DOI 10.36446/af.e1048

ON THE GROUNDS FOR CALLING ADDICTION A DISEASE*

Acerca de las razones para ver la adicción como una enfermedad

Federico Burdman [a]
https://orcid.org/0000-0002-8506-9271
fburdman@uahurtado.cl

[a] Departamento de Filosofía, Universidad Alberto Hurtado, Santiago, Chile.

Abstract

In this paper, I look into the debate about the status of addiction as a disease. Although addiction is widely regarded as a disease, several authors have put forward reasons for agnosticism or skepticism about the appropriateness of the disease label. Any attempt to address this issue directly is complicated by its relationship to several other contentious issues, both on the side of theories of addiction and on the side of theories of disease. My primary aim in this paper is to identify the major points of contention. My secondary aim is to offer a limited defense of the disease view. The crux of the debate is whether addiction is aptly pictured as the result of psychological dysfunction. The main hurdle for the psychological dysfunction claim is that we currently lack a relatively unified account of the behavioral trait that most strongly suggests dysfunctional processes: the loss of expectable ability to control behavior. I argue that there is a sense in which the dysfunction claim may be warranted even in the absence of a unified causal account. However, this requires assumptions that are to some extent controversial, most notably, the impaired control view of addiction and the acceptability of dysfunction-talk based on personal- level behavioral traits.

Key words: Addiction; Disease; Impaired Behavioral Control; Dysfunction; Harm.

Resumen

En este artículo, echo una mirada al debate acerca del estatus de la adicción como enfermedad. Aunque la adicción es comúnmente vista como una enfermedad, varios autores han esgrimido razones para el agnosticismo o el escepticismo acerca de la corrección de esta etiqueta. Cualquier intento de abordar esta discusión directamente se complica por su relación con varios otros debates abiertos, tanto del lado de las teorías de la adicción como del lado de las teorías de la enfermedad. Mi objetivo principal en este artículo es identificar cuáles son los principales puntos de controversia. Mi objetivo secundario es ofrecer una defensa limitada de la tesis de la adicción como enfermedad. El punto neurál-

* This article is the result of research funded by research projects PICT-2019-02605 (Agencia Nacional de Promoción Científica y Tecnológica, Argentina) and FONDECYT Iniciación 11250503 (Agencia Nacional de Investigación y Desarrollo, Chile).

gico del debate es si la adicción es el resultado de una disfunción psicológica, y el principal obstáculo para afirmar que hay tal disfunción psicológica es que actualmente carecemos de una explicación relativamente unificada del rasgo conductual que más fuertemente sugiere la ocurrencia de algo disfuncional: la merma de la capacidad esperable para controlar la propia conducta. Según sostengo, hay un sentido en el que la afirmación de disfunción puede estar justificada incluso en ausencia de una explicación causal unificada. Esto requiere, sin embargo, suposiciones que son hasta cierto punto controversiales, incluyendo la visión de la adicción en términos de deterioro del control conductual y la aceptabilidad de predicar disfunción a partir de rasgos conductuales a nivel personal.

Palabras clave: Adicción; Enfermedad; Control conductual deteriorado; Disfunción; Daño.

1. Introduction

Addiction is widely considered to be a disease[1]. It has been consistently listed in every edition of the DSM, including the latest one (American Psychiatric Association, 2022), it is listed in the World Health Organization's classification of diseases, the ICD-11 (World Health Organization, 2019), and it is commonly referred to as a disease by numerous other influential medical and research institutions around the globe. In psychiatric literature, support for the view is almost universal, and it is also widely accepted among treatment providers (Barnett et al., 2022). While popular conceptions of addiction have traditionally considered it to be a moral failing, the view of addiction as a disease has of late gained greater acceptance among the general public (Connelly et al., 2018; Pescosolido et al., 2010; Rise & Halkjelsvik, 2019). In addition, many people with addiction think of their condition as a disease (Snoek, 2017b, p. 189). Even non-medical institutions and support groups such as AA and NA refer to addiction as a disease (Flanagan, 2013).

Contrary to the prevailing opinion, some have remained agnostic or outright skeptical about the appropriateness of the disease label. Many have criticized the view of addiction as a disease *of the brain* (Field et al., 2019; Heather et al., 2022; Levy, 2013; Pickard, 2022a, 2022b), while others

[1] A brief note on terminology: I will follow the prevailing trend in the technical literature by focusing on cases of drug addiction, although what I say is intended to be relevant to other types of addiction as well. As for the term 'drugs,' I will use it liberally to refer to all substances that can be the object of addictive behavior, including alcohol, nicotine, and other substances not commonly referred to as 'drugs' outside the technical literature. People with addiction are often the target of stigmatizing attitudes that are both inhumane and detrimental to their chances of recovery. I intend my use of the term 'addiction' to carry no negative connotations about the people who suffer from it.

have questioned whether it is aptly seen as a disease at all (Heather, 2013, 2017; Heyman, 2009; Lewis, 2015, 2017; Ross, 2020; for an early defense of the view, see Fingarette, 1988). Skeptics have challenged the arguments advanced in defense for the disease view, pointing to disanalogies with paradigmatic medical conditions and claiming that addiction does not meet plausible criteria for inclusion. In addition, they have sometimes suggested that labeling addiction as a disease may not be in the best interest of those who suffer from it.

So, is addiction a disease? Any attempt to answer this question directly is complicated by its relationship to several other contentious issues, both on the side of theories of addiction and on the side of theories of disease. My primary aim in this paper is to lay out what the major points of contention are. My secondary aim is to offer a limited defense of the disease view. The crux of the debate is whether addiction is aptly pictured as the result of psychological dysfunction, and the main hurdle for the psychological dysfunction claim is that we currently lack a relatively unified account of the behavioral trait that most strongly suggests dysfunctional processes: the loss of expectable ability to control behavior. I argue that there is a sense in which the dysfunction claim may be warranted even in the absence of a unified causal explanation. This requires, however, assumptions that are to some extent controversial—most notably, the impaired control view of addiction and the acceptability of dysfunction-talk based on personal-level behavioral traits.

I begin by looking at the disease question itself. I suggest that it is not exhausted by either institutional or purely pragmatic considerations, and I introduce the hybrid account of disease as *harmful dysfunction* in terms of which the ensuing discussion is framed (Section 2). Subsequently, I turn to the question of whether addiction satisfies harm and dysfunction criteria. Addiction is obviously harmful, but one might ask whether it involves the *right sort* of harm to qualify as disease—namely, harm that is intrinsic to the condition. Although addiction-related harms are significantly exacerbated by modifiable socioenvironmental factors, I argue that it is highly unlikely that addiction is a condition whose only relationship to harm is mediated by extrinsic factors (Section 3). I then turn to what is arguably the central issue in the disease debate: whether addiction is aptly pictured as the result of dysfunction in the sense relevant for disease. I first discuss brain dysfunction. Critics of the disease view have objected on the grounds that we currently have no neural-level explanation of what is dysfunctional in the neural underpinnings of addiction. Although I accept this observation as true, I argue that on a plausible view of brain dysfunction, a claim of brain dysfunction need not presuppose that we have a way of identifying

dysfunction that is independent of psychological or behavioral criteria (Section 4.1.). The final battleground, then, is whether we have reason to believe that addiction involves psychological dysfunction. If the impaired control view of addiction is correct, I argue, then the loss of expectable ability to control behavior strongly suggests psychological dysfunction. One might worry that dysfunction-talk cannot be warranted unless we have a precise etiological explanation for impaired control. I concede that we do not currently have one, and it is an open possibility that a precise etiological explanation will not be forthcoming—the causes of impaired control in addiction may be too complex, varied and multifactorial to satisfy the requirement that we identify a precise etiological mechanism. I suggest that impaired behavioral control may nevertheless be sufficient for a claim of dysfunction if one drops the requirement of a relatively unified causal account (Section 4.2.).

2. What is a disease?

At risk of letting the reader down at the outset, I'll begin by stating the inevitable: whether addiction is a disease depends in large part on what one takes a disease to be. This is a question that has long been debated by philosophers of medicine. The main contenders in the debate are naturalism, normativism, and several variations thereof. *Naturalist* accounts take disease to be definable in purely medical or scientific terms. In a popular articulation of the view, a key criterion of disease is dysfunction, which naturalists believe can be glossed over in purely medical or scientific language (e.g., Boorse, 2014). *Normativists* see disease as a fundamentally value-laden concept. On most normativist accounts, calling a condition a disease implies that the condition is undesirable, or that the person suffering from it is unfortunate for having the condition, with judgments about these matters involving an inescapable reference to evaluative or otherwise normative standards (e.g., Cooper, 2020).

One point on which these venerable foes agree is that the question 'what is a disease?' should be approached as one whose answer depends on philosophical or otherwise conceptual considerations. This implies that the question is distinct from, and in some sense prior to, the question of whether a given condition is treated by competent clinicians and recognized as a disease by relevant medical institutions. It also implies that the question of whether a condition is a disease is not exhausted by pragmatic considerations, such as whether it is beneficial to regard the condition as a disease. I will assume as much in what follows. Before moving on, it is worth defending this assumption briefly.

Consider a coarse institutionalism about disease, the view that there is nothing to being a disease other than being accorded the status of disease by the world of medical institutions and/or competent medical practitioners. If being a disease is identified with its being treated as such, this view seems to commit us to the unpalatable implication that conditions *become* diseases when the relevant institutional arrangements are made and cease *to be* diseases when they are no longer regarded as such, rather than these decisions tracking prior facts relevant to disease status.

Institutionalism allows, however, for more interesting variations that leave enough room for the possibility of questioning actual institutional arrangements and practices. One stems from the observation that when clinicians and medical institutions debate whether a condition should be included in nosological classifications, they tend to focus on whether it would be *beneficial* to adopt the relevant practices and institutional arrangements. One might think then that all there is to the question of whether a condition is a disease is whether it would be beneficial for medical institutions—or for society at large— to regard it as such.

Pragmatic considerations of this sort play a major role in medical thinking. Consider, for instance, the distinction between clinical and subclinical problems in regulating drug use. The distinction between a disordered and a normal-range difficulty in regulating drug use is largely driven by considerations of whether the person would benefit from consultation or treatment by a competent physician. Even though treatability is an important consideration, some reflection suggests that it cannot be all there is to being a disease. For some diseases, we do not currently know of any beneficial medical intervention. And people can sometimes benefit from consultation with health professionals even though they do not suffer from a recognizable pathology. Many people seek therapy to cope better with a range of 'problems in living.' They benefit from therapeutic intervention without having, or being seen as having, a psychiatric disorder (Bortolotti, 2020).

Similar considerations apply to variations of the argument that focus on other potential benefits of seeing a condition as a disease. For instance, claims about the impact of the disease label on social attitudes toward addiction have often been presented as central to the disease debate. For the disease view, it has been argued that framing addiction as a disease leads to less stigmatizing attitudes among the public, as people come to see addiction as the purview of doctors rather than law enforcement (Leshner, 1997; McLellan et al., 2000; Volkow & Koob, 2015)[2].

[2] On a different key, it has also been argued that calling addiction a disease is central

Those who argue against the disease view have countered these claims. The alleged responsibility-undermining implications of the disease view have been cited as a reason against it (Holden, 2012). And whether adopting the disease label reduces stigma has also been disputed, with empirical studies yielding mixed results, at best (Corrigan & Watson, 2004; Kvaale et al., 2013; Meurk et al., 2014; Pescosolido et al., 2010). What is the impact of the disease label on social attitudes remains an open empirical question. Suppose, however, that the real effect of adopting the disease label is to increase the perception of 'otherness' of people with addiction in the eyes of the public. Conceptually, it would make perfect sense to continue to maintain that addiction is a disease, acknowledging at the same time some undesirable effects of the disease label. We would then be very concerned about combating the stigma of addiction (as we should be), but this concern does not seem to be a proper basis for arguing that addiction is not a disease.

It seems then we cannot ask whether addiction is a disease without taking a stance on what it is for a condition to be a disease in the first place. One might hope to do so, however, without taking a *substantive* stance. Consider the view of disease as a family resemblance concept (Lilienfeld & Marino, 1995). In this view, asking whether addiction is a disease amounts to asking about the extent to which it shares the features that conditions we think of as diseases typically have. Addiction arguably presents both analogies and disanalogies with paradigmatic medical conditions. For instance, the fact that pharmacological treatments are in use and that hereditary factors account for a portion of susceptibility to addiction seem to bring us closer to core components of the disease concept (Leshner, 1997; McLellan et al., 2000; Volkow & Koob, 2015). On the other hand, the fact that recovery from addiction, when successful, is a highly agential achievement seems to distinguish addiction from many paradigmatic diseases (J. Davies, 2018; Holden, 2012; Satel & Lilienfeld, 2017). Moreover, many people seem to benefit from 12-step programs, which approach addiction in a non-medical way—even though they often refer to it as a disease. In sum, it seems that addiction is in some ways similar to paradigmatic medical conditions, but

to securing funding for addiction research and promoting the search for more effective treatments (Berridge, 2017; Heilig et al., 2021), as well as a means of getting health insurance companies to cover addiction-related expenses (McLellan et al., 2000; Volkow & Koob, 2015). Not surprisingly, these claims have also been disputed. Its effectiveness in securing funding for addiction research has been criticized on the grounds that it leads to a disproportionate focus on brain-centered projects at the expense of research focused on social and environmental variables, and as effective, at best, at the cost of leaving matters of social justice out of sight (Hart, 2017; Lie et al., 2022).

it is also in some ways similar to life problems that we tend to approach in non-medical ways. If we could do no better than place addiction in this similarity space, we would probably find its place somewhere on the fringes of the domain of disease. How well addiction fits into the disease category would then be a matter of degree.

The above reasoning highlights an important point: if addiction is a disease, it is at best a somewhat unusual member of the category[3]. I resist, however, the family resemblance approach for much the same reasons given above: there seems to be content to the concept of disease that is not exhausted by placing particular conditions in varying degrees of approximation to paradigmatic medical conditions in resemblance space.

So, we are back at square one, facing a choice between substantive philosophical accounts of disease. In what follows, I will adopt a *hybrid* theory of disease (Wakefield, 1992). The hybrid account renders the core of the disease concept as the presence of a *harmful dysfunction*, thus integrating both naturalistic and normativist elements. My choice of a hybrid account is based on considerations of intrinsic plausibility, but it also has a strategic payoff. The two main criteria for disease on the hybrid account —harm and dysfunction— coincide with the main candidate criteria favored by other accounts (which nonetheless find reason to object whether one or the other is necessary). Thus, discussing whether addiction meets harm and dysfunction criteria will make the ensuing discussion still relevant to people committed to most other accounts of disease.

One more terminological note: I will not be making a principled distinction between disease and disorder. Although there is not universal agreement on the use of these terms, it is sometimes assumed that calling a condition a disease implies that there is a known explanation for its defining symptoms, whereas calling it a disorder implies no such commitment. I do not follow this usage here. One reason is that, I think, we are often content to call certain conditions diseases even when we lack a consensus explanation of their defining symptoms. Moreover, whether it is appropriate to call a condition a disease in the absence of a precise causal explanation of its symptoms is one of the key points under debate. It is therefore preferable not to settle the issue by terminological fiat.[4]

[3] I thank an anonymous reviewer for pressing me to clarify this point.

[4] Importantly, Wakefield's influential hybrid theory is officially a theory of disorder, even though it proposes a dysfunction criterion and it is meant to apply broadly to the medical domain. I thank an anonymous reviewer for pressing me to clarify this point.

3. Harm

In the standard formulation of the hybrid account, for a condition to be a disease it must be harmful to the individual who has it in the light of appropriate sociocultural standards (Wakefield, 1992, 2021). The harm criterion stands for the normative component in the theory of disease: part of what we say when we call a condition a disease is that it is bad for you. Consider *situs inversus totalis*, a rare condition in which the lateralized position of organs within the thorax is inverted (Eitler et al., 2022). *Situs inversus* is plausibly seen as a case of a physiological (developmental) process gone wrong. However, people who have it typically live their entire lives without experiencing any harm as a result of the condition, and therefore it is arguably inaccurate to say that they suffer from a disease.

Whether addiction meets the harm criterion might seem like a point of no debate. No one familiar with severe cases could argue that having an addiction is not harmful. Some see harm as a definitional feature of the condition, arguing that someone cannot properly be said to have an addiction if the way they use drugs does not cause them significant harm (e.g., Sinnott-Armstrong & Pickard, 2013). Indeed, harmful consequences are part of the definition of the Substance Use Disorder label in the DSM-5 and are reflected in the diagnostic criteria. Even if it were better thought of as not defining of addiction, it would be hard to deny that harm is a central element in typical addiction case-histories. In severe cases, people with an addiction often enter a destructive spiral that leads to death or other dramatic consequences, including loss of health, loss of housing and livelihood, damage to important relationships, among others. Even theorists who reject the view of addiction as a disease on other grounds agree that it is "unquestionably destructive" (Lewis, 2015, p. x).

A more difficult question arises when we ask whether addiction involves the *right sort* of harm to be a disease. Intuitively, it must result from the condition in a way that does not depend on contingent features of the social, environmental, or cultural setting. Consider the neurodiversity view of autism. Proponents typically acknowledge that people with autism experience difficulties in navigating a social world constructed to meet the needs of neurotypical individuals. But—the argument goes—it is not the condition that is intrinsically harmful, but rather the social attitudes and environmental arrangements that fail to adequately accommodate the needs of neurodiverse individuals (Dominus, 2019). These contrasting ways of thinking about the relevant harms have importantly different implications. The former suggests that it is the person who should be treated to better adapt to environmental conditions, whereas the latter

suggests that it is the environment, not the person, that should be modified.

A further complication arises when we consider what goes into an environment being modifiable. Upon closer examination, the question points to more than the mere feasibility of envisioned changes in environmental conditions. Any such changes are bound to involve costs of various kinds—economic, social, moral. Thus, when we ask, 'Is the harm experienced by people with this condition an intrinsic feature of the condition?' the question we pose is partly normative in nature. Consider homosexuality, another *locus classicus* in this debate. In the context of a homophobic society, being gay can certainly lead to harm. But it is the homophobic attitudes that need to be amended, not gay people. In saying this, one is not simply assuming that changes in homophobic attitudes are possible. Indeed, it is clear that the struggle against homophobic mindsets and attitudes remains a challenging task. To say that it is the homophobic environment rather than the person that needs to be changed is to take a *moral* stance. One way to articulate this stance is to claim that the cost of getting people to leave homophobic attitudes behind is morally justified.

Neil Levy (2013) has suggested that something resembling the above argument may apply to addiction. What gets the argument off the ground is that there is considerable evidence that environmental circumstances have a significant impact on addiction. For instance, the proximity of opportunities to use are known to modulate the intensity of craving (Dar et al., 2010; Juliano & Brandon, 1998) and craving episodes are highly context-sensitive in further ways, as they are easily triggered by a variety of environmental cues that the person associates with drug use (Skinner & Aubin, 2010). Moreover, social and economic variables also play a role. Substance use problems disproportionately affect those suffering from housing and employment instability (Saloner & Cook, 2013), and recovery is much more likely for those who can benefit from social support networks and who have credible alternatives to a drug-focused lifestyle (Hart, 2013). In addition, many believe that stigmatizing attitudes toward people with addiction play a role in explaining why it is so difficult to quit (Hadland et al., 2018; Yang et al., 2017).

Addiction problems would look very different in a more just society, one in which opportunities were available to all and in which people with addiction were not rejected and stigmatized. Levy's argument does an important service by undermining the comfortable assumption that society as a whole has little to do with harm in addiction—there is indeed much that could be done to alleviate these harms. Nevertheless, it is a stretch to conclude that addiction-related harms would disappear under ideal social

conditions. It is difficult to imagine a scenario in which addiction might turn out to be like homosexuality as a trait whose *only* relationship to harm is mediated by modifiable factors.

First, some of the harms associated with addiction do not seem to depend on conditions external to the agent in any robust sense. Many people experience various types of health problems as a result of a history of drug abuse. There is evidence of neurotoxic effects associated with sustained long-term use of some drugs, particularly alcohol and stimulants (Ersche et al., 2013; Xiao et al., 2015), and long-term drug use can have adverse effects on the organism beyond the brain. The adverse health effects associated with long-term alcohol and nicotine use are well known.

Other types of harm result from social and environmental factors that do not appear to be truly modifiable. It is probably true that drug availability can make a difference in people's ability to control their use, but it is hard to imagine a scenario in which drugs could be made truly unavailable (as opposed to outlawed). It is also hard to imagine changing the environment so that it is completely free of cues that trigger craving, since craving can be triggered by all sorts of things that the person has come to associate with drug use, including the time of day (Palij et al., 1996). For other types of cues, the magnitude of the envisioned environmental change may be so great as to be almost unfeasible.

To recapitulate, on a plausible construal of the harm criterion for disease, whether addiction is a disease depends in part on the extent to which we think of addiction-related harms as the result of modifiable social and environmental conditions. The more we are inclined to think that there are accessible possible worlds—as Levy would put it—in which relevant changes in environmental conditions would lead to the elimination of these harms, the less addiction will seem to involve the kind of harm relevant to disease. But it is difficult to imagine a scenario in which suitable changes to environmental conditions would lead to the elimination, rather than reduction, of addiction harms. Some of these do not seem to depend on environmental conditions in a robust sense, and some of the environmental conditions that do play a role appear not to be truly modifiable. Some of the harms experienced by people with addiction are best thought of as intrinsic to the condition, even though stigma and social distancing undoubtedly add unnecessary pain to the harms experienced by people with addiction.

This is not the same as saying that addiction is a disease, since not all inherently harmful conditions are aptly called diseases. It is time now to move on to the other main component of the hybrid account: the dysfunction criterion.

4. Dysfunction

The question of whether addictive behavior is the result of dysfunctional processes naturally leads to the question of how we determine what the relevant norms of proper functioning are. This is, alas, another notoriously contentious issue. Naturalistic views of function and dysfunction have tended to cluster around two main families of views. *Selectionist* or *etiological views* identify the function of a trait or mechanism with the production of the effect for which it was selected over evolutionary history (Millikan, 1984; Neander, 1991; Wakefield, 1992). *Systemic* or *causal views* see function as depending not on historical facts but on the actual causal role of the trait or mechanism in the context of wider organismic operation (Boorse, 1975; Cummins, 1975). As noted above, Wakefield himself favors a selectionist account, and thus the standard formulation of the hybrid account of disease incorporates a selectionist understanding of dysfunction. In principle, however, one could side with Wakefield in thinking that disease requires both harm and dysfunction, and still endorse a different account of the latter. Another way in which I will depart from the standard hybrid approach is that I will not assume that an acceptable account of function must necessarily be value-free.[5]

(My own view is that the causal account comes closer to capturing the sort of contrast between well-functioning and dysfunctional traits that is at play in judgments about health and disease. A case discussed by Wakefield (2020) serves as an illustration. Imagine a gosling that, upon emerging from its shell, first encounters a fox and subsequently develops the stereotypical mother-oriented behaviors of geese toward the fox. In a sense, the imprinting mechanism is failing to do what it was evolutionarily meant to do, e.g., it has led the gosling to pair with the wrong creature in a way that will negatively affect its fitness. On the other hand, there is a sense in which the mechanism is performing as expected: it has successfully registered the visual impression of the first creature it encountered and has subsequently been successful in tracking and following around that creature. My intuitive reading of the case is that, while there is a clear sense in which something has gone awry with the gosling's development, the sense in which her imprinting mechanism has failed her is not of the sort we think of as disease or disorder[6]. Regretfully, I cannot properly defend this claim here.)

[5] For the view that function itself is a value-laden notion, see Amundson (2000). Garson (2024) argues that the systemic view is not value-free (which he sees as a reason to favor a selectionist account).

[6] This contrasts with Wakefield's proposed reading of the case as motivating a notion of mental disorder without an underlying neurological dysfunction.

The discussion of brain dysfunction and psychological dysfunction raises different issues, so I will discuss them in turn. To foreshadow, my view is that whether brain dysfunction is involved in addiction depends, among other things, on whether we have reason to believe there is dysfunction at the psychological level. With respect to psychological dysfunction, the main difficulty is not how we define norms of proper functioning, but the prior question of what mechanism is responsible for the dysfunctional outcomes. I argue that a plausible case for psychological dysfunction can be made given two critical assumptions: that addiction involves a loss of expectable ability to control behavior, and that dysfunction-talk is acceptable in the absence of a relatively unified etiological explanation.

4.1. Brain dysfunction

Proponents of the *Brain disease model of addiction* (BDMA) have emphasized that addiction correlates with a series of structural and functional changes in relevant brain systems, both at the synaptic and circuitry levels (Berridge, 2017; Everitt & Robbins, 2016; Heilig et al., 2021; Hyman, 2007; Kalivas & Volkow, 2005; Leshner, 1997; Volkow et al., 2016; Volkow & Koob, 2015). Most accounts focus on changes in the brain's reward system—a collection of neural circuits in the midbrain, including parts of the ventral tegmental area, the nucleus accumbens, the basal ganglia, the thalamus, and their projections to the medial prefrontal cortex. The reward system is widely believed to play a critical role in motivated behavior, particularly as mediated by various forms of reinforcement learning.

There is indeed a large body of research showing that, on average, the brains of people with addiction differ from those of non-addicted controls in several ways. Although the evidence is generally considered to be quite robust, a couple of caveats should be noted.

First, the neural basis of addiction is not easily distinguished from the neurotoxic effects of long-term drug ingestion. As noted above, there is evidence that sustained use of large amounts of alcohol (Xiao et al., 2015) and stimulants (Ersche et al., 2013) can cause grey matter abnormalities, particularly in the prefrontal cortex. It is plausible that this is a *secondary effect* of a history of substance abuse. The neurotoxic effects are thought to occur only after long-term drug exposure, making addiction conceptually distinct from, and perhaps causally prior to, such effects. Assuming this is correct, these brain changes should not be confused with the neural signature of addiction proper. Distinguishing one from the other is a challenging task.

A second difficulty is that, once differences in certain parameters of brain structure and function have been identified between people with addiction and non-addicted controls, it is not easy to determine whether we are looking at the neural basis of addiction or at the neural correlates of some other trait—say, higher than average impulsivity—that increases its possessor's susceptibility to addiction. Empirical studies of the brain basis of addiction recruit subjects who are known to have or have had the condition, but we usually lack information about what was going on in people's brains *before* they developed addiction. Moreover, we know that people can differ in their susceptibility to addiction due to a number of factors, including genetic ones (Verhulst et al., 2015). Thus, it cannot be easily ruled out that some of the structural and functional differences cited in the empirical literature were already present as precursors to addiction and thus, again, not to be confused with the neural signature of addiction itself[7].

Despite these challenges, neuroscientific studies have made remarkable progress in identifying brain differences that are credibly pictured as correlated with addictive behavior. At the macro level, addiction correlates with distinct patterns of metabolic activity in dorsal striatal and frontal circuits. At the micro level, there is evidence for changes in baseline levels of key neurochemicals and in the availability of neuroreceptors at synaptic junctions.[8]

The key question is whether evidence of brain differences is aptly interpreted as evidence of brain dysfunction. Critics of the BDMA have countered that we have no neural-level reason to believe that the said brain changes are dysfunctional. Or, to put it another way, that the only reason we have to believe that there is anything dysfunctional about the relevant brain processes is that they correlate with addictive behavior. In this respect, things look different from, say, brain tissue damage or axon demyelination, which cause deviations from standard brain structure and function that can be described as pathological before weighing their neuropsychological effects. Importantly, proponents of the BDMA agree—a "specific pathognomonic brain lesion" has not been identified as the neural basis of addiction (Heilig et al., 2021, p. 4).

Deviation from the statistical norm is not sufficient for dysfunction. The reason is simple: brains are highly plastic organs, designed to change in response to experience. It is to be expected that the brains of highly specialized agents will differ in certain respects from the baseline of

[7] The point is made by Hall et al. (2015, p. 107), among others.

[8] See Heilig et al. (2021) for a summary of the evidence.

control subjects[9]. Neuroscientist Marc Lewis has argued that the brain changes we see in addiction are the neural underpinnings of a form of "entrenched learning" (2015, 2017). More cautiously, Hanna Pickard has argued (2022a, p. 328) that at present we simply do not know enough about normal brain function to demarcate pathological from non-pathological deviations from standard parameters in the absence of gross anatomical abnormalities.

The criticisms raised by Lewis and Pickard are correct in an important sense: we do not currently have a brain-level account of what is pathological about brain differences in addiction. The judgment that these brain changes amount to brain dysfunction depends substantially on the observation that they are correlated with addictive behavior. Thus, in an important sense, claims about brain dysfunction in addiction are based on behavioral and psychological criteria.[10]

One may wonder, however, whether being identifiable by purely neural-level means is a prerequisite for something being a brain dysfunction. Anneli Jefferson (2022) has convincingly argued that it is not. If there is an identifiable type of brain process that realizes a psychological trait or process that we have reason to regard as dysfunctional, then that amounts to a good reason to regard that type of brain process as dysfunctional. It is true that there are cases of brain diseases where it is possible to explain what has gone wrong in purely neural terms. Jefferson cites brain tumors and neurosyphilis as paradigmatic examples: there is a clearly identifiable deviation from normal brain structure and function which is plausibly depicted as pathological before weighing its psychological effects, and which precedes and is causally responsible for such effects. But to think that conformity to this picture is required for any claim about brain dysfunction is to set the bar unduly high. Neural processes can go wrong in a myriad of different ways that are compatible with the absence of macro-level signs of dysfunction. And it is far from obvious why the means through which we are (at present) able to identify neural processes that have gone wrong should be definitional of what it *is* for neural processes to fail to function properly.

[9] This point was famously illustrated a few years ago by a study of taxi drivers. Looking at scans of their brains, the study reported that the posterior hippocampi of taxi drivers were significantly larger than those of controls (Maguire et al., 2000). This is not evidence of a pathogenic process by any plausible standard.

[10] Importantly, a similar case could be made for most of the conditions included in standard psychiatric nosology. At present, virtually no psychiatric condition can be traced back to specific neural signatures that we have independent reason to believe are pathological in nature. That is, one is tempted to add, why we think of them as psychiatric rather than neurological disorders.

I agree with Jefferson on this point: if it can be shown that there are brain differences —brain processes that can be type-individuated—that realize psychological traits or processes that we have reason to think are dysfunctional, then it seems warranted to call the relevant brain differences brain dysfunctions. Importantly, this does not answer the question of dysfunction in addiction, but rather moves it to the psychological level.

4.2. Psychological dysfunction

Beyond the BDMA, theories of addiction have focused on several different psychological processes that appear to work anomalously in this context. A short list includes anomalies in desire and motivation (Burdman, 2024; Holton & Berridge, 2013; Kavanagh et al., 2005), decision-making processes (Redish et al., 2008), temporal discounting (Ainslie, 2000), attentional processes (Cox et al., 2016), habit formation (Schroeder & Arpaly, 2013; Tiffany, 1990), and thought and belief formation processes (Levy, 2014; Pickard, 2016; G. Segal, 2013; Sripada, 2022). A comprehensive discussion of psychological dysfunction in addiction would look into all candidate processes, but this is beyond the scope of a single paper. For current purposes, I will focus on what is arguably the main suspect for psychological dysfunction in addiction: the impairment of the ability to control drug-related behavior. This is, I hasten to add, an important limitation of my argument. There is a particularly important tradition of research on the idea that addiction reflects some sort of malfunction of the mechanisms of desire that certainly deserves more attention than I can give it at present. My focus on impaired control is motivated by the thought that the 'compulsive' nature of addictive behavior is probably the main reason why psychiatrists tend to think of addiction as a disorder[11].

Impaired ability to control drug-related behavior is considered a key feature of addiction by highly influential institutional sources, including the DSM-5 and the ICD-11. In the DSM, four of the eleven diagnostic criteria for Substance Use Disorder refer to impaired behavioral control (American Psychiatric Association, 2022, p. 545). First-person narratives of addiction also typically speak of impaired control over drug use (Hammer et al., 2012; Hänninen & Koski-Jännes, 1999; Snoek, 2017a, among many others). In the technical literature the idea is so widespread that it would be unwarranted to attribute it to any one source. A qualitative study of

[11] A similar point is made by Gene Heyman and Verna Mims (2017, p. 386). Gabriel Segal has suggested that this may also be true of the BDMA (2013, note 5). Although the BDMA is couched in the language of neural dysfunction, it is actually aimed at identifying the neural basis of compulsive behavior.

expert opinion found that virtually all respondents endorsed the view that addiction impairs control over drug use (Carter et al., 2014).

There are, however, respectable dissenting voices. *Choice theories* of addiction propose to explain apparent evidence of impaired control by reference to relevant aspects of the circumstances under which decisions to use are made—circumstances that make use seem, from the agent's perspective, the most valuable among realistically available options (Ainslie, 2000; Becker & Murphy, 1988; Field et al., 2019; Foddy & Savulescu, 2010; Heyman, 2009; Pickard, 2018, 2022a). On such a view, it is not necessary to posit an impairment of behavioral control to explain addictive behavior. Elsewhere I have argued that positing *some degree of* impairment in behavioral control is necessary to explain addictive behavior (Burdman, 2022), but I cannot do justice to the relevant dialectic here. For present purposes, I will assume that the mainstream view is correct and that impaired control is a key feature of (severe) addiction.

Impaired control is an elusive concept. Addictive behavior is not a reflex-like occurrence, and the cravings experienced by people with addiction are not literally irresistible. In a suitable sense of these terms, addictive behavior is plausibly described as intentional, goal-directed action, and is explainable in part by reference to what the person most wants to do and the choices she makes. In addition, addictive behavior is sensitive to incentives and is typically flexible in the face of environmental events and situational pressures. In these respects, addictive drug use differs markedly from purely involuntary behavior.[12] At the same time, there is some hard to come by sense in which the flexible, intentional, and choice-based nature of addictive behavior appears to coexist with a significant impairment of control.[13] The most striking illustration of this is just how difficult it is for people with addiction to refrain from using, even when they are fully committed to the judgment that using is bad for them and that quitting is the only way to improve their situation.

On the face of it, loss of expectable ability is at least *suggestive* that something is wrong in the mechanisms that lead to behavior. Consider the kinds of behaviors that are considered of interest from a clinical perspective.

[12] Arguably, the common description in the psychiatric literature of such behavior as compulsive is not intended to deny these observations. Berridge (2017) and Heilig and colleagues (2021) explicitly acknowledge that compulsion-talk need not negate that there is a sense in which the agent can be described as making choices.

[13] Many have attempted to make ends meet, proposing various theoretical moves to reconcile the intentional, choice-based aspect of addictive behavior with a relevant sense of impaired control (see Burdman, 2022; Henden, 2018; Holton & Berridge, 2013; Kennett, 2013; Sinnott-Armstrong, 2013; Sripada, 2022).

Picture someone who consistently has difficulty in curbing her drug use, and often ends up using more than intended. Or someone who has great difficulty refraining from using, even when doing so clearly interferes with other things the person very much wants to do, is expected to do, and are part of a social role with which the person identifies. Imagine further that the motivation she experiences to use is so great that she engages in hazardous and risky behaviors in pursuit of drugs, even though she is aware that she is putting herself at risk. Now suppose that the same person frequently does all of the above, as will often be the case when addiction is severe. Such things seem to be credible grounds for claiming that her behavior manifests a failure to function adequately in a relevant sense of the term: it amounts to a significant disruption of an important ability we normally expect people to have—the ability to control their own behavior.

Is that enough? That depends on one's view of what is required for a dysfunction claim. One might object that function and dysfunction are predicated of particular subpersonal mechanisms, so dysfunction-talk would be unwarranted unless we can clearly identify a subpersonal mechanism that is not functioning properly. On this view, a behavioral-level trait such as impaired control may indeed suggest that something has gone wrong in the subpersonal machinery, but it would be insufficient to justify a dysfunction claim. This approach is consistent with popular philosophical accounts of function, and it coheres with a popular view of disease—what Dominic Murphy calls the *strong* interpretation of the medical model (Murphy, 2009). In brief, the strong interpretation takes diseases to have (biologically) definable essences: for all true diseases, it should be possible to trace observable symptoms back to underlying pathogenic processes. We may not know at present the exact etiological explanation for some diseases, but all true diseases must have one. Again, neurosyphilis can be taken as a paradigmatic example. A plurality of symptoms, both physiological and psychological, can ultimately be traced back to a single underlying pathogenic process—a bacterial infection[14].

When we turn to theories of addiction, it is unclear what might be cast in such a role. Theories of addiction abound, variously placing their bets on cognitive, motivational, valuational, affective, situational and social explanations. It is possible that in due course we will discover a single mechanism capable of explaining all the superficial variability of addiction, with a result that approximates the strong model of disease. My present argument is certainly not sufficient to rule out this possibility, since I have not considered the prospects of important theories that focus on some of the

[14] The example is discussed in Jefferson (2022), chapter 2.

candidates for dysfunction. But the interesting question is what we would make of this if a precise etiological mechanism were not forthcoming. It is certainly an open possibility that the condition we currently call addiction is simply too heterogeneous, and the explanation for the impairment in behavioral control that we take to be constitutive of it may be too complex, too varied, and too multifactorial. There would be then two main ways forward. One would be to insist, along the lines of a strong model, that conditions are individuated by etiological criteria. Thus, in due course, we would come to see that there is no such thing as addiction as a unified kind (Pober, 2013). We would end up with a host of different succeeding labels, each defined by a specific etiological mechanism, only loosely unified by the fact that the loss of expectable ability to control drug-related behavior is a common symptom of them.

The main other way forward would be to drop the requirement of a unified etiological explanation. Importantly, this is consistent with another important understanding of disease, that Murphy calls the *minimal interpretation* of the medical model (2009). The minimal interpretation privileges symptoms over underlying causal pathways. On such a view, it may be sufficient for disease that a cluster of symptoms tends to occur together, unfold in characteristic ways, and be beneficially addressed by a similar kind of medical response, whether or not these are explained by common causal mechanisms. This would be consistent with retaining the idea of addiction as we currently think of it, on the assumption that the correct approach to explaining impaired control will turn out to be *pluralistic*, rather than reducible to a single key factor[15].

What would we make of the dysfunction claim then? I suggest that we could retain the basic intuition that loss of expectable ability is sufficient evidence that something has gone wrong in the mechanisms leading to behavior, even if this does not translate into a unified or even a restricted set of causal explanations. The more contentious move in the argument is to allow for dysfunction-talk to be grounded in personal-level psychological and behavioral traits. In its defense, this move is consistent with minimal model-inspired accounts that place high-level constructs at the center of the theory of mental disorder, such as a decrease in reasons-responsiveness (Graham, 2010) or the undermining of capacities that are critical to being a person (Edwards, 2009) or necessary for a flourishing life (Sisti et al., 2013). In such views, the crucial consideration is how the person is able

[15] There is a growing number of voices favoring such an approach to the mechanisms underlying psychiatric illness. For a recent review of the literature, see A. Segal et al. (2024).

to function in her environment, regardless of whether a precise etiological mechanism has been identified. This is a different concept of (dys)function than the one often at play in the disease debate, since it departs from a naturalistic understanding. From a descriptive point of view, however, it is not implausible that a concept of (dys)function that focuses on the way persons are able to operate in their environment is a crucial component of our intuitive understanding of mental disorder.

Importantly, some of the objections raised against the disease view of addiction fail to gain much traction when the disease view is articulated in terms of a pluralistic approach to impaired control. Consider the charge of reductionism (J. Davies, 2018; Field et al., 2019; Heim, 2014; Lie et al., 2022; Satel & Lilienfeld, 2014). This may be an apt criticism of the BDMA, but it need not be if one adopts a pluralistic approach to explaining impaired control. Such an approach is in principle consistent with a wide range of options in the metaphysics of mind, including granting psychological properties a fair degree of autonomy from neural-level properties.

Second, and relatedly, the worry that the disease view downplays the role of social, environmental, cultural, and historical context (Field et al., 2019; Satel & Lilienfeld, 2014, 2017) also proves unfounded on this way of approaching the disease view. Again, the worry is plausibly raised against standard formulations of the BDMA, which tend to adopt a strongly internalistic approach in which the explanation of addictive behavior is fundamentally a matter of dysregulation of the inner workings of the reward system. If psychological dysfunction is grounded directly in the loss of expectable ability to control behavior, then the claim is consistent with viewing the causal explanation of impaired control as extending beyond the internal to include social, environmental, or other external factors that play a crucial role in explaining addictive behavior (Glackin et al., 2021). More broadly, this way of framing the disease view is consistent with externalistic approaches to psychiatry (W. Davies, 2016; de Haan, 2020; Maiese, 2021; Roberts et al., 2019).

5. Conclusion

I have articulated and, in a limited sense, defended a version of the view of addiction as a disease. The main limitation of my argument is that it hangs on several assumptions concerning contentious issues in theories of disease, theories of biological or mental function, and theories of addiction. Nevertheless, I hope to have made some progress in identifying some of the major points of contention and the assumptions that go into articulating a plausible way of defending the disease view.

I began by assuming a hybrid theory of disease and then looked into the grounds for claiming that addiction meets its two criteria: harm and dysfunction. On a plausible interpretation of the harm criterion, this requires that the condition reliably leads to harm in a way that does not depend on modifiable features of social and environmental context. While the harmfulness of addiction is beyond doubt, one might wonder about its dependence on socioenvironmental factors. I argued that some of the harms associated with addiction do not seem to depend much on environmental factors, and that some of the environmental factors that do seem to play a critical role in addiction—such as drug availability and the presence of craving-triggering cues—do not seem to be modifiable in the sense at issue. This, I suggested, is sufficient to claim that addiction meets the harm criterion for disease.

I then turned to dysfunction. I sided with critics of the BDMA in their observation that we currently lack a neural-level explanation of what is dysfunctional about the types of brain processes that correlate with addictive behavior. On a plausible view of brain dysfunction, however, claims of brain dysfunction need not require that dysfunction be identifiable by neural-level means. This transfers the question to the psychological level. I focused on what is arguably the most important candidate for psychological dysfunction: the impairment of behavioral control. The loss of expectable ability to control behavior certainly suggests that something is wrong with the mechanisms that lead to behavior, but it may not be enough if one requires that dysfunction-talk is backed by a precise etiological explanation. Indeed, the main hurdle for the psychological dysfunction claim is that we do not currently have a sufficiently unified explanation for impaired control. One might continue to bet that a relatively unified explanation will be eventually be uncovered, and thus defend the dysfunction claim on the basis of that promise. I have suggested that there is a way in which we might endorse the dysfunction claim even if a relatively unified explanation of impaired control were not forthcoming. This requires accepting dysfunction-talk in the absence of a precise etiological story. For those willing to make such a move, the loss of expectable ability to control behavior that we see in addiction may be sufficient for psychological dysfunction. This way of articulating the dysfunction claim avoids some important criticisms raised by those who resist the disease view—reductionism and excessive internalism—, thus faring better in this regard as well.

References

Ainslie, G. (2000). A research-based theory of addictive motivation. *Law & Philosophy*, *19*, 77-115. http://www.jstor.org/stable/3505175

American Psychiatric Association (2022). *Diagnostic and statistical manual of mental disorders: DSM-5-TR*. American Psychiatric Association.

Amundson, R. (2000). Against normal function. *Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences*, *31*(1), 33-53. https://doi.org/10.1016/S1369-8486(99)00033-3

Barnett, A., Savic, M., Pickersgill, M., O'Brien, K., Lubman, D. I., & Carter, A. (2022). Addiction treatment providers' engagements with the brain disease model of addiction. In N. Heather, M. Field, A. C. Moss, & S. Satel (Eds.), *Evaluating the brain disease model of addiction* (pp. 283-294). Routledge. http://doi.org/10.4324/9781003032762-30

Becker, G., & Murphy, K. (1988). A theory of rational addiction. *Journal of Political Economy*, *96*(4), 675-700. https://doi.org/10.1086/261558

Berridge, K. C. (2017). Is addiction a brain disease? *Neuroethics*, *10*(1), 29-33. https://doi.org/10.1007/s12152-016-9286-3

Boorse, C. (1975). On the distinction between disease and illness. *Philosophy & Public Affairs, 5*(1), 49-68. http://www.jstor.org/stable/2265020

Boorse, C. (2014). A second rebuttal on health. *Journal of Medicine and Philosophy*, *39*(6), 683-724. https://doi.org/10.1093/jmp/jhu035

Bortolotti, L. (2020). Doctors without 'disorders.' *Aristotelian Society Supplementary Volume*, *94*(1), 163-184. https://doi.org/10.1093/arisup/akaa006

Burdman, F. (2022). A pluralistic account of degrees of control in addiction. *Philosophical Studies*, *179*(1), 197-221. https://doi.org/10.1007/s11098-021-01656-7

Burdman, F. (2024). Recalcitrant desires in addiction. In S. Amaya, D. Shoemaker, & M. Vargas (Eds.), *Oxford Studies in Agency and Responsibility, vol. 8* (pp. 58-80). Oxford University Press. https://doi.org/10.1093/oso/9780198910114.003.0004

Carter, A., Mathews, R., Bell, S., Lucke, J., & Hall, W. (2014). Control and responsibility in addicted individuals: What do addiction neuroscientists and clinicians think? *Neuroethics*, *7*(2), 205-214. https://doi.org/10.1007/s12152-013-9196-6

Connelly, M., Malato, D., Benz, J., Rein, D., Emery, S., Meit, M., Oppenheimer, C., Tompson, T., Kantor, L., Ganesh, N., & Swanson, E. (2018). *Americans recognize the growing problem of opioid addiction*. The Associated Press-NORC Center for Public Affairs Research.

Cooper, R. (2020). The concept of disorder revisited: Robustly value-laden despite change. *Aristotelian Society Supplementary Volume*, *94*(1), 141-161. https://doi.org/10.1093/arisup/akaa010

Corrigan, P. W., & Watson, A. C. (2004). At issue: Stop the stigma: Call mental illness a brain disease. *Schizophrenia Bulletin*, *30*(3), 477-479. https://doi.org/10.1093/oxfordjournals.schbul.a007095

Cox, W. M., Klinger, E., & Fadardi, J. S. (2016). Nonconscious motivational influences on cognitive processes in addictive behaviors. In N. Heather & G. Segal (Eds.), *Addiction and Choice* (pp. 259-285). Oxford University Press. https://doi.org/10.1093/acprof:oso/9780198727224.003.0015

Cummins, R. (1975). Functional analysis. *The Journal of Philosophy*, *72*(20), 741-765. https://doi.org/10.2307/2024640

Dar, R., Rosen-Korakin, N., Shapira, O., Gottlieb, Y., & Frenk, H. (2010). The craving to smoke in flight attendants: Relations with smoking deprivation, anticipation of smoking, and actual smoking. *Journal of Abnormal Psychology*, *119*(1), 248-253. https://doi.org/10.1037/a0017778

Davies, J. (2018). Addiction is not a brain disease. *Addiction Research & Theory*, *26*(1), 1-2. https://doi.org/10.1080/16066359.2017.1321741

Davies, W. (2016). Externalist psychiatry. *Analysis*, *76*(3), 290-296. https://doi.org/10.1093/analys/anw038

de Haan, S. (2020). An enactive approach to psychiatry. *Philosophy, Psychiatry, & Psychology*, *27*(1), 3-25. https://doi.org/10.1353/ppp.2020.0001

Dominus, S. (2019, February 19). Open Office. *The New York Times Magazine*.

Edwards, C. (2009). Ethical decisions in the classification of mental conditions as mental illness. *Philosophy, Psychiatry, & Psychology*, *16*(1), 73-90. https://doi.org/10.1353/ppp.0.0219

Eitler, K., Bibok, A., & Telkes, G. (2022). Situs inversus totalis: A clinical review. *International Journal of General Medicine*, *Volume 15*, 2437-2449. https://doi.org/10.2147/IJGM.S295444

Ersche, K. D., Williams, G. B., Robbins, T. W., & Bullmore, E. T. (2013). Meta-analysis of structural brain abnormalities associated with stimulant drug dependence and neuroimaging of addiction vulnerability and resilience. *Current Opinion in Neurobiology*, *23*(4), 615-624. https://doi.org/10.1016/j.conb.2013.02.017

Everitt, B. J., & Robbins, T. W. (2016). Drug addiction: Updating Actions to habits to compulsions ten years on. *Annual Review of Psychology*, *67*(1), 23-50. https://doi.org/10.1146/annurev-psych-122414-033457

Field, M., Heather, N., & Wiers, R. W. (2019). Indeed, not really a brain disorder: Implications for reductionist accounts of addiction. *Behavioral and Brain Sciences*, *42*, e9. https://doi.org/10.1017/S0140525X18001024

Fingarette, H. (1988). *Heavy drinking: The myth of alcoholism as a disease*. University of California Press. ark:/13960/t58d0j65h

Flanagan, O. (2013). Phenomenal authority. In N. Levy (Ed.), *Addiction and self-control* (pp. 67-93). Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199862580.003.0005

Foddy, B., & Savulescu, J. (2010). A liberal account of addiction. *Philosophy, Psychiatry, & Psychology*, *17*(1), 1-22. https://doi.org/10.1353/ppp.0.0282

Garson, J. (2024). Brain disorders, dysfunctions, and natural selection: Commentary on Jefferson. *Philosophical Psychology*, *37*(3), 558-569. https://doi.org/10.1080/09515089.2023.2232375

Glackin, S. N., Roberts, T., & Krueger, J. (2021). Out of our heads: Addiction and psychiatric externalism. *Behavioural Brain Research*, *398*, 112936. https://doi.org/10.1016/j.bbr.2020.112936

Graham, G. (2010). *The disordered mind: An introduction to philosophy of mind and mental illness*. Taylor & Francis. https://doi.org/10.4324/9780203857861

Hadland, S. E., Park, T. W., & Bagley, S. M. (2018). Stigma associated with medication treatment for young adults with opioid use disorder: A case series. *Addiction Science & Clinical Practice*, *13*(1), 15. https://doi.org/10.1186/s13722-018-0116-2

Hall, W., Carter, A., & Forlini, C. (2015). The brain disease model of addiction: Is it supported by the evidence and has it delivered on its promises? *The Lancet Psychiatry*, *2*(1), 105-110. https://doi.org/10.1016/S2215-0366(14)00126-6

Hammer, R. R., Dingel, M. J., Ostergren, J. E., Nowakowski, K. E., & Koenig, B. A. (2012). The experience of addiction as told by the addicted: Incorporating biological understandings into self-story. *Culture, Medicine and Psychiatry*, *36*(4), 712-734. https://doi.org/10.1007/s11013-012-9283-x

Hänninen, V., & Koski-Jännes, A. (1999). Narratives of recovery from addictive behaviours. *Addiction*, *94*(12), 1837-1848. https://doi.org/10.1046/j.1360-0443.1999.941218379.x

Hart, C. (2013). *High price*. Harper Perennial.

Hart, C. (2017). Viewing addiction as a brain disease promotes social injustice. *Nature Human Behaviour*, *1*(3), 0055. https://doi.org/10.1038/s41562-017-0055

Heather, N. (2013). Is alcohol addiction usefully called a disease? *Philosophy, Psychiatry, & Psychology*, *20*(4), 321-324. https://doi.org/10.1353/ppp.2013.0050

Heather, N. (2017). Q: Is addiction a brain disease or a moral failing? A: Neither. *Neuroethics*, *10*(1), 115-124. https://doi.org/10.1007/s12152-016-9289-0

Heather, N., Field, M., Moss, A., & Satel, S. (Eds.) (2022). *Evaluating the brain disease model of addiction*. Routledge.

Heilig, M., MacKillop, J., Martinez, D., Rehm, J., Leggio, L., & Vanderschuren, L. J. M. J. (2021). Addiction as a brain disease revised: Why it still matters, and the need for consilience. *Neuropsychopharmacology*, *46*(10), 1715-1723. https://doi.org/10.1038/s41386-020-00950-y

Heim, D. (2014). Addiction: Not just brain malfunction. *Nature*, *507*, 40. https://doi.org/10.1038/507040e

Henden, E. (2018). Addiction as a disorder of self-control. In H. Pickard & S. Ahmed (Eds.), *The Routledge Handbook of Philosophy and Science of Addiction* (pp. 45-53). Routledge.

Heyman, G. M. (2009). *Addiction: A disorder of choice*. Harvard University Press.

Heyman, G. M., & Mims, V. (2017). What addicts can teach us about addiction: A natural history approach. In N. Heather & G. Segal (Eds.), *Addiction and choice* (pp. 385-408). Oxford University Press. https://doi.org/10.1093/acprof:oso/9780198727224.003.0021

Holden, T. (2012). Addiction is not a disease. *Canadian Medical Association Journal*, *184*(6), 679. https://doi.org/10.1503/cmaj.112-2033

Holton, R., & Berridge, K. (2013). Addiction between compulsion and choice. In *Addiction and self-control*. Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199862580.003.0012

Hyman, S. E. (2007). The neurobiology of addiction: Implications for voluntary control of behavior. *American Journal of Bioethics*, *7*(1), 8-11. https://doi.org/10.1080/15265160601063969

Jefferson, A. (2022). *Are mental disorders brain disorders?* Routledge.

Juliano, L. M., & Brandon, T. H. (1998). Reactivity to instructed smoking availability and environmental cues: Evidence with urge and reaction time. *Experimental and Clinical Psychopharmacology*, *6*(1), 45-53. https://doi.org/10.1037/1064-1297.6.1.45

Kalivas, P. W., & Volkow, N. D. (2005). The neural basis of addiction: A pathology of motivation and choice. *American Journal of Psychiatry*, *162*(8), 1403-1413. https://doi.org/10.1176/appi.ajp.162.8.1403

Kavanagh, D. J., Andrade, J., & May, J. (2005). Imaginary relish and exquisite torture: The elaborated intrusion theory of desire.

Psychological Review, *112*(2), 446-467. https://doi.org/10.1037/0033-295X.112.2.446

Kennett, J. (2013). Just say no? Addiction and the elements of self-control. In N. Levy (Ed.), *Addiction and self-control* (pp. 144-164). Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199862580.003.0008

Kvaale, E. P., Haslam, N., & Gottdiener, W. H. (2013). The 'side effects' of medicalization: A meta-analytic review of how biogenetic explanations affect stigma. *Clinical Psychology Review*, *33*(6), 782-794. https://doi.org/10.1016/j.cpr.2013.06.002

Leshner, A. I. (1997). Addiction is a brain disease, and it matters. *Science*, *278*(5335), 45-47. https://doi.org/10.1126/science.278.5335.45

Levy, N. (2013). Addiction is not a brain disease (and it matters). *Frontiers in Psychiatry*, *4*. https://doi.org/10.3389/fpsyt.2013.00024

Levy, N. (2014). Addiction as a disorder of belief. *Biology and Philosophy*, *29*(3), 337-355. https://doi.org/10.1007/s10539-014-9434-2

Lewis, M. (2015). *The biology of desire: Why addiction is not a disease*. Public Affairs. ark:/13960/s23dq7f9rfr

Lewis, M. (2017). Addiction and the brain: Development, not disease. *Neuroethics*, *10*(1), 7-18. https://doi.org/10.1007/s12152-016-9293-4

Lie, A. K., Hansen, H., Herzberg, D., Mold, A., Jauffret-Roustide, M., Dussauge, I., Roberts, S. K., Greene, J., & Campbell, N. (2022). The harms of constructing addiction as a chronic, relapsing brain disease. *American Journal of Public Health*, *112*(S2), S104–S108. https://doi.org/10.2105/AJPH.2021.306645

Lilienfeld, S., & Marino, L. (1995). Mental disorder as a Roschian concept: A critique of Wakefield's "harmful dysfunction" analysis. *Journal of Abnormal Psychology*, *104*(3), 411-420. https://doi.org/10.1037/0021-843X.104.3.411

Maguire, E. A., Gadian, D. G., Johnsrude, I. S., Good, C. D., Ashburner, J., Frackowiak, R. S. J., & Frith, C. D. (2000). Navigation-related structural change in the hippocampi of taxi drivers. *Proceedings of the National Academy of Sciences*, *97*(8), 4398-4403. https://doi.org/10.1073/pnas.070039597

Maiese, M. (2021). An enactivist reconceptualization of the medical model. *Philosophical Psychology*, *34*(7), 962-988. https://doi.org/10.1080/09515089.2021.1940119

McLellan, A. T., Lewis, D. C., O'Brien, C. P., & Kleber, H. D. (2000). Drug dependence, a chronic medical illness. *JAMA*, *284*(13), 1689-1695. https://doi.org/10.1001/jama.284.13.1689

Meurk, C., Carter, A., Partridge, B., Lucke, J., & Hall, W. (2014). How is

acceptance of the brain disease model of addiction related to Australians' attitudes towards addicted individuals and treatments for addiction? *BMC Psychiatry*, *14*(1), 373. https://doi.org/10.1186/s12888-014-0373-x

Millikan, R. (1984). *Language, thought and other biological categories*. The MIT Press.

Murphy, D. (2009). Psychiatry and the concept of disease as pathology. In M. Broome & L. Bortolotti (Eds.), *Psychiatry as Cognitive Neuroscience: Philosophical perspectives* (pp. 103-118). Oxford University Press. https://doi.org/10.1093/med/9780199238033.003.0006

Neander, K. (1991). Functions as selected effects: The conceptual analyst's defense. *Philosophy of Science*, *58*(2), 168-184. https://doi.org/10.1086/289610

Palij, M., Rosenblum, A., Magura, S., Palij, M., Handelsman, L., & Stimmel, B. (1996). Daily cocaine use patterns: Effects of contextual and psychological variables. *Journal of Addictive Diseases*, *15*(4), 13-37. https://doi.org/10.1300/J069v15n04_02

Pescosolido, B. A., Martin, J. K., Long, J. S., Medina, T. R., Phelan, J. C., & Link, B. G. (2010). "A disease like any other"? A decade of change in public reactions to schizophrenia, depression, and alcohol dependence. *American Journal of Psychiatry*, *167*(11), 1321-1330. https://doi.org/10.1176/appi.ajp.2010.09121743

Pickard, H. (2016). Denial in addiction. *Mind and Language*, *31*(3), 277-299. https://doi.org/10.1111/mila.12106

Pickard, H. (2018). The puzzle of addiction. In H. Pickard & S. Ahmed (Eds.), *The Routledge Handbook of Philosophy and Science of Addiction* (pp. 9-22). Routledge.

Pickard, H. (2022a). Addiction and the meaning of disease. In N. Heather, M. Field, A. Moss, & S. Satel (Eds.), *Evaluating the brain disease model of addiction* (pp. 321-338). Routledge.

Pickard, H. (2022b). Is addiction a brain disease? A plea for agnosticism and heterogeneity. *Psychopharmacology*, *239*(4), 993-1007. https://doi.org/10.1007/s00213-021-06013-4

Pober, J. M. (2013). Addiction is not a natural kind. *Frontiers in Psychiatry*, *4*(4), 123. https://doi.org/10.3389/fpsyt.2013.00123

Redish, A. D., Jensen, S., & Johnson, A. (2008). A unified framework for addiction: Vulnerabilities in the decision process. *Behavioral and Brain Sciences*, *31*(4), 415-487. https://doi.org/10.1017/S0140525X0800472X

Rise, J., & Halkjelsvik, T. (2019). Conceptualizations of addiction and moral responsibility. *Frontiers in Psychology*, *10*, 1483. https://doi.

org/10.3389/fpsyg.2019.01483

Roberts, T., Krueger, J., & Glackin, S. (2019). Psychiatry beyond the brain: Externalism, mental health, and autistic spectrum disorder. *Philosophy, Psychiatry, & Psychology*, *26*(3), E-51-E-68. https://doi.org/10.1353/ppp.2019.0030

Ross, D. (2020). Addiction is socially engineered exploitation of natural biological vulnerability. *Behavioural Brain Research*, *386*, 112598. https://doi.org/10.1016/j.bbr.2020.112598

Saloner, B., & Cook, B. L. (2013). Blacks and hispanics are less likely than whites to complete addiction treatment, largely due to socioeconomic factors. *Health Affairs*, *32*(1), 135-145. https://doi.org/10.1377/hlthaff.2011.0983

Satel, S., & Lilienfeld, S. (2014). Addiction and the brain-disease fallacy. *Frontiers in Psychiatry*, *4*, 141. https://doi.org/10.3389/fpsyt.2013.00141

Satel, S., & Lilienfeld, S. (2017, June 22). Calling it a "brain disease" makes addiction harder to treat. *The Boston Globe*.

Schroeder, T., & Arpaly, N. (2013). Addiction and blameworthiness. In N. Levy (Ed.), *Addiction and self-control* (pp. 214–238). Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199862580.003.0011

Segal, A., Tiego, J., Parkes, L., Holmes, A. J., Marquand, A. F., & Fornito, A. (2024). Embracing variability in the search for biological mechanisms of psychiatric illness. *Trends in Cognitive Sciences*. https://doi.org/10.1016/j.tics.2024.09.010

Segal, G. M. A. (2013). Alcoholism, disease, and insanity. *Philosophy, Psychiatry, & Psychology*, *20*(4), 297-315. https://doi.org/10.1353/ppp.2013.0059

Sinnott-Armstrong, W. (2013). Are addicts responsible? In N. Levy (Ed.), *Addiction and self-control* (pp. 122-143). Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199862580.003.0007

Sinnott-Armstrong, W., & Pickard, H. (2013). What is addiction? In K. W. M. Fulford, M. Davies, R. G. T. Gipps, G. Graham, J. Z. Sadler, G. Stanghellini, & T. Thornton (Eds.), *The Oxford handbook of philosophy and psychiatry*, *vol. 1* (pp. 851-864). Oxford University Press. https://doi.org/10.1093/oxfordhb/9780199579563.013.0050

Sisti, D., Young, M., & Caplan, A. (2013). Defining mental illnesses: Can values and objectivity get along? *BMC Psychiatry*, *13*(1), 346. https://doi.org/10.1186/1471-244X-13-346

Skinner, M. D., & Aubin, H. J. (2010). Craving's place in addiction theory: Contributions of the major models. *Neuroscience and Biobehavioral Reviews*, *34*(4), 606-623. https://doi.org/10.1016/j.

neubiorev.2009.11.024

Snoek, A. (2017a). *Addiction, self-control and the self: An empirical, ethical study*. Castricum.

Snoek, A. (2017b). How to recover from a brain disease: Is addiction a disease, or is there a disease-like stage in addiction? *Neuroethics*, *10*(1), 185-194. https://doi.org/10.1007/s12152-017-9312-0

Sripada, C. (2022). Impaired control in addiction involves cognitive distortions and unreliable self-control, not compulsive desires and overwhelmed self-control. *Behavioural Brain Research*, *418*, 113639. https://doi.org/10.1016/j.bbr.2021.113639

Tiffany, S. T. (1990). A cognitive model of drug urges and drug-use behavior: Role of automatic and nonautomatic processes. *Psychological Review*, *97*(2), 147-168. https://doi.org/10.1037/0033-295X.97.2.147

Verhulst, B., Neale, M. C., & Kendler, K. S. (2015). The heritability of alcohol use disorders: A meta-analysis of twin and adoption studies. *Psychological Medicine*, *45*(5), 1061-1072. https://doi.org/10.1017/S0033291714002165

Volkow, N. D., & Koob, G. (2015). Brain disease model of addiction: Why is it so controversial? *The Lancet Psychiatry*, *2*(8), 677-679. https://doi.org/10.1016/S2215-0366(15)00236-9

Volkow, N. D., Koob, G. F., & McLellan, A. T. (2016). Neurobiologic advances from the brain disease model of addiction. *New England Journal of Medicine*, *374*(4), 363-371. https://doi.org/10.1056/nejmra1511480

Wakefield, J. C. (1992). The concept of mental disorder: On the boundary between biological facts and social values. *American Psychologist*, *47*(3), 373-388. https://doi.org/10.1037/0003-066X.47.3.373

Wakefield, J. C. (2020). Addiction from the harmful dysfunction perspective: How there can be a mental disorder in a normal brain. *Behavioural Brain Research*, *389*, 112665. https://doi.org/10.1016/j.bbr.2020.112665

Wakefield, J. C. (2021). Must social values play a role in the harm component of the harmful dysfunction analysis? Reply to Rachel Cooper. In L. Faucher & D. Forest (Eds.), *Defining mental disorder: Jerome Wakefield and his critics* (pp. 553-576). The MIT Press. https://doi.org/10.7551/mitpress/9949.003.0034

World Health Organization. (2019). *International Classification of Diseases, Eleventh Revision (ICD-11)*. https://icd.who.int/browse11

Xiao, P., Dai, Z., Zhong, J., Zhu, Y., Shi, H., & Pan, P. (2015). Regional gray matter deficits in alcohol dependence: A meta-analysis of voxel-based morphometry studies. *Drug and Alcohol Dependence*, *153*, 22-28. https://doi.org/10.1016/j.drugalcdep.2015.05.030

Yang, L. H., Wong, L. Y., Grivel, M. M., & Hasin, D. S. (2017). Stigma and substance use disorders. *Current Opinion in Psychiatry*, *30*(5), 378-388. https://doi.org/10.1097/YCO.0000000000000351

Received 1st December 2023; revised 14th November 2024; accepted 19th November 2024.