RESEARCH ARTICLE

Drug overdose deaths, addiction neuroscience and the challenges of translation [version 2; peer review: 1 approved, 2 not approved]

Previously titled: Crisis, what crisis? Addiction neuroscience and the challenges of translation

Samuel McLean, Nikolas Rose

Department of Global Health and Social Medicine, King’s College London, London, WC2R 2LS, UK

Abstract
In this article, we argue that the rapid rise in drug overdose deaths in America is a tragedy that draws attention to fundamental conceptual and experimental problems in addiction science that have significant human consequences. Despite enormous economic investment, political support and claims to have revolutionised addiction medicine, neurobiological models are yet to produce a treatment for substance addiction. This is partly, we claim, because neurobiology is unable to explain essential features of addiction and relapse that neurobehavioral models of addiction are better placed to investigate. We show how addiction neuroscience turned to long-term memory to explain the chronicity of addiction and persistent relapses long after neurochemical traces have left the body. The turn to memory may in time help to close the translational gap facing addiction medicine, but it is our view in this article that the primary value of memory theory lays in its potential to create new critical friendships between biological and social sciences that are attuned to the lived experience and suffering of stigmatised people. The value of the memory turn may rest upon the capacity of these critical friendships to wean addiction science off its long-term dependence on disease concepts of human distress.

Keywords
Drug addiction, opioid crisis, translation, addiction neuroscience, the memory turn, dopamine theory, critical friendship

Open Peer Review

Reviewer Status

Invited Reviewers

<table>
<thead>
<tr>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>X</td>
<td>X</td>
<td>✓</td>
</tr>
</tbody>
</table>

version 2
(revision)
29 Jan 2021

version 1
14 Sep 2020

1. Matt Field, University of Sheffield, Sheffield, UK
2. Philippe Bourgois, University of California, Los Angeles, Los Angeles, USA
3. Harold Kalant, University of Toronto, Toronto, Canada

Any reports and responses or comments on the article can be found at the end of the article.
Introduction

Drug addiction is a disorder of long-term memory


In this article we review the claim that there is an ‘addiction crisis’ [NIDA, 2018] and argue that the crisis represents both a challenge and an opportunity for addiction neuroscience, as it seeks to translate its emerging memory concept of addiction into medical treatments. We argue, however, that the practical and theoretical value of thinking about addiction as a problem of long-term memory extends beyond its translational utility. Its potential, we claim, is located in its capacity to establish a critical friendship of social and biological science, which produces a new kind of addiction science that does not see human relations with drugs through a conceptual schema of disease, disorder and death. The so-called addiction crisis concerns rapid rises in drug overdose fatalities in the United States fuelled by opioid use. This poses fundamental questions of knowledge and translational research that are by no means specific to the United States, or even addiction. Most medical and psychiatric research on addiction since the 1970s has been informed by a neuroscientific theory focused upon dopamine which is linked to excessive pleasure seeking [Wise, 1978]. We argue that research based on this hypothesis has not fulfilled the revolutionary promises made on its behalf by organisations such as National Institute on Drug Abuse [NIDA] and the American Psychiatric Association [APA]. This failure to deliver on grand promises and unrealistic expectations is partly because they do not accord with human lived experience of addiction, now commonly termed a severe form of substance use disorder [Volkow et al., 2016]. It is also related, we argue, to a dependence on reductive neurobiological models and pharmacological interventions that are isolated from social conditions of distress that so often condition lives marked by addiction. The translational problems facing addiction neuroscience, however, have deeper epistemological roots in how human problems are perceived and known in our ‘neurobiological age’ (Rose, 2013).

Despite the emergence of a science that claims addiction is not a matter of will or personality but has a biological basis in the brain, and despite decades of research into this neurobiological basis of substance addiction that remains generously supported and funded by the National Institute of Health [NIH], this knowledge has not yet been converted into treatments or policies that tackle addiction and reduce relapses. Thus, proposals to tackle opioid-related deaths are dishearteningly familiar. Reduce prescriptions of opioid based pain killers for pain so that they do not leak into the illegal market. Convert those who have become opioid users to alleviate chronic pain to non-pharmacological means of pain control - CBT, mindfulness, acupuncture. Try to mitigate the dangers of illegal use by harm reduction, moving those who have become dependent on opioids onto less addictive drugs such as buprenorphine. Educate prescribers, educate children, educate those experiencing chronic pain, educate actual and potential ‘addicts’ as to the dangers. And from the social scientists, the familiar critique of ‘medicalisation’ and of the powers of big pharma to influence prescribing practices now exemplified in the rise and rise of OxyContin and its consequences.

The first half of this article outlines the tragedy of rising drug overdose mortality and why the focus on errant prescription practices in popular discourse at the expense of social determinants of suffering is part of the problem. The second half looks at how conceptual and experimental developments in addiction neuroscience which we term ‘the memory turn’ are better placed to investigate and comprehend the chronic, relapsing kind of addiction that resists medicalised and non-medicalised treatments. Significant changes are taking place in addiction neuroscience that are not yet well-known within social studies of addiction despite the fact that they date back at least two decades. As Steve Hyman, then director of National Institute of Mental Health, said in 2005: ‘Based on the available

---

1 The authors would like to thank the reviewers, Philippe Bourgeois and Harold Kalant for their reviews of the first revision of this article. Their contributions have had a substantive impact on the revised article and certainly improved it.

2 We use the terms ‘drug addiction’ and ‘substance addiction’ interchangeably.

3 The terminology is contested [Hasin et al., 2013; O’Brien, 2011]. We have, however, chosen to speak of addiction rather than dependence. Dependence is a ‘normal’ adaptive response. Everyone who takes psychoactive substances for an extended period becomes dependent, but only a small percentage of people develop compulsive, chronic drug-seeking that characterize the clinical category of addiction. ‘Substance use disorder’ is the clinical name for drug addiction in the Diagnostic and Statistical Manual of Mental Disorders IV, but the overall title is “Substance and Addiction-Related Disorders.” [APA, 2013].
neuroscientific evidence from the molecular to the behavioural, addiction is best explained as a disorder of memory and learning”. In this ‘memory turn’ drug memory science is developing memory models that conceptualise the ways that changes in memory systems mediate interactions of brain, behaviour and environment that underlie addiction and relapse.

Memory theory has begun the difficult transition from lab to clinic, and there are indications that it may prove to underpin effective treatment strategies. And yet one of the most telling contributions of drug memory science is to articulate at the level of neurobehavioral systems why addiction is actually so resistant to treatment. We conclude, however, by imagining another kind of future for the memory turn that leads us to a new kind of addiction science. This science would be based in neurosocial collaborations that interrogate the ways in which histories of violence, exclusion and stigmatisation express themselves in functional and structural changes to brain, nervous system and body.

The argument advanced in this article is based upon four years of research and collaboration with addiction neuroscientists (McLean, 2019). Integral to the method was the “history of concepts” (Foucault, 1966 [1991]). This is the study of how concepts emerge within a field of knowledge and how they reshape the ways that researchers within that field perceive and think about their object of enquiry. This research was underpinned by interviews and conversations with addiction neuroscientists and historians, and time spent in a leading neuroscience laboratory in the UK. The analysis was developed in a number of workshops organised by the Neuroscience and Society Network (NSN) based at King’s College London. This is an interdisciplinary network which facilitates collaborations between researchers in the life sciences and the human sciences in the spirit of ‘critical friendship (Rose & Abi-Rached, 2013). Indeed, this article grows out of an international NSN workshop titled Memory/Habit/Addiction that took place in early 20204.

Drug overdose deaths
Disease concepts of human experience tend to have a bad reputation in the social sciences and humanities. And for good reasons. In such approaches, social conditions of distress, from socioeconomic deprivation to racism and social exclusion, are almost always reduced to matters of individual pathology. This has the effect of turning our critical attention away from the ways in which power and violence are often manifested in states of health and sickness. That neuroscientific research is often subject to the same kind of criticisms is unsurprising, since concepts of disease and illness are increasingly determined by how neurobiology perceives and knows them. Addiction is no different.

Despite many differences within addiction neuroscience, there are certain beliefs that the majority of addiction researchers can agree upon irrespective of discipline, method, or training. One is that some human beings develop relationships with drugs they wish to change, but have great difficulty doing so ‘organically’ without specialist support. Another is that only a small proportion of drug users experience adverse life-changing problems due to drug use – support networks breaking down, jobs and homes being lost. Few researchers in the field would accept the stereotypical vision of the drug user as a person whose life has been emptied of meaning and value beyond finding the next ‘hit’ or ‘score’. Drug addiction is far less common than domestic drug policy in the United States or United Kingdom would suggest (Garriott, 2011; Nutt, 2012), or one would imagine based on news reporting. This is especially true in the case of ‘hard’ illegal drugs such as heroin, crack cocaine and methamphetamine (Reinarman & Levine, 1997). And yet when debate and theorising about the ‘causes’ of addiction cools down, we are still left with the reality. That human beings suffer from unattended misery they themselves attribute to drug use, and typically desire relief from their affliction which is hard to find. No critique of disease concepts – however justified or well-intentioned – changes this painful truth of social life today.

According to the National Survey on Drug Use and Health, approximately 19.7 million people in the United States, aged twelve and older, battled with “substance use disorder” in 2017 (NSDUH, 2018). The National Institute on Drug Abuse (NIDA, 2018) describes drug addiction as a “national crisis” in the United States. The costs to health, society and economic welfare, NIDA argues, are “devastating”. This is represented most strikingly in drug-related mortality statistics over the past two decades. From 1999 to 2018, 817,000 people are estimated to have died from drug overdoses, making it a leading cause of injury-related death in the United States (Hedegaard et al., 2020). The Centers for Disease Control and Prevention situate the “opioid epidemic”. In the same period of time, opioid-related overdose deaths – those including prescription opioids, heroin, and synthetic opioids like fentanyl – increased almost six times (Hedegaard et al., 2020).

In 2018, opioids were involved in seventy per cent of all drug overdose deaths in the United States, with synthetic opioids (other than methadone) accounting for nearly two-thirds of these drug overdose deaths (Wilson et al., 2020). The total “economic burden” of prescription opioid misuse alone in the United States is estimated at $78.5 billion a year (Florence et al., 2016). According to the European Monitoring Centre for Drugs and Drug Addiction, although the number of drug overdose deaths is far lower in the United Kingdom, similar trends appear to be developing (EMCDDA, 2017). In 2018, England and Wales recorded the highest number and the highest annual increase in drug overdose deaths since the time series began in 1993 (ONS, 2019). Britain also leads Europe in drug overdose deaths: almost 1 in 3 drug overdose deaths in Europe occurred in the UK in 2018 – twice the number of those occurring in Germany, which is in second place (EMCDDA, 2017).

The problems of suffering and death due to drugs are real enough, as these statistics show. Yet discourse of crisis that pervades debates around addiction – especially in the United

---

4 This proceedings of this Workshop are reported at http://somatosphere.net/2020/memory-habit-addiction.html/
States - is unfortunate since it may reinforce moral panic [Cohen, 1972] around a population of people classified as addicted to drugs who are already overburdened by stigma and chronic harms of social exclusion [Bourgois, 2000]. For this reason, the language of crisis ought to be avoided, replaced by an emphasis on the human tragedy of drug overdose fatalities.

A social autopsy of drug overdose deaths

The rise in drug overdose deaths fuelled by opioid use has largely been framed – at least in popular discourse – as a “crisis of over prescription” [Nutt, 2012]. This diagnosis is not merely naïve, it has unfortunate consequences, for it turns our critical attention away from the profound social distress at the root of drug overdose deaths. It is important to recognise that these deaths are not specific to opioids5. And while the over prescription of OxyContin and other opioids clearly require urgent redress, this problem is itself a symptom of the structural distribution of social despair. We claim no originality in arguing this. Social scientists working on health and medicine have been pointing this out since the early decades of the twentieth century [Agar, 1973; Bourgois, 2003; Campbell & Shaw, 2008; Lindensmith, 1947; Preble & Casey, 1969], when addiction science first emerged as a distinct field of knowledge codified by biomedical discourse [Campbell, 2007], and the medical concept of addiction first aligned with the common use of the word addiction in the United States and Britain [Room, 1969]. But it has taken the recent work of two economists to ensure that this argument and its implications can no longer be avoided.

In Deaths of Despair and the Future of Capitalism [2020], Anne Case and Angus Deaton argue that, uniquely among high-income nations, life-expectancy in the United States has declined for three consecutive years for the first time since 1918. This, they argue, can be explained through the dramatic rise of “deaths of despair” as a result of drug overdose, suicide and alcoholism over the past two decades. Case and Deaton draw conclusions that recall those of Durkheim (1897[2002]) in his classic sociological study of suicide at the end of the 19th century. The oversupply of opioids, they argue, did not create the conditions for this despair, they filled the gaping hole opened up in white working-class communities by four decades of social and economic exclusion and fragmentation. For those without a college education, not only have median wages declined since 1979, and work security deteriorated, they have found themselves less valued in the economy and disconnected from the ‘American Dream’. Drug overdose deaths, like suicide and self-harm, they propose, follow from a loss of belonging, of self-worth and of hope amidst social and economic upheaval.

Framing drug overdoses as “deaths of despair” is gaining traction. For all the talk of ‘accidental overdose’ from prescription opioids, the relationship between suicide and opioid use is emerging as a significant area of research (Oquendo & Volkow, 2018). Early signs indicate that they are entangled in multiple ways (Volkow & Gordon, 2019). A 2017 study based on national survey data found two things of particular importance. That suicidal ideation was 40–60% higher for people who misused prescription opioids than those who do not. And that individuals with a prescription opioid use disorder were twice as likely to attempt suicide as those who did not misuse prescription opioids (Ashrafion et al., 2017). Although causality remains an area of debate, the underlying claim in these reports is that suicidal ideation and suicide are responses to the suffering of substance addiction that is itself an expression of social despair.

Isolating the opioid deaths from the wider problem of drug overdose and prescription opioids from illegal drugs not only diverts critical attention away from social despair, it may deepen social divisions by stratifying drug users into two broad groups based, it seems, more on prejudice and discrimination than knowledge and evidence. On the one hand, there are ‘bad addicts’ mixed up with illegal opioids, to be punished with the full force of the law. And on the other hand, ‘unfortunate dependents’ deserving of social and medical support having become unwillingly dependent upon legal opioids. This division has the potential to exacerbate hostile public attitudes towards drugs and those who use them, and to entrench stigma that remains a major barrier preventing people from seeking out and receiving support and treatment for drug-related life-problems (Hadland et al., 2018; Lloyd, 2010; Yang et al., 2017). It could even fuel already tense racial divisions. Note, as Chris McGreal does in American Overdose (2019), how the first group is typically associated in the public imagination with African American and Hispanics, and the second group with Whites.

This should not surprise anyone. For dividing opioid users into deserving and undeserving groups helped to create the conditions for a never-ending ‘war of drugs’ that has always had an ugly racial dimension. As David Courtwright (1982: 1) demonstrates in his history of opioid addiction in America, our modern medical concept of addiction emerged in the United States between 1865 and 1935 out of a fundamental change in medical perceptions of the ‘typical addict’. In this period the “addict profile” shifted from the “addicted matron”, middle-aged women of the middle or upper class, to the “street criminal”, lower-class urban men perceived to be African American or Hispanic. Indeed, we might ask if the opioid crisis would be viewed as an urgent public health problem by political and medical elites if problems with Fentanyl and OxyContin addiction were entangled in the public imagination with the lives of African American men in Harlem or Detroit.

Rise and fall of the dopamine hypothesis

It would, of course, be misguided to attribute the failure to address the structural distribution of despair underlying the ongoing opioid tragedy to the epistemic authority of neuroscientific research in addiction, that is to say, its power to determine how addiction is perceived and known. There are, however, striking parallels between the diagnosis of over prescription as a determining factor of the tragedy, on the one

5 From 2012 through 2018, drug overdose deaths involving cocaine more than tripled. In 2017, cocaine accounted for nearly 1 in 5 overdose deaths. And between 2017 and 2018, there was a thirty-seven per cent increase in overdoses involving psychostimulants such as methamphetamine.
hand, and on the other hand, neurobiological concepts of addiction focused on dopaminergic circuitry in the ‘reward system’. Both these ways of thinking tend to ‘externalise’ questions about the social conditions and lived experience that give rise to problematic patterns of opioid use and the addiction and dependence that can result. And both explanations are conducive to a culture of public policy which fetishizes ‘quick fixes,’ and that ultimately leave the underlying social causes of human suffering unchanged.

Nonetheless, as we shall see, not all neuroscientific models of addiction are conceptually closed to sociocultural factors, and certainly not to the same degree. You would not necessarily know it from reading Nature or Science, or most of the social scientific literature, but the models of addiction have evolved significantly since Alan Leshner [1997] prematurely claimed neurosciences were already “revolutionising” addiction science and medicine. While one neurotransmitter, dopamine, and one brain region, the striatum remain central, the experimental frame of reference now includes all brain regions and neurotransmitter systems. Glutamate [the so-called ‘memory molecule’] and prefrontal cortex have become particularly prominent [Kalivas, 2009; Kalivas & O’Brien, 2008; Kalivas & Volkow, 2005]. It is also now common for models to integrate various levels of analysis: neural, molecular, systemic, hormonal, behavioral and psychological [Volkow et al., 2016].

Behavioral and psychological mechanisms have had prominent roles in mainstream neuroscientific theories since at least the mid-1990s [Everitt & Robbins, 2005; Koob & Le Moal, 2001; Robinson & Berridge, 1993]. But the one neuroscientific theory that has dominated the representation of neuroscience in the critical social sciences [Campbell, 2010; Courtright, 2019; Raikhe, 2015; Vrecko, 2010] is one that is conceptually closed to the social and fuels stigmatising tropes about substance addiction. This is the model that was heavily promoted by the National Institute on Drug Abuse [NIDA] to the extent that the historian, David Courtright [2010: 137], calls it the “NIDA paradigm”. It views addiction as a “chronic, relapsing brain disease characterised by a loss of control over drug-taking”. According to this theory, it is high volumes of dopamine (the so-called ‘pleasure molecule’) in the brain’s ‘reward system’ that underlie the loss of self-control said to define addiction.

Because of the importance given to the dopamine pathway from midbrain to striatum, neuroscientists refer to this as the “dopamine theory” of addiction, or “dopamine hypothesis” (Marsden, 2006; Nutt et al., 2015; Robbins & Everitt, 1999; Wise & Rompre, 1989; Wise, 2008). Despite its historical prominence, many contemporary neuroscientists are highly critical of this way of thinking and the effect of its dominance on the translational efficacy of addiction neuroscience over four decades. The humbling truth, David Nutt et al. [2015] argue, is that despite significant investment and hype, dopamine theory is yet to produce a “single effective treatment” for addiction. Of course, the translation problem – the difficulty turning knowledge into treatments that ameliorate the suffering of the estimated 8-12 per cent of opioid users who develop opioid addiction – is hardly specific to dopamine theory or the neurosciences6. The reality on-the-ground is that addiction continues to evade medicalised and nonmedicalized treatments, which may suggest a fundamental intractability. This may be true. For several reasons, however, we consider the historical dominance of the dopamine hypothesis to be an important part of the translational problem. Let us state a few of these reasons before showing how addiction neuroscience is responding to some of the conceptual and experimental problems of the dopamine hypothesis.

The pleasure concept

First and foremost, the dopamine hypothesis is governed by a pleasure concept of addiction that is highly stigmatising. According to this concept, the loss of self-control that is said to define addiction coexists with excessive pleasure-seeking. Excessive – that is to say, pathological – since the result are significant harms to health, which place life itself at risk. It is sometimes said that moral explanations of addiction have given way to neurobiological alternatives [Campbell, 2011]. We see something different. We can observe in the NIDA paradigm traces of the normative content of the biblical myth of “fallen human nature” corrupted by the hedonic desire for ephemeral gratifications of the flesh, now translated into the language of brain chemistry not sin. The concept is stigmatising, since it shares with the myth, the judgement that while all humans desire hedonic pleasure, there is something peculiarly errant about those who become ‘prisoners’ to it. In this way, the pleasure concept reaffirms stigmatising tropes that can be found in Anglo-American cultures since at least the late eighteenth century that intoxicants and those who use them are inherently deviant [Levine, 1978; Reinarman & Levine, 1997; Valverde, 2008]. These tropes are themselves major impediments to drug recovery.

The pleasure concept of addiction also holds to a reductive account of human motivation, one that conceives of pleasure and pain as polarities. Yet this is at odds with the human experience of addiction, as it is documented in authoritative clinical, neuroscientific, and social studies (Bourgois, 2000; Everitt & Robbins, 2005; Everitt, 2014; Marsden et al., 2018; Volkow et al., 2016). This is particularly true with human opioid addiction. Two brief examples, one neurobiological, the other anthropological, make this point for us. First: neurobiology. The pleasure-seeking concept is at odds with both George Koob’s (2008, with Le Moal, 2008) influential work on stress and the dysregulation of affect, and basic knowledge of the neurochemical function of opioid receptors, first demonstrated by Candice Pert & Solomon Snyder (1973). This work suggests that opioids function in the nervous system of the body or in specific brain receptors to reduce the intensity of pain, something that accords closely with the lived experience of heroin addiction.

---

6 It should be pointed out that there is considerable debate regarding the efficacy of drug therapy. While some reviewers conclude that it has little value in the treatment of addiction [Kalant, 2015], others consider drug therapy of significant clinical value [Pierce et al., 2012; Schacht et al., 2018].
Second: anthropology. In Pastoral Clinic (2010: 5), a deeply moving ethnography of heroin addiction in New Mexico’s Española Valley, “home to the highest rate of heroin addiction and fatal overdoses in the United States”, Angela Garcia documents in elegiac detail, the motivating power of suffering in human addiction, and the complex ways grief and pain become intertwined with memories of pleasure and the uncertain promises they make. The promise of drugs here is not of pleasure or happiness, nor even satisfaction, but at best, temporal relief from life’s misery.

The pleasure concept is incompatible with human experience of addiction, we argue, in part because the dopamine theory it informs is conceptually closed to sociocultural and psychological dimensions of this experience. And yet, as noted above, dopamine theory has not only proven highly influential in shaping the direction of addiction neuroscience, but also has shaped the ways that social scientists, with few exceptions, perceive the neuroscience of addiction (Courtwright, 2019).

This is unfortunate because the field of addiction neuroscience does not operate with a unified conception of addiction. There are not only multiple neuroscientific theories of addiction (Campbell, 2011; Raikhel, 2015), there are competing dopamine models with different concepts of dopamine function. Most notable here, perhaps, Wolfgang Schultz’s (2007a, 2007b, 2011) influential work on “prediction error” in the striatum, amygdala and frontal cortex, which accounts for the difference between an “expected reward” based on prior experience (and thus memory of it) and the “actual reward” provided by the drug or behaviour. But most significantly, memory models of addiction are reshaping neurobiological styles of thought.

In the words of the Carmack et al., 2017: 523 “The overall hypothesis that addiction persists as a memory or memory-like process long after drug exposure has become the widely accepted position in the field.”

**Beyond the pleasure principle**

As the dopamine hypothesis begins to mutate, it is worth asking why it is’s epistemic authority was so long-lasting, despite its evident limitations? In part, perhaps, it may be because, as we have noted, the dopamine hypothesis retains resonances of the moralistic image of the insatiable quest for the pleasure of the hit overwhelming the will, conscience and the demands of propriety and civility: the science of addiction here retained the sense that what was at stake was a ‘disease of the will’ [Valverde, 2008]. Perhaps its because it linked so well with more general arguments in the neuroscience of psychopathologies, which focusses much attention on the functioning of individual neurotransmitters, often moving seamlessly from the experimental reductionism of laboratory-based research on animal models to a kind of metaphysical reductionism that extrapolates, without intermediate experimental work, to the vital lives of human beings. Yet human brains are many orders of magnitude greater than such experimental animals, their neural development is many orders of time longer, and their environmental transactions are immensely more rich and complex. And human lives are unintelligible without recognising that the world they inhabit is one of meanings and memories, shaped by myths and stories of lives, their own and others, of hopes and aspirations, of dreams and despair. To address these problems is not to reject neuroscience, but to argue for a different neuroscience that accepts the scientific necessity of moving beyond what the late Carl Woese [2004] termed ‘metaphysical reductionism’ in order to replace human brains in human organisms as they make their lives in a human social world.

No doubt many factors have been involved, as they always are in the overturning of a paradigm, or the mutation in a style of thought. Anomalies certainly accumulated, as we would expect from a reading of Thomas Kuhn (1970), and a younger group of research scientists, if not revolutionaries, certainly sought to mutate the old ways of thinking into something new. But not least among the factors that led to the waning of the dopamine hypothesis its inability to deliver what the neuropsychiatric establishment said it would. Namely, treatments that tackle addiction and relapse, and no less significantly, to reduce the stigma of addiction by showing it to be a ‘brain disease’. In a major intervention in the New England Journal of Medicine Volkow et al. [2016: 363] claimed that the ‘brain disease model’ has delivered ‘scientific advances in the prevention and treatment of substance-use disorder and related developments in public policy’. Yet the evidence for this claim, as the translational gap indicates, is thin at best [Hammer et al., 2013; Heather et al., 2018], and for one of the reasons the authors cite themselves. ‘Compulsive and impulsive behaviours’, they write, that are essential to the clinical state of addiction, ‘are not clearly tied to neurobiology’ [Volkow et al., 2016].

This major gap in our knowledge and treatment of substance addiction will not be closed by directing yet more money and resources to neurobiological and animal models. The compulsive and impulsive behaviors that are hallmarks of chronic, relapsing addiction cannot be explained through neural maladaptations alone. This is hardly surprising, since addiction is the product of an extraordinary range of factors [Kalant, 2010] that include unconscious and conscious memories of addiction-related experience established between drugs and their associated cues within the lived environment and body. In the next section, we show how through its turn to memory, addiction neuroscience is moving beyond its initial pleasure concept and the limitations of the dopamine hypothesis.

**The memory turn**

The turn to memory began at least twenty-five years ago. We can date it to a paper by Norman White [1996] in which he argues that ‘drug-related information’ [e.g. affects and cognitions] is ‘stored in anatomically distinct memory systems that contribute in distinct ways to the overall clinical state of drug addiction and relapse’. This view has received extensive support [Goodman & Packard, 2016] and is now one of several neuroscientific models in which maladaptations in long-term memory are vital to explaining addiction and relapse. But as is so often the case in the history of ideas and science, the ‘new’ is often a novel repetition of the ‘old’ in new language [Deleuze, 1995; Foucault, 1969]. The memory turn is
a case in point. Its neurobehavioral theory is powered by the resurgence of learning models of addiction [Everitt & Robbins, 2016; Lewis, 2018]. The link between learning and memory, on the one hand, and clinical states of addiction, on the other, is hardly new. It is nearly as old as addiction science itself. As Charles O’Brien [2014] reminds us, learning and memory were essential to Abraham Wikler’s classic neuropsychiatric studies of conditioned withdrawal and relapse [Wikler, 1961; Wikler, 1965].

More than a century ago, Pavlov [1927] demonstrated why separating learning and memory experimentally is difficult, perhaps impossible, to establish. For learning is empirically observable in patterns of behavioral response to stimuli such as tolerance and sensitisation while memory is not [Kandel, 2007]. Memory, rather, is hypothesised to explain how learning occurs without being directly observable [Hebb, 1949]. The distinction, therefore, is conceptual. It is also crucial to memory theories of addiction. Since while learning and memory are viewed as co-dependent, causal emphasis is placed upon memory before learning. The underpinning hypothesis is that maladaptations in long-term memory are provide the best available neuroscientific account of treatment resistant addiction and relapse. That is to say, these ‘neurobehavioral’ theories provide the most cogent explanation as to why addiction-related habits are as resilient as they are, why addiction-related associations are more easily re-established if disrupted than when they are established for the first time, and why learning displayed in the acquisition of non-addiction-related habits is so difficult to achieve and maintain over time.

Of course, neurobiological and neurobehavioral theories are not mutually exclusive. It is well-known that prolonged intake of large amounts of psychoactive substances lead to functional even structural changes in the brain that impact on the severity of addiction and the risk of relapse [Berrios-Cárcamo et al., 2020]. But by the mid-1990s, neuroscientists and psychologists were already arguing that memory theory is well-suited to explain vital features of addiction that neurobiological models of dopamine [or any neurotransmitter or molecule for that matter] are unable to account for. Two features in particular. First, why compulsive drug-seeking persists despite explicit knowledge of often severe harms to health and risks to life. And second, and no less important, why compulsions and cravings that result in relapses, can persist for years, even decades, after all detectable neurochemical traces have left the body. Most neuroscientists and psychologists recognised that to explain these features of addiction and relapse, their models needed to investigate how brain mechanisms identified through neurochemical animal models interact with behavioral, environmental and psychological factors [Everitt & Robbins, 2016; Hyman, 2005; Robinson & Berridge, 1993; Robinson & Berridge, 1998; White, 1996] that both ‘trigger’ these mechanisms and guide them towards specific targets [Kalant, 2010; Kalant, 2015]. This is what memory theory aims to do.

Around the same time, addiction and memory scientists started to converge on a shared realisation that addiction and long-term memory share molecular, neural and cellular pathways and systems (Hyman et al., 2006; Kelley, 2004). To understand one, is to know the other. Hence this intimate relationship between chronic, relapsing addiction on the one side, and long-term memory on the other, may provide a basis for exploring and making sense of those vital features of substance addiction (Nestler, 2001; Nestler, 2013; Nestler et al., 1993; Robbins et al., 2008). This insight formed the basis for a new field of knowledge – drug memory science – that is organised around three distinct if mutually reinforcing lines of enquiry. The first is neurobehavioral studies of the roles played by drug memory systems and conditioned learning in addiction and relapse [Everitt & Robbins, 2016; Robbins & Everitt, 2001; White, 1996]. The second is the analysis of how ‘drug memories’ – i.e. memories associated with addiction via the affects and cognitions of drugs such as opioids –underlie compulsive ideation and behaviour and weaken the capacity to acquire new habits (Milton & Everitt, 2010). The third is research on the ways that substance addiction damages neurobiological mechanisms involved in long-term memory from synaptic plasticity to glutamate transmission (Kalivas & O’Brien, 2008; Kauer & Malenka, 2007; Lüscher & Malenka, 2011).

Underlying this field of knowledge is the memory concept. In sum this is the idea that drug-related experiences that underlie addiction and relapse are represented in memories of these experiences and in the multiple memory systems that make these memories possible. Memory models seek to understand the interactions of drug memories and drug memory systems in different parts of the nervous system [from midbrain to prefrontal cortex] and at different levels of analysis [from molecules to behavior]. Drug memories come in different kinds. They can be declarative and explicit; – those people consciously recall and associate with their drug-related experience – the euphoria and the suffering, satisfaction and displeasure. But drug memories are for the most part procedural and implicit; unconscious, somatic memories of drug-related experience (Squire, 2004). The latter are causally more significant. Declarative memories function by strengthening both procedural memories and existing associations between drugs and environmental and physiological cues. Drug memories are maladaptive (Milton & Everitt, 2012) because they are strongly affective and extremely difficult to disrupt and change once consolidated the first time (Tronson & Taylor, 2013). Functional, perhaps even structural, changes to memory systems in the brain underlie aberrant drug memories and pathological learning (Robbins et al., 2008; Robbins & Everitt, 2001).

To make sense of this process, neural, molecular and behavioural knowledge of the mesolimbic dopamine system (the pathway connecting the ventral tegmental area in the midbrain to the ventral striatum of the basal ganglia in the forebrain) has been integrated into new knowledge in our understanding of “corticostriatal systems” that connect the striatum to the prefrontal cortex (Graybiel, 2008; Graybiel & Grafton, 2015). It might be argued that the memory turn is giving new meaning to the longstanding description of drug addiction as a ‘habit’. As Robbins & Everitt (1999) put it: “Bad habits add up.”
From bench to clinic
This turn to memory, we are arguing, is a significant scientific event in addiction neuroscience. A claim reflected in both what many senior neuroscientists and psychiatrists say and experimental work itself [Lewis, 2018]. Indeed, as Everitt & Robbins [2016] argue, theories of learning and memory inform all major neuroscientific theories of drug addiction and relapse. A new way of seeing and knowing substance addiction as a problem of memory has taken shape, which understands he great difficulty in forgetting drug memories and learning new drug-free habits as due to the pathological usurpation of multiple memory systems. Yet given the extreme difficulty treating chronic addiction and relapse, and the time it takes to turn basic research into treatments, it is too soon to know what the therapeutic value of memory models could be.

What we do know is that the translational research is underway that seeks to use the evidence from memory studies to prevent addiction relapse [Miller & Marshall, 2005; Volkow et al., 2016; Xue et al., 2012]. There are signs memory theory is starting to inform the clinical classification of addiction. The latest edition of the Diagnostic and Statistical Manual of Mental Disorders (DSM-5, 2013: 481) includes for the first time “powerful memories” as part of the core symptomatology of “substance use disorder”[1]. So far, clinical research informed by the memory concept of addiction is largely focused on pharmacological interventions, in particular the search of chemical agonists and antagonists able to disrupt drug memories. This work by either by introducing reconsolidation so they can be modified, or extinction where associations between environmental cues and conditioned responses are broken altogether (Lee et al., 2005; Merlo et al., 2014; Miller & Marshall, 2005; Torregrossa & Taylor, 2012). This resembles the development of propranolol research for post-traumatic stress disorder (PTSD) in the late 1990s and early 2000s. Propranolol beta-blocker, was found to be effective in disrupting fear memories through a process of memory reconsolidation; it does not cure PTSD, but can relieve some of its somatic symptoms [Debiec et al., 2006; Nader et al., 2000]. (Debiec et al., 2006; Nader et al., 2000).

The memory turn has also started to make the vital if difficult transition from animal models to clinical human research. Here we would like to highlight one example of how drug memory science is beginning to inform clinical human research. The example is drawn from interdisciplinary research we are doing as part of the Neuroscience and Society Network [NSN] at King’s College London cited earlier in this article. The clinical work of Professor John Marsden and colleagues at the Institute of Psychiatry, Psychology and Neuroscience (IoPPN) at King’s College London on cocaine use disorder (CUD) (2018), shows how drug memory science is moving beyond established therapeutic targets of physical dependence and withdrawal towards understanding the long-lasting effects addiction has upon cognition and behaviour.

Marsden and colleagues have developed and piloted a novel Memory-focused Cognitive Therapy (MFCT) to tackle cocaine use disorder [CUD]. Given the enduring effects of cocaine-related conditioning in patients, cue-induction procedures are used to elicit cocaine-related cognitions in patients, with the aim of reducing craving for the stimulant. The rationale is essentially Pavlovian: when repeated cue-exposure goes unrewarded, strong associations between stimulant-cues-responses established in addiction are broken over time. MFCT also adapts trauma-focused cognitive therapy successfully developed for post-traumatic stress disorder (PTSD) to reduce the intensity of affective responses to trauma-related memories. Crucially, MFCT uses a range of techniques to restructure how patients think about, perceive and relate to the stimulant and its sensory associations. That is to say, to help them grasp the situational dependence of their habits and avoid or reframe those situations that evoke the habitual craving for the substance. In an important way, patients become necessary colluders in the process of discovering and reducing the social and affective cues that maintain the compulsive use of drugs long after the initial goal has dissipated.

The memory turn, then, has arrived at the clinic, and may prove efficacious. Reductions in unnecessary human misery are obviously desirable. But medical treatments are only one possible future for drug memory science. Despite the current direction of addiction neuroscience, medical interventions are not the natural goal of addiction science nor its ultimate measure of value. If most neuroscientists believe otherwise it is in large part because the historical alliance of addiction science and medicine dates to their birth [Berridge, 2013; Campbell, 2007; Courtwright, 1982]. In the next section, we suggest another future for drug memory science. This would entail a critical friendship among biological and social research able to interrogate how social suffering expressed in stubbornly high numbers of opioid overdose fatalities become embodied in particular lives and communities.

Neurosocial collaborations
Memory theory, we have been arguing, reflects a significant development on more reductive neurobiological theories of addiction. We have used the dopamine hypothesis – often called ‘NIDA paradigm’ in the social science literature – as our example of such a theory, because of the influence it has had within the neurosciences and in shaping perceptions of addiction neuroscience within the critical social sciences. Memory models are superior in that they seek to understand neurobiological processes in relation to behavioural, psychological and environmental factors. This more integrated approach follows from the memory concept, whose primary conceptual innovation is to situate the body in the lived environment. This has created a conceptual space in which to think about addiction as an emergent process that does not privilege neurobiological determinants. It enables one to think of the process of addiction as consisting of reciprocal relations of brain mechanisms, behaviours and life. These neurosocial relations are recorded in the drug memories and memory systems that drug memory science models; they are strengthened through habitual repetition and become increasingly difficult to disrupt over time.
The emphasis of these models on long-term memory not short-term euphoria, the inclusion of diverse molecular targets like memory-related glutamate, and the incorporation of cortisotriatral systems are indicators of an important shift in addiction neuroscience towards more integrated neurobehavioral models concerned with chronic, relapsing addiction. For this reason, drug memory science is better suited than dopamine theory to a translational focus on substance addiction [i.e. severe cases of substance-use disorder] that continue to resist medicalised and nonmedicalized treatments. Further, as we have noted, memory theory is also more attuned than dopamine theory to the lived experience of addiction. We highlighted how pleasure and pain coexist in drug memories, forming powerful alliances that are mutually reinforcing, and in ways that contradict the idea that pleasure and pain are two distinct sources of motivations. Disentangling wanting to ‘feel good’ from the desire ‘not to feel bad’ under conditions of addiction is complicated at best.

These analytical advantages establish the translational potential of memory models noted above. But the destination we have in mind is not simply a turn to memory. For us, what is important is a movement in theory beyond the conceptual and experimental divisions and limits of both biological and social sciences. Treatment resistance is a symptom of a fundamental weakness with addiction science in its current form, based in its bifurcation into biological and sociological forms of inquiry. Such divisions of knowledge inevitably produce concepts and models that are at odds with human life. Because, of course, human life cannot be divided according to the disciplinary and conceptual binary of ‘biology’ and ‘society’.

As the structural distribution of despair evidenced in drug overdose deaths shows, the process of addiction is far more complex than any neuroscientific concept or model could explain or explore experimentally. This is not something that can be overcome through the concept of environment in existing neurobehavioral science. At least in part, this is because the concept of environment that functions in contemporary neuroscience is not intended to include the structural violence of socioeconomic inequality, xenophobia and stigma of lived experience that we argue express themselves in addiction, relapse and the horrors of drug overdose deaths in its frame of reference. Of course, no single field of inquiry can reasonably be expected to encompass all dimensions of such complex phenomena. What is needed are new concepts, models and collaborations that can take the neurobiological sciences of addiction beyond their analytical limits.

For its conceptual and experimental advances on neurobiological theories of addiction, contemporary drug memory science remains within the domain of neuroscience and is therefore subject to the same kind of limitations and problems. For while the memory turn is a significant shift in knowledge in addiction neuroscience it is not an “epistemological break” [Bachelard, 1986 [1938]] – a fundamental rupture in its conceptual edifice. The memory concept as it functions in drug memory science remains a disease concept, even if is one less stigmatising than the pleasure concept of dopamine theory for reasons described above. Addiction remains a neuropsychiatric disorder and this conceptual inheritance has human consequences. It means human lives defined by such a concept also inherit the burden of grief and distress that come with being the bearer of not only a disease, but one that is morally charged. And yet, drug memory science contains the seeds of a more fundamental shift in perception and knowledge worthy of terms like ‘epistemological break’ and ‘paradigm shift’.

If the memory concept creates a conceptual opening beyond neurobiological descriptions of addiction, a neurosocial approach can take us experimentally beyond addiction neuroscience itself. The ‘liminal’ ontology of neurosocial relations implies that they cannot be owned by any single field of knowledge. Neurosocial relations, qua relations, are the reciprocal interconnectivity of brain, behaviour and life that constitute the process of addiction. The epistemic authority of addiction neuroscience in the United States and Britain far outstrips its actual capacity to know addiction. Thus the potential of the turn to memory cannot be fulfilled under conditions, in which the concept of addiction remains organised according to the pathological refrain of disease, disorder, and death.

Whatever its current limits, we argue that the memory turn is an opportunity to imagine another kind of addiction science, one that starts with a critical friendship of social and biological science. A friendship sustained by an ethos of humility and self-reflection, and which aims like the Neuroscience and Society Network highlighted earlier to cultivate interdisciplinary collaborations. What these neurosocial collaborations will look like in practice cannot be determined in advance. They would have an exploratory research design grounded in lived experience. But we envisage the growth of vital models of neurosocial relations – that is to say models that are grounded in an understanding of the vital lives of human beings as they make their way in a material, interpersonal, semantic and social world, both conceptual and experimental. Such models would seek to determine how unconscious and conscious processes, mediated by memory, work in consort to establish a kind of ‘second nature’ that undermines human efforts to establish freer and less destructive relationships with drugs.

Neurosocial collaborations, which would involve not only the range of addiction neuroscientists and clinicians, along with social scientists, medical humanities scholars, policy makers and, crucially, those with lived experience of addiction, would advance positive and theoretical knowledge of addiction and relapse. They may succeed in weaning addiction neuroscience off its dependence on clinical practice, and even generate an archive of concepts and methods that allow us to analyse human relations to intoxicants in all their particularities, without perceiving them through a conceptual schema of disease. They may have translational value. But perhaps, in the end,

---

7 Our notions of the memory concept and neurosocial relations are indebted to Deleuze’s metaphysics of the virtual as set out in Difference and Repetition.
only substantial changes in social structures and priorities are necessary to do justice to the human tragedy of every single drug overdose death.

Conclusion: Vital life
Drug memory science taps into an existential truth of memory (see Heidegger, 1978 [1927]; Nietzsche, 1974[1883]) that the human experience of chronic, relapsing addiction discloses: memories hold us together, but they can also tear us apart— from ourselves, others, and life itself. Such is the paradox of memory.

Twenty-five years ago, memory was a marginal object of enquiry in addiction neuroscience. The memory concept of addiction is now central to addiction neuroscience. Eric Kandel, the most prominent neurobiologist of memory of his generation, has no doubt. In his recent book, *The Disordered Mind*, he writes “Drug addiction is a problem of long-term memory”: “The memory of pleasure” he continues, “persists long after an addicted person has stopped taking the drug” (Kandel, 2018: 201-2). As Kandel’s quote shows, addiction neuroscience is not post-pleasure. The memory turn does not represent a straightforward shift from pleasure to memory. It does, however, contradict the Utilitarian concept of pleasure-seeking inherent in dopamine theory of addiction. The human experience of addiction studied by sociologists and anthropologists such as Angela Garcia call into question Bentham’s (2015[1789]) famous dictum that pleasure and pain constitute the two “sovereign masters” of human nature and motivation. The memory turn shows rather that the pleasure of addiction coexists with pain and suffering in memories of drug-related experience. Drug memories which appear to strengthen and become more sensitive to being “triggered” by the associations that build up between drugs and cues over time.

In the realm of affliction, conceptual innovation such as memory theory of addiction is only as valuable as it is valuable to the human beings and communities concerned. Neurosocial collaborations would help to address the translation weakness in addiction science. They should be forged with the purpose of attending to the affliction of addiction— not to advance the interests or epistemic authority of any one field of knowledge. On this, Nietzsche (2008[1871]) was right— ‘objective’ knowledge means seeing with many eyes. The memory concept of addiction, we argue, makes drug memory science a good candidate for cultivating these collaborations and models. In part because the relations of brain-life and pleasure-suffering are built into the concept itself in ways social sciences and humanities can help develop. If the memory turn is to aid addiction studies in fulfilling this ambition, neurosocial collaborations must develop new conceptual and experimental models with social life at their heart. Such collaborations sustained by “critical friendship” (Rose & Abi-Rached, 2013), must proceed from an unrelenting ethical concern with the vital lives of human beings.

Data availability
All data underlying the results are available as part of the article and no additional source data are required.

References


Reference Source


Open Peer Review

Current Peer Review Status: ❌ ❌ ✅

Version 2

Reviewer Report 15 February 2021

https://doi.org/10.21956/wellcomeopenres.18205.r42387

© 2021 Bourgois P. This is an open access peer review report distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Philippe Bourgois

Center for Social Medicine and Humanities, University of California, Los Angeles, Los Angeles, CA, USA

Unfortunately the revision does not improve the argument. It weakened it by making it clearer. 1) The literature review around which the article anchors its argument is longer and more inclusive but continues to judgmentally dismiss complex literatures (harm reduction especially) and also misses relevant social literatures (religious treatment? pharmaceutical racism? decriminalization harm reduction? pre-industrial psychoactive drug use/shamanism?. 2) It fails to practice the "ethos of humility" that it calls for in a potentially innovative call for a neurosocial dialogue with laboratory-based neuroscientists. 3) They continue to insist on the failure of "NIDA" neuroscientists' contributions to opioid clinical treatments. (Ironically opioids may be the only substance for which there are some partially magic bullet big pharma solutions. Most notable useful treatments are methadone, buprenorphine, naloxone. There are also the public health clinician harm reduction outreach workers or psychiatrist/psychologist CBT providers who offer different options of diversified opioid prescriptions including heroin injection and 24 hour drop in centers. The same applies to social upstream approaches to substance use disorders such as the Marxist/Gramscian/Foucauldian/ anarchist/human rights inspired movements for the decriminalization of drug consumption and decarceration of petty retail sellers; or the literature on the disciplined service work of routine public health infrastructures of confidential mobile syringe exchanges and state-supported (or risqué semi-illegal) safe injection sites and state-supported or underground advocacy programs sometimes often headed by peer health outreach workers some of whom use substances actively or used to use chronically. 4) They overemphasize the importance of de-pathologizing the disease concept. 5) They overemphasize the “pleasure” principle (dopamine) "NIDA" focus as the likely cause of the pathologization of drug users and 6) They recognize that the "memory turn" scientists continue to erase/externalize the social causes of addiction but conclude in a semi nonsequitor to say that the “memory turn" neuroscientists like the the "pleasure block dopamine" neuroscientists are still primarily focusing on magic bullet medications which continue to externalize the social. I suspect the memory turn laboratory scientists, like the pleasure blocker lab scientists, (if those divisions hold six years from now?) are likely to care much about having a more productive humble and clinically translationally
useful interdisciplinary dialogue with humanities-leaning social scientists and the laboratory-leaning neuroscientists. Especially if their methods and literatures and scientific jargons continue to externalize the social whether they focus on memory or pleasure in their understanding of what drives addiction.

They dismiss in less than a sentence the entire sprawling field of “harm reduction” (which is a very diverse and contradictory multidisciplinary field although dominated by often psychological/individualist reductionist public health researchers based in psychology or epidemiology or health services. Harm reduction has had multiple generations of interpretations and debates).

De-pathologizing and de-stigmatizing the approach of researchers may be a good commonsensical humane principle but I suspect it is largely irrelevant to altering the translational potential of laboratory scientists to conceptualize it to be their discipline's duty or capacity to address the fuller complexity of the social structural violence of addiction or its existential suffering/ecstatic dimensions.

Their mention of pain and pleasure being two sides of the same coin is insightful (S&M exists because of that!) but I don't see how pain's relationship to pleasure is particularly more useful for depathologizing disease approaches in a society that pathologizes non-normativity, poverty, dependence and unemployment. We all need to be much much more humble about whatever the word "addiction" means. Theory is never right or wrong - it is just a useful or counterproductive abstraction. We have certainly always failed to wrap our head around solutions for chronic unemployment, suffering, rising social inequality and racism, which I think they are at the heart of chronic substance use disorders and overdoses today in the US and are the substance of the populist term "diseases of despair".

The memory turn section is more elaborated but it dead-ends into the inconsistency of an argument that recognizes that the memory turn laboratory sciences continues to externalize the social. I'm not sure if laboratory scientists need to care or know that addiction is stigmatized and pathologized socially. A depathologization/de-stigmatization of neuroscience theories of addiction is probably as irrelevant as it is unlikely to lead (if it were to occur) to a translational clinical awareness of the need for expensive public care of unmet social, physical, mental/emotional, and economic needs (including subsidized housing and universal access to healthcare). Instead it will produce more expensive big pharma profitable "me-too" drugs for insured whites in the US with complex chronic conditions.

**Competing Interests:** No competing interests were disclosed.

**Reviewer Expertise:** Anthropology/social medicine/social inequality-racism/substance use disorders/violence/homelessness/incarceration/serious mental illness/political economy.

I confirm that I have read this submission and believe that I have an appropriate level of expertise to state that I do not consider it to be of an acceptable scientific standard, for reasons outlined above.
The concept of addiction has never been a single universally accepted one, and various complementary or competing concepts – moral, medical, psychological, social and others - have coexisted over past decades or even centuries. From the mid-twentieth century to the 1970s, the various definitions adopted by the Expert Committee of the WHO were essentially pharmacological in nature, attributing addiction primarily to the actions of the drugs themselves and to individual differences in susceptibility to those actions (Jaffe, 1980; World Health Organization, 1974)\(^1,2\). In this generally well written and extensively documented conceptual analysis, McLean and Rose raise some important questions about the recent evolution of the concept.

As they point out, the pharmacological concept of addiction was made more specific from the 1970s onward by the neurobiological focus on addiction as pleasure-seeking based on the effects of drugs on a postulated dopaminergic reward system (Wise, 1980)\(^3\) and its offspring, the brain
disease concept of addiction (Leshner, 1997)\textsuperscript{4}. They criticise this focus on the grounds that (i) it is not in accord with the lived experience that distress rather than pleasure predominates in addicted drug use, (ii) the current epidemic of drug overdose deaths can not be attributed to overprescribing alone because suicidal thinking and attempts are common in overdose cases, (iii) it leads to differentiation between “good” addicts who are the victims of overprescribing and “bad” addicts who seek pleasure from illicit drugs, and (iv) the dopamine hypothesis has not led to any effective new therapy for addiction. They then argue that a much better explanation of addiction is provided by the theory evolved over the last twenty-five years that it is a functional disorder of long-term memory. This model, they believe, is already providing new and more effective approaches to treatment of addiction.

The first two of these points are valuable corrections to commonly held but perhaps erroneous beliefs. The authors present abundant evidence to support them, and readers must give them serious consideration. However, the evidence provided does not make clear whether suicidal thinking preceded the start of addicted use, or was a consequence of it. The authors might wish to clarify that question, since it affects the strength of their argument. The third point is less convincing, because the differentiation between those who actively seek and use illicit drugs and those who use drugs medically prescribed for them rests on a variety of social and psychological criteria, stereotypes and attitudes, and can not be attributed exclusively to the dopamine theory or other neurobiological models. The fourth point is perhaps debatable. While some reviewers conclude that drug therapy has been of little value in the treatment of addictions (Kalant, 2015)\textsuperscript{5}, others consider it of significant clinical value (Pierce \textit{et al.}, 2012; Schacht \textit{et al.}, 2017)\textsuperscript{6,7}. The present authors should consider making their comments on this point less absolute.

The remainder of the paper, dealing with the importance of memory in the development of addiction, is probably the most important part. This is an older concept than the authors imply: the existence of links between learning and memory on the one hand and various elements of the clinical state of addiction on the other, had already been demonstrated well before the 1990s (Finkelberg \textit{et al.}, 1978; Kalant \textit{et al.}, 1971; Lê \textit{et al.}, 1982; van Wimersma-Greidanus \textit{et al.}, 1975)\textsuperscript{8-11}. Nevertheless, there is great value in this part of the paper, with its detailed and lucid explanation of the role of memory in the generation of addiction, and the extensive documentation of the more recent supporting evidence. Since this is a conceptual review rather than an experimental paper, I do not know how to answer the questions about study design, methods and analysis. Assuming that analysis refers to conceptual analysis rather than to analytical techniques, my answers to question 2 would be “Partly” and to question 3 "Yes". A similar recent review that examines addiction as a maladaptive form of learning draws on much of the same source literature (Lewis, 2018)\textsuperscript{12} and should be cited.

The main limitation of this portion of the present paper is the tendency to view the memory theory of addiction and the neurobiologically based concepts as mutually exclusive, rather than to consider the possibility of complementary causal roles of both factors. As expressed elsewhere (Kalant, 2010, 2015)\textsuperscript{13,15}, in most cases the neurobiological studies identify the mechanisms of brain responses while the behavioral and environmental studies explore the causes that call the mechanisms into play and direct them towards specific targets. However, the prolonged intake of large amounts of drugs of various kinds does lead to functional and even structural changes in the brain that can affect the severity of addiction and its susceptibility to treatment and the risk of relapse. For example, a recent review describes evidence that heavy consumption of ethanol, nicotine and other drugs of addiction can give rise to neuroinflammation and oxidative stress in
the brain, which induce increased consumption, thus setting up a self-perpetuating cycle (Berrios-Cárccamo et al., 2020)\textsuperscript{14} that can be stopped by administration of anti-inflammatory and antioxidant medications (Israel et al., 2019)\textsuperscript{15}. McLean and Rose would improve the value of their review substantially if they could discuss the issue of neurobiological and behavioral mechanisms as interacting determinants of addiction rather than as conflicting and mutually exclusive explanations.

A few minor corrections should be made if the paper is revised:
- Page 7, right hand column, paragraph 3, line 6 – MCFT should be MFCT.
- Bottom paragraph, line 3 - “this turn” should be “this turns”.
- Page 8, left hand column, paragraph 2, line 2 - “can be explained” should be “can not be explained”.
- Page 8, paragraph 4, line 8 - “war of drugs” should be “war on drugs”.
- Final paragraph, line 1 - “collaborations between life and human sciences” is inappropriate, because life sciences can also be human sciences. Perhaps better to say “between biological and social sciences”?

References
5. Kalant H: Neurobiological research on addiction: What value has it added to the concept?. The International Journal of Alcohol and Drug Research. 2015; 4 (1). Publisher Full Text

**Is the work clearly and accurately presented and does it cite the current literature?**
Yes

**Is the study design appropriate and is the work technically sound?**
Partly

**Are sufficient details of methods and analysis provided to allow replication by others?**
Yes

**If applicable, is the statistical analysis and its interpretation appropriate?**
Not applicable

**Are all the source data underlying the results available to ensure full reproducibility?**
No source data required

**Are the conclusions drawn adequately supported by the results?**
Partly

**Competing Interests:** No competing interests were disclosed.

**Reviewer Expertise:** Behavioral Pharmacology. Biological and Behavioral studies of addiction.

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however I have significant reservations, as outlined above.

Reviewer Report 06 October 2020

https://doi.org/10.21956/wellcomeopenres.17871.r40450

© 2020 Bourgois P. This is an open access peer review report distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.
Philippe Bourgois
Center for Social Medicine and Humanities, University of California, Los Angeles, Los Angeles, CA, USA

This article identifies the shift in neuroscience from a reductionist dopamine model to memory pathways as a potentially paradigm-breaking moment that opens possibilities for a new interdisciplinary dialogue between neuroscientists and social scientists to make a call for the translational utility of what it calls “a neurosocial approach inspired by social medicine”. It tantalizingly holds up the laudable principle of a “…a new relation of ‘critical friendship, between neuroscientists and social scientists of addiction.” The prominence of the senior author and the Wellcome Institute venue and his thesis on the discursive power of neuroscience all through society but especially in disease would be a useful one for drawing greater attention to the potential of this dialogue. The emerging neurosocial literatures (not all of which would use that term but all claim to represent “the social” is simply ignored - when it is not ridiculed - by most US social scientists (who ignore laboratory science literatures), front-line harm reductionists, and epidemiologists working on substance use disorders and/or addiction/treatment/outreach/social services.

The “crisis” title and first half of the abstract is something of a red herring. As is the valence of the term “pursuit of pleasure” and the dismissive critiques of existing attempts at translational to effective treatments whether in mainstream medicine or harm reduction services. Those oversimplifications and inordinate valences to discursive power distract from the valuable important (utopian?) primary contribution of the article concerning the potential for the neurosocial dialogue:

1) Crisis “hook”

The article does not provide an argument that there is no “crisis,” as commonly understood, namely a distinct rise in the numbers of opioid death in a defined span of time. On page 4, indeed the authors detail the rapid increase in opioid deaths that begs for response. They are right to emphasize that the increase in mortality is not well described as a simple technical “crisis of over prescription”. There is however a question with respect to: a) what kind of crisis is this, and b) whether or not “crisis language” is counterproductive and distracts from the root causes of the social/existential roots of the problem. Critiques of “the crisis model” (Roitman et al.) argue that “crisis” misconstrues long-standing and socially structurally embedded conditions that are not well described in the sudden language of “emergency” that produces short term palliative measures that are ultimately ineffective. (There is also the more old fashioned longstanding sociology of substance use disorder literature critiquing the stigmatizing/criminalizing effects of “moral panics”.)

2) Overprescription and the novelty of “structural distribution of despair”

The hegemonic model of addiction in the social sciences has already been for some decades a model of the “structural distribution of social despair” long before Case and Deaton’s valuable initial demographic documentation of the drop in life expectancy among least “educated” whites in 2015 (that appears to have been what prompted Deaton’s Nobel prize that same year). Few if any social scientists studying the lived experience of substance use disorders would ascribe the disaster of opioid deaths to something as naïvely technical and reductionist as “over prescription.”
The critical ethnographic literature goes back at least to the 1930s in the US.

3) Do neuroscience models have significant impact on stigma and translational science vs old fashioned racism, class power and big pharma profiteering driving translational applications (in the US dominated globalized/medicalized and even the spiritual-religious field) if (in)effective “evidence based” treatments.

I could not help wondering why shifting from addiction being a “brain disease” or “disease of the will” to a model emphasizing “damages” within “the fundamental neurobiological mechanisms involved in long-term memory” would be likely to be less likely to be translationally interpreted into “effective treatments” that are less stigmatizing in racist and/or xenophobic, in highly unequal societies where rapidly rising social inequality has been perversely normalized in popular consciousness (especially the US) as “good for the economy” and the only way to motivate human autonomy and self-realization?

4) Impact of “brain disease pursuit of pleasure” model in the social sciences.

The NIDA paradigm is not the paradigm of most ethnographers and/or critical theory social scientists of addiction. Most have never viewed or have critiqued reductionist interpretations of “addiction is a disease/brain disease” model with a narrow definition of pleasure (or don’t even know what that refers to). Many if not most social scientists (who are not psychologists/health services behavioral scientists since at least the 1930s) have treated it as a social and political crisis. Focusing on “pleasure” as a strawman for memory also seems misplaced a false binary. Memory of pleasure is part of the memory prompt which few would reduce to a simple positive bounded definition of pleasure but commonsensically/psychologically/biologically socially might involve relief from pain or flight from suffering distress and traumatic memories. Incidentally the increasingly translational literature on “structural vulnerability” that is going more mainstream into left-leaning reinventions of medical education and clinical practice and addressing so-called “social determinants of health” can be of some use here as it applies to “addiction”.

5) Translational failure of “Effective Treatments”

The trouble with the translational field of treatment (both the neuroscientific and the social quantitative and the social epidemiological) is that in medicalized and non-medicalized addiction treatment fields nothing and almost everything counts as evidence-based treatment from methadone/buprenorphine/naloxone to faith and conversion to CBT psychological therapy to narcotics anonymous to harm reduction. This is because, as the authors argue, addiction is a profoundly social phenomenon of personal existential suffering and population-level despair that is driven by forces and institutional structures. Consequently dismissing the NIDA brain disease pleasure or memory models for that matter which arguably produced or shaped or coincided with/influenced/legitimized biomedical/pharma driven methadone/buprenorphine naloxone treatments and even free spiritual narcotics anonymous treatments. Furthermore the clinicians and medicalized harm reductionists the neurosocial would want to reach out to refer to those as “evidence-based treatments”. Most importantly dismissing them is not consistent with the humility of anticipating that the neurosocial to advance translational effective treatments is to show how resistant to treatment addiction has been. I agree with their take on it overall but it is problematic in their presentation of it.
In conclusion, and sorry for the length of this review. I enjoyed reading the article and I appreciate the authors’ optimism over the potential of a neurosocial dialogue and am intrigued and hopeful that the memory turn could possible open up a greater possibility for critically respectful dialogue across the current epistemological/interdisciplinary/disciplinary vacuums that lie between laboratory sciences and in vivo (one currently exists with more bio-psychologically reductionist and narrowly medicalized social sciences that has not been effective). I am not as optimistic as the authors and would like a bit more clarity/detail on how the memory turn will not encounter the same structural reductionist failures of the dopamine/pleasure model and ultimately (in the US at least) are likely to continue to be trumped (ugh) by moralizing populist conceptions of sin/worthiness/racism/carceral punishment/profiteering and genetic/biological reductionism.

**Is the work clearly and accurately presented and does it cite the current literature?**
Yes

**Is the study design appropriate and is the work technically sound?**
Yes

**Are sufficient details of methods and analysis provided to allow replication by others?**
Yes

**If applicable, is the statistical analysis and its interpretation appropriate?**
Yes

**Are all the source data underlying the results available to ensure full reproducibility?**
Yes

**Are the conclusions drawn adequately supported by the results?**
Partly

**Competing Interests:** No competing interests were disclosed.

**Reviewer Expertise:** Anthropology/social medicine/social inequality-racism/substance use disorders/violence/homelessness/incarceration/serious mental illness/political economy.

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however I have significant reservations, as outlined above.

Reviewer Report 30 September 2020
https://doi.org/10.21956/wellcomeopenres.17871.r40446

© 2020 Field M. This is an open access peer review report distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.
I am sympathetic to several of the themes that run through this paper, particularly the emphasis on “deaths of despair” as determinants of the opioid crisis, and the call for collaboration between neuroscientists and social scientists. I also appreciated the attempt to raise awareness among social scientists of the limitations of relatively simplistic neuroscientific accounts of addiction, and of the limited innovations in addiction treatment that have arisen from dopamine theory.

However, I felt that the linkage between different arguments was unclear or overly speculative (or both), and in particular I was not persuaded by the authors’ argument that recent work on the ‘memory turn’ is likely to yield a solution to the opioid crisis. I elaborate on these arguments below.

1. The authors provide a brief accessible summary of the “rise and fall of the dopamine hypothesis” (p5). However, contemporary addiction neuroscientists are fully aware that the neurobiology of addiction is complicated, and goes well beyond dopaminergic adaptations. Therefore any attack on the dopamine hypothesis is a bit of a straw man, assuming that the intended purpose was to criticise neuroscience accounts of addiction more broadly. A recent and concise overview of the neurobiology of addiction is provided by Volkow et al. (2016), and the current paper might be strengthened by broadening the coverage of neuroscientific accounts of addiction. The paper would also be strengthened by referring to other reviews which demonstrate that neuroscience research more broadly has failed to deliver novel effective treatments for addiction, and this might be partly because of its inability to model the complexity of addiction, including the importance of social factors (Field & Kersbergen, 2020; Heilig et al., 2016; Borsboom et al. 2020), the latter includes a relevant commentary by Field et al. (p18) that discusses the opioid crisis and “deaths of despair” in the context of reductionist neuroscience approaches to addiction).

2. Alternatively, perhaps the focus on the dopamine hypothesis was because it has “proven highly influential in shaping perceptions of addiction neuroscience in the social sciences (Courtwright, 2019). It is how, with few exceptions, social scientists conceive of addiction neuroscience” (p6).

If so, the purpose of the criticism of the dopamine hypothesis might be signposted more clearly, along the lines of “the neuroscience has moved on, and it is important that social scientists catch up”.

3. The line of argument in the paper then moves on to argue for the importance of “the memory turn” as a potentially important way to conceive of addiction and to bridge neuroscience with social science. I broadly agree. However, the evidence presented in favour of the notion that a focus on memory may yield new treatments for addiction is selective and rose-tinted. The recent work by Marsden and colleagues on MCFT is preliminary and has not yet been subjected to a definitive clinical trial, therefore the findings from small pilot studies should be interpreted with caution. The authors should also consider other translational addiction interventions that aimed to disrupt drug-related memories, and generally did not translate well “from bench to bedside” (e.g. Das & Kamboj)5. More recent memory-focussed interventions have also yielded promising
findings (e.g. Das et al., 2019). These findings might be discussed. However in the absence of well-powered trials, the clinical promise of interventions translated from “the memory turn” should not be overstated.

4. In my opinion the most significant limitation of the paper is its central argument, summarized on p8:

“the memory turn is thus much better suited to interrogating the more destructive cases of substance addiction that represent the opioid crisis. It is more attuned to the lived experience of human addiction as shown in social and clinical studies referred to above. It taps into the complex nature of motivation in which pleasure and pain coexist in drug memory. It is not simply that the person wishes to ‘feel good’, or to be ‘free of pain’, it is that pleasure and pain become entwined in reinforcing ways. And it does so, by situating the neural, cognitive and affective systems involved within a neuropsychological schema in which the social’ plays a decisive role”.

Where is the supportive evidence for this grand claim? I think that “the social and clinical studies referred to above”, refers to the literature discussed on p5-6 which demonstrate that opioids function as negative reinforcers (they relieve physical and emotional pain in addition to producing euphoria). The notion that negative reinforcement underpins addiction has a long history and indeed the underpinning neuroscience is well described in work by Koob and colleagues, as cited here. It is also plausible that negative reinforcement models of addiction can account for “deaths of despair”, including the opioid crisis. But why does this suggest that memory-based treatment interventions are likely to provide the answer? This link was not clear to me, and the paper would be greatly improved if the authors could spell out the relevant evidence and its implications more clearly. In addition, alternative explanations of the link between social deprivation and drug addiction should be considered, including relative absence of alternative (substance free) reinforcement (e.g. Leventhal et al., 2015; Acuff et al., 2019).

5. The authors made other claims that were not clearly supported by the evidence, for example on p8:

“These relations represent a challenge for addiction research: the need to develop conceptual and experimental models of the ways that unconscious and conscious processes interact through memory to undermine human efforts to develop new relationships (cognitive and affective) with drugs”.

6. Also, I would agree with the following claim, as many would, but I did not understand how this claim is related to the “memory turn”, or how it follows from the arguments presented (p8):

“...harm reduction activism (see Campbell, 2020) that is led by service users, but supported by neurosocial collaborations, would be a positive move forward.”

7. The authors claim that “The compulsive, chronic, relapsing hallmarks of substance addiction can be explained through neurochemical maladaptations alone”. Although this is certainly the conventional view among addiction neuroscientists, the notion of addiction as a chronically relapsing disorder of compulsion has been challenged in recent years (e.g. Heather, 2017), and this debate should be mentioned here.

8. Throughout the paper and particularly on page 7, the authors place technical terms from the
neuroscience and memory literatures in quotation marks (e.g. “corticostriatal systems”, “declarative”). I'm not sure what the quotation marks are supposed to achieve here – it would be better to define the terms in a way that would be useful for the intended audience of this paper.

9. On page 7 authors state that “powerful memories” are part of one of the diagnostic criteria in DSM 5. I'm not sure that this is correct. The authors include a footnote in relation to this claim but I could not see the footnote anywhere.

References

Is the work clearly and accurately presented and does it cite the current literature? Partly

Is the study design appropriate and is the work technically sound? Partly

Are sufficient details of methods and analysis provided to allow replication by others? No

If applicable, is the statistical analysis and its interpretation appropriate? Not applicable

Are all the source data underlying the results available to ensure full reproducibility? No source data required

Are the conclusions drawn adequately supported by the results?
No

**Competing Interests:** No competing interests were disclosed.

**Reviewer Expertise:** Psychology, addiction, behaviour change.

I confirm that I have read this submission and believe that I have an appropriate level of expertise to state that I do not consider it to be of an acceptable scientific standard, for reasons outlined above.