of he *National Epidemiologic Survey on Alcohol and Related Conditions* (NESARC) in the United States (Kessler et al 2008), reported population prevalence of 1.3%. When HLR apply best-practice econometric sample-selection bias correction to the data, using the same conceptualization of risk, they estimate the population prevalence at 7.7%.

This analysis highlights what I regard as the most important complication facing critics who reject application of the concept of disease to addiction altogether. 'Disease' is not a concept over which clinicians and the researchers focused on serving them can claim exclusive 'ownership'. Arguably, in fact, this community has no need in the first place for a scientifically regimented concept of disease, since their mission calls upon them to apply their characteristic interventions, including pharmaceutical interventions, wherever they have evidence that these can relieve patient distress of any kind (and are in addition ethically appropriate). By contrast, public health researchers *require* rigorous models of characteristic epidemiological patterns. To the extent that a condition that reduces public health spreads by contagion in the manner typical of classic vector-borne diseases, public policy interventions can be effectively modeled epidemiologically (Brockman and Helbing 2013). Recent examples include obesity (Huang et al 2016) – and narcotic use for recreation and/or psychological self-medication (Galea, Nandi, and Vlahov 2004).

I conjecture, based on unsystematic reading, that most epidemiologists would not wish to insist that unhealthy behavioral patterns at population scale such as obesity and addiction are 'diseases' in more than a metaphorical sense. But it not clear that the relevant distinction where scientific discourse is concerned is between the metaphorical and the literal; what is arguably more important is the distinction between casual and technically specialized concepts. From that point of view, there is a philosophical case to be made that (1) the only technical regimentation of the disease concept in the scientific literature is the epidemiologist's; and that (2) if the concept is empirically identified in the best models of a phenomenon that undermines public health, then the concept will be applied. Whether some applications are called 'metaphorical' may signal nothing more than the fact that the applications in question are novel. It is the nature of the dynamics of concept innovation in science that technically applicable 'metaphors' come, as it were, pre-frozen. There is not yet a preponderance of clear evidence that any specific addiction, let alone all addictions, in fact show classic disease patterns of contagion. The point is that this open empirical prospect is likely to play a more decisive role than the normative opinions of clinicians in determining whether campaigns to stop conceptualizing addiction as a disease are ultimately successful.

Some further qualification is in order here. Notwithstanding philosophers' tendencies to want to serve as universalizing conceptual janitors, cross-disciplinary conceptual pluralism is a common phenomenon. It might turn out to be natural and sensible to say that addiction is a disease from the perspective of public health but is best not handled in the clinical context according to the standard protocols of medicalized conditions, i.e., diseases in the non-technical sense.

Further reflection on our main example of gambling addiction suggests that this relatively simple bifurcation of conceptual practice might still be too simple. A 2010 prevalence study of problem gambling in large South African cities (Sharp et al 2015) did not use a diagnostic-stem question, and thus allows for meaningful comparison of population prevalences of more and less severe manifestations of gambling problems according to the PGSI operationalization of risk. Approximately 3% of the sample were found to be at high risk for clinical diagnosis, while approximately 18% were assigned risk scores of >0. Kincaid et al (2013) performed a taxometric analysis of these data, and found that a distinct taxon emerged at the upper end of the risk distribution if PGSI scoring criteria were made more stringent, according to a rule that reduced the high-risk group to approximately 2%. The authors speculate, on the basis of secondary data from the study, that this taxon might mark the sub-population that is 'truly' addicted, in the sense of having relatively entrenched neuroadaptation to gambling and relative recalcitrance to intervention.

Let us consider the suggested pluralistic approach to application of the disease concept in light of these data. From a public health perspective, the welfare loss expected to be associated with the 19% of the 'at-risk' taxon who are conjectured to not be 'truly addicted' is likely to be substantially greater than that expected to be associated with the 'truly addicted' taxon, and so is likely to be the more efficient target for intervention if resource scarcity calls for triage at the policy level (Sulkunen et al 2019; Harrison, Lau, and Ross 2019). Furthermore, and more importantly in the present context, only the larger taxon is likely to be sufficiently widely distributed to support application of an epidemiological model. However, if a medicalized response model, particularly one involving in-patient care supported by pharmaceuticals, were to be applied anywhere it would be to people in the smaller taxon. These reflections on real data for policy certainly do not directly pull a rug out from under Lewis's view. After all, his opinion is that the disease model does not apply to addiction at the clinical and individual patient scale, period. But the case *does* lead us to recognize how *profoundly* disunified the disease concept might be in practice on the pluralistic interpretation: according to it, in the instance at hand the *most* plausible targets for application of the clinical concept of disease are precisely those to whom the public health concept of disease might *not* apply.

The general conclusion of the discussion to this point is that debates over whether addiction should be conceptualized as a disease resist settlement because disease itself is a context-sensitive and relatively unstructured concept. The most important reasons for caution in applying it to addiction involve normative assessment of the consequences for addicts, not the implications for scientific clarity. As will be seen in Section 4, however, the discussion of the disease conception, conjoined with reflections on addicts' agency in Section 3, *will* help to shed light on a less elusive and more important question, on the kind of *disorder* that addiction exemplifies, and that justifies motivation for interventions.

3. Is addictive behavior chosen?

Defenses of the conception of addiction as a disease almost always include the claim that this undermines its moralization and cultural tendencies to stigmatize and blame addicts (Lewis 2015). The alleged basis for this is that if the addict is taken to be in the grip of a disease, then she is less likely to be thought to have chosen her condition and its consequences. Reflection on unambiguous cases of diseases that typically result from voluntary behavior, such as HIV/AIDS, should remind us that the effectiveness of this argument is likely to be limited.

There are two different senses in which addicts are often held responsible for their addictions. That addicts typically choose the behavior that triggers their addiction is disputed only by skeptics about free will in general.[4] Where this is the intended sense of responsibility, the question of whether a disease model is maintained is logically irrelevant, unless the model incorporates a scientifically unsupported doctrine of strong genetic determinism specifically about addiction. Policy-relevant debate instead focuses on whether addicts choose to be addicted. The most common narrative associated with addiction as a chosen condition is that addicts make series of discrete choices to go on consuming addictive targets even when they have become aware that such choices maintain their addiction. Their continuing addiction can consequently be held to manifest weakness of willpower. In response to this, a disease model is often invoked to emphasize that paralysis of will is a non-chosen consequence of addiction, in roughly the sense that paralysis of limbs might be a consequence of a stroke. People do not suppose that stroke victims choose not to walk, even in cases where they believe that the stroke itself was caused by chosen behavior.

As with concepts of disease, both popular and scientific concepts of choice feature heterogeneity. However, whereas the concept of disease is analytically specified only in epidemiology, alternative choice concepts have been rigorously modeled within disciplines that respectively deploy them (Ross 2011). Parts of cognitive science dominated by computational psychologists and philosophers often associate choice with conscious deliberation. On this conception, addicts can straightforwardly be held not to choose addictive behavior if it is thought to be generally caused by subconscious, automatic processes. By contrast, economists (other than those behavioral economists who are distinguished precisely by their embrace of conceptual frameworks borrowed from psychology) deploy a less stringent concept of choice, according to which a course of behavior counts as chosen just in case it is modifiable by incentives. On this conception, empirical evidence speaks decisively in favor of addictive behavior as chosen: addictive consumption is clearly sensitive to changes in costs and benefits (Chaloupka 1991;

[4] Almost everyone acknowledges some common classes of exceptions: children born with addictions derived from addicted mothers, and people who become addicted due to decisions made by medical professionals.

DeGrandpere et al 1992; Leung and Phelps 1993; Carroll 1996; Bickel et al 1998; Chaloupka et al 1999; Higgins et al 2004; Heyman 2009, pp. 65-88).

Literatures on the relationship between choice and responsibility are enormous. I refer to "literatures" because there is more than one, linked to one another in the scientific network only by a few relatively isolated nodes. Economists and some philosophers have generated a large cluster of theoretical work that is preoccupied with the extent to which addictive choice satisfies application conditions for various models of objective rationality; Elster and Skog (1999) provide a representative sample of such research. It derives such policy relevance as it has from its connection to concerns about paternalism: to the extent that some behavior is regarded as potentially or actually rational, thought in the liberal tradition inclines to the view that it should be regulated only with respect to externalities it generates. I will here set these disputes aside on grounds that they are tangential to the question of how we should best conceptualize addiction in the first place, and because where policy and intervention are concerned they are directly relevant only to the relatively small subset of addicts who are content to remain addicts, or who believe that any potential modification of their behavior they might choose is no one else's business.

Fortunately there is a growing literature that interprets relationships between choice and responsibility more broadly and flexibly. In this literature, technical and a priori standards of rationality are replaced by a naturalistic understanding of intelligibility by reference to actual, varying, interpersonal and cultural standards. Pioneering work of this kind stems from Ainslie (1992, 2001), who emphasizes the typical drama of addiction as revolving around maintenance of 'personal rules'. These are principles by which people maintain levels of consistency and predictability in their choices that constitute dynamic equilibria between social expectations that are applied to them, and their own preferred points of trade-off between stability and adventure. The dynamics of personal-rule maintenance *are*, on Ainslie's account, the phenomena to which people refer when they deploy the simplified cultural construct of 'willpower'.

The pernicious aspect of addiction, Ainslie argues, is that its characteristic neurodynamics undermine addicts' confidence in their ability to maintain personal rules. A smoker, for example, might decide that smoking is no longer consistent with the sort of person she wishes to be, in light of her own passage through her life-cycle and shifting social standards of health with which she identifies. Nicotine addiction, however, might then set up unanticipated challenges for her. On one typical pattern, she might form a personal rule to become abstinent, but then discover that because of the neuroadaptations her smoking has caused, her policy requires a level of self-policing against cues for craving that amount to self-tyranny. On an alterative pattern that is also common, she might design a rule to smoke only while drinking and socializing, and then find that her addiction leads her to over-exploit this loophole and spend unsustainable time in the pub. Ainslie refers to his model of personal-rule dynamics as 'picoeconomics', because he formally frames it in terms

of alternative patterns of 'bundling' of the objects of reward (e.g. a sequence of discrete smoking experiences versus a year's self-permission to smoke) that effect curvatures of revealed intertemporal discount rates. But the picoeconomics label can be motivated more generally where addiction is concerned, simply by observing that addiction systematically and significantly shifts the relative costs of different personal rules (Bénabou and Triole 2004; Ross 2010).

Ainslie is a psychiatrist, and his work within that disciplinary frame has been relatively narrowly focused on the phenomenology of self-control, framed for incorporation into behavioral economics. However, I referred to it as 'pioneering' because a philosophical literature has subsequently developed that carefully interprets normative responsibility in terms of constructs more or less identical to Ainslie's 'personal rules' (Keefe 1996; Velleman 2009; Tiberius 2018). This emerging perspective is nicely synthesised by Ismael (2016), into what she calls a 'self-governance model of personhood' (SGP). According to the model, persons are constructed through natural social dynamics that are driven by pressures to distribute, reward, and sanction responsibility, and a human individual achieves and maintains personhood to the extent that she judges herself, and is judged by others, to govern herself well with respect to the responsibilities she acknowledges. Everyday conflicts among values ensure that such personhood is an ideal against which everyone falls short. But we treat as mental disorders standing conditions of people's brains that seriously undermine their capacity for self-governance. Addiction is a leading such disorder; and the mechanism by which it interferes with responsibility is essentially identical to that previously identified by Ainslie, though Ismael does not cite him. We can interpret the most common life-cycle of addiction in terms of picoeconomics and the SGP.

Addiction begins with escalating frequency of use and dosage sizes over periods of months or years, usually beginning in adolescence or early adulthood. The majority of drug addicts either become abstinent or develop controlled use (which, contrary to crude variants of the disease model, is not rare) in middle age (Heyman 2009). Self-engineered recovery is significantly less frequent in addicts who suffer from co-occurring psychiatric disorders such as major depression. As this group of addicts is over-represented in clinical samples, Heyman (2009) argues that the image of addiction as an intractable, chronic condition is based on the salience of convenient clinical populations in prevalence and treatment research. At the same time, controlling for age, relapses among recovered addicts lead to faster escalation of use frequency and dosage than is observed in the early stages of dependency among people without addiction histories. This indicates that the reward circuits of recovered addicts do not entirely forget their training. This is the element of truth in the otherwise false representation of addiction as a chronic disorder.

Addicts' efforts, between onset and recovery, to escape from their addictions should be framed as attempts to regain personal governance. The normal process that eventually leads to abstinence or stable and controlled consumption is one of discovery of effective personal rules. Addicts experience their initial unsuccessfully

maintained personal rules as failures that undermine their confidence in themselves. However, the high rate of recovery, which applies to the majority who never receive clinical assistance, and the fact that eventual success follows an average of seven 'serious' campaigns for control, each ending in relapse (Heyman 2009), is evidence that these failures are occasions for learning. Attempted personal rules are typically modifications of previous ones that add explicit patches to what are perceived as design flaws in the prototypes. For example, the gambling addict may try to limit casino visits to the first Friday after payday, then have the rule collapse because she finds herself framing every bonus earning, including those where she can influence the timing, as paydays. This might lead her to launch a new control campaign after she has arranged with her employer to defer all of her accrued bonuses into an annual extra payment. In general, the most effective kind of personalized assistance that can be provided to addicts is to direct their cognitive attention to the trial-and-error learning provided by their unsuccessful personal rules. This seems to eventually convince most addicts that self-governance, however diminished it has become for them, is not altogether eroded and thus has potential to be strengthened.

Personal addiction, on this account, is a consequence, grounded in neural reward learning, of sequences of choices, at least some of which are consequences of personal rules that are poorly adapted to the ecology of available rewards in the person's environment. It emerges as a disorder *with respect to social, psychological and economic expectation of self-governance.* It is not a disorder 'of the brain', in the sense that no part of the addict's brain is directly damaged *by addictive neuroadaptation* (though of course it might be damaged by alcohol, or pulmonary disease from tobacco, or the baseball bats of the gambling provider's goons), and no part of the brain is malfunctioning with respect to its evolutionary proper function.[5] Addicts typically choose to become un-addicted, and successfully implement this choice, but only through trial-and-error learning of personal rules that are better adapted to their environments. They typically rely quite heavily on assistance in this learning from family and close friends. A lucky minority get further useful help from therapists and clinicians.

4. Addiction as a disorder of society

The characterization of personal addiction just given emphasizes negative welfare effects that result from the normal functioning of the reward learning systems of the brain in certain modern environments. In closing I will direct attention to these environments.

All mammals share similar neural reward circuits, so all are in principle vulnerable to addictive neuroadaptation, though non-humans cannot in addition suffer from undermined personhood. Indeed, much of the crucial evidence concerning addictive neuroadaptation is derived from experiments with rats (Beninger 2018). Natural selection tolerated this vulnerability because the kind of environmental contingency that triggers addiction does not arise naturally. When wild elephants or baboons

[5] For analysis of this concept, see Millikan (1984).

encounter low-toxicity sources of alcohol, for example berries that occasionally ferment on the vine, they indulge in benders that they evidently find sufficiently rewarding to tolerate (at least in the case of baboons) significantly enhanced predation risk while drunk. They are at no risk of addiction, however, because they cannot cultivate sources of low-toxicity alcohol. Their parties are windfalls, the frequency of which they cannot influence.

Modern humans encounter addiction in large numbers because they can engineer addictive environments. Where most addictive targets – particularly alcohol – are concerned, they have not engineered these environments *for the sake of* their addictive properties. Stockpiled alcohol has provided a range of enhancements to human welfare for millennia (Hockings and Dunbar 2020), which are traded off against non-zero frequency of alcohol addiction. Contemporary corporate producers of alcoholic beverages, while open to serious ethical criticism for encouraging natural risk-loving behavior in young people, do not attempt to engineer higher rates of addiction per se.

Producers and vendors of various other addictive targets, however, enjoy profits that are directly driven by their deliberate design and continuous, research-driven 'improvement' of addictive environments. This has clearly been the core business model of the cigarette industry since its inception (Brandt 2009). And the intentional fostering of increased rates of gambling induction is the currently dominant dynamic in the commercial gambling industry. As documented by Schüll (2012), each successive short generation of electronic gaming machine (EGM) technology has been engineered to make the player's experience more addictive. Particularly effective in this regard are multiple paylines, each carrying only a small stake on a given play, that fully engage the gambler's cognitive resources, thereby drawing her quickly and deeply into 'the zone', and that produce continuous strings of small wins to maximally arouse dopamine circuit response. Of course, the sum of small wins is smaller in expectation than the sum of small losses. The massive displacement of traditional gambling forms by EGMs on casino floors in Las Vegas, Macau, and elsewhere, is strong indicative evidence of the profit increases being derived from engineered addiction.

As discussed in Section 2, here, at the epidemiologist's population scale, is where we might unambiguously locate disease and disorder – *social* disease and disorder. I argued that if someone prefers to regard application of these concepts at the public health scale as merely metaphorical, that is largely a matter of rhetoric rather than substance. Application of the epidemiological disease concept to addiction is compatible with regarding addiction as a phenomenon of inefficient, under-informed and especially *manipulated and exploited* choice at the personal scale. Empathy with addicts can indeed be based partly on viewing them as victims – not of their brains, but of other, predatory, humans.

References

Abbott, M. (2017). Beyond Reno: A critical commentary on Hancock and Smith. *International Journal of Mental Health and Addiction* 15: 1177-1186.

Ainslie, G. (1992). *Picoeconomics*. Cambridge University Press.

Ainslie, G. (2001). *Breakdown of Will*. Cambridge University Press.

Alexander, B. (2010). *The Globalization of Addiction.* Oxford University Press.

American Psychiatric Association (2013). *Diagnostic and Statistical Manual of Mental Disorders 5.* APA Press.

Bénabou, R., & Triole, J. (2004). Willpower and personal rules. *Journal of Political Economy* 112: 848-886.

Beninger, R. (2018). *Life's Rewards*. Oxford University Press.

Bickel, W., Madden, G., & Petry, N. (1998). The price of change: The behavioral economics of drug dependence. *Behavior Therapy* 29*:* 545-565.

Borsboom, D., Cramer, A., & Kalis, A. (2018). Brain disorders? Not really… Why network structures block reductionism in psychopathology research. *Behavioral and Brain Sciences* 24: 1-54.

Brandt, A. (2009). *The Cigarette Century.* Basic Books.

Brockmann D., & Helbing, D. (2013). The hidden geometry of complex, network-driven contagion phenomena. *Science* 342: 1337–1342.

Carroll, M. (1996). Reducing drug abuse by enriching the environment with alternative non-drug reinforcers. In L. Green & J. Kagel (eds), *Advances in Behavioral Economics: Vol. 3. Substance Use and Abuse*, Norwood, NJ: Ablex; pp. 37-68.

Chaloupka, F. (1991), Rational addictive behavior and cigarette smoking, *Journal of Political Economy* 99: 722-742.

Chaloupka, F., Grossman, M., & Tauras, J. (1999). The demand for cocaine and marijuana by youth. In F. Chaloupka, W. Bickel, M. Grossman, and H. Saffer (eds), *The Economic Analysis of Substance Use and Abuse: An Integration of Econometric and Behavioral Economic Perspectives*, Chicago: University of Chicago Press; pp. 133-156.

DeGrandpre, R., Bickel, W., Hughes, J., & Higgins, S. (1992). Behavioral economics of drug self-administration: III. A reanalysis of the nicotine regulation hypothesis. *Psychopharmacology* 108: 1-10.

Elster, J., & Skog, O.-J., eds. (1999). *Getting Hooked: Rationality and Addiction*. Cambridge University Press.

Ferris, J., & Wynne, H. (2001). *The Canadian Problem Gambling Index Final Report* Canadian Center on Substance Abuse at www.ccsa.ca/pdf/ccsa-008805-2001.pdf.

Fingarette, H. (1989). *Heavy Drinking.* University of California Press.

Galea, S., Nandi, A., & Vlahov, D. (2004). The social epidemiology of substance use. *Epidemiology Review* 26:36–52.

Hancock, L. (2011). *Regulatory Failure? The Case of Crown Casino.* Australian Scholarly Publishing.

Hancock, L., & Smith, G. (2017). Critiquing the Reno model I-IV: International influence on regulators and governments (2004–2015)— the distorted reality of responsible gambling. *International Journal of Medical Health and Addiction* 15: 1151-1176.

Harrison, G., Lau, M., & Ross, D. (2019). The risk of gambling problems in the general population: A reconsideration. *Journal of Gambling Studies*, forthcoming. https://doi.org/10.1007/s10899-019-09897-2

Heyman, G. (2009). *Addiction: A Disorder of Choice.* Harvard University Press.

Higgins, S., Heil, S., & Lussier, J. (2004). Clinical implications of reinforcement as a determinant of substance use disorders. *Annual Reviews of Psychology* 55: 431-461.

Hockings, K., & Dunbar, R., eds. (2020). *Humans and Alcohol.* Oxford University Press.

Huang, H., Yan, Z., Chen, Y., & Liu, F. (2016). A social contagious model of the obesity epidemic. *Scientific reports* 6: 37961. doi:10.1038/srep37961

Ismael, J. (2016). *How Physics Makes Us Free.* Oxford University Press.

Keefe, J. (1996). *Divided Minds and Successive Selves.* MIT Press.

Keller, E. (2000). *The Century of the Gene.* Harvard University Press.

Kessler, R., Hwang, I., LaBrie, R., Petukhova, M., Sampson, N., Winters, K., & Shaffer, H. (2008). DSM-IV Pathological Gambling in the National Comorbidity Survey Replication, *Psychological Medicine* 38: 1351-1360.

Kincaid, H., Daniels, R., Dellis, A., Hofmeyr, A., Rousseau, J., Sharp, C., & Ross, D. (2013). A taxometric analysis of problem gambling data from a South African national urban sample. *Journal of Gambling Studies* 29: 377-392.

Lesieur, H., & Blume, S. (1987). The South Oaks Gambling Screen (The SOGS): A new instrument for the identification of problem gamblers. *American Journal of Psychiatry* 144: 1184-1188.

Leung, S.-F., & Phelps, C. (1993). "My kingdom for a drink . . . ?" A review of the estimates of the price sensitivity of demand for alcoholic beverages. In M. Hilton and

G. Bloss (eds), *Economics and the Prevention of Alcohol-Related Problems*, Rockville, MD: National Institute on Alcohol Abuse and Alcoholism Research Monograph No. 25. NIH Pub. No. 93-3513; pp. 1-31.

Lewis, M. (2015). *The Biology of Desire.* Public Affairs.

MacKillop, J., & Munafó, M., eds. (2013). *Genetic Influences on Addiction: An Intermediate Phenotype Approach.* MIT Press.

Millikan, R. (1984). *Language, Thought, and Other Biological Categories.* MIT Press.

Potenza, M. (2008). The neurobiology of pathological gambling and drug addiction: an overview and new findings. *Philosophical Transactions of The Royal Society B Biological Sciences* 363:3181-3189.

Ross, D. (2010). Economic models of addiction. In D. Ross, H. Kincaid, D. Spurrett and P. Collins, eds., *What is Addiction?* MIT Press, pp. 131-158.

Ross, D. (2011). Estranged parents and a schizophrenic child: choice in economics, psychology and neuroeconomics. *Journal of Economic Methodology* 18: 215-229.

Ross, D., Sharp, C., Vuchinich, R., & Spurrett, D. (2008). *Midbrain Mutiny: The Picoeconomics and Neuroeconomics of Disordered Gambling.* MIT Press.

Sharp, C., Dellis, A., Hofmeyr, A., Kincaid, H., & Ross, D. (2015). First evidence of comorbidity of problem gambling and other psychiatric problems in a representative urban sample of South Africa. *Journal of Gambling Studies* 31: 679-694.

Schüll, N.D. (2012). *Addiction by Design.* Princeton University Press.

Sulkunen, P., Babor, T., Örnberg, J.C., Egerer, M., Hellman, M., Livingstone, C., Marionneau, V., Nikkinen, J., Orford, J., Room, R., & Rossow, I. (2019). *Setting Limits: Gambling, Science, and Public Policy.* Oxford University Press.

Sutton, R., & Barto, A. (2018). *Reinforcement Learning: An Introduction.* 2nd edition. MIT Press.

Tiberius, V. (2018). *Well-Being as Value Fulfillment.* Oxford University Press.

Velleman, J.D. (2009). *How We Get Along.* Cambridge University Press.

Watters, E. (2010). *Crazy Like Us: The Globalization of the American Psyche*. Free Press.

White, W. (1998). *Slaying the Dragon: The History of Addiction Treatment and Recovery in America.* Bloomington: The Chestnut Health Systems / Lighthouse Institute.

Williams, R., Volberg, R., & Stevens, R. (2012). *The Population Prevalence of Problem Gambling: Methodological Influences, Standardized Rates, Jurisdictional Differences, and Worldwide Trends*. Report Prepared for the Ontario Problem Gambling Research Centre and the Ontario Ministry of Health and Long Term Care, May 2012; available at www.gamblingresearch.org/applydownload.php?docid=12579?.