

Getting Out of Your Head: Addiction and the Motive of Self-Escape

Daniel Morgan and Lucy O'Brien
Mind and Language 31 (3) (2016): 314-334.

Introduction¹

It is an interesting fact about philosophical work aimed at understanding the nature of addiction that very little space is given to trying to capture, and give significance to, the changes to our consciousness that addictive substances tend to bring about. Part of the aim of this paper is to constitute a correction in that direction. It is an important and obvious fact about alcohol, and other drugs, that they *change our states of consciousness*: they change what it is like to be us. But they do not only change what it is like to be us by producing new experiences of one kind or another. They have in common that they are capable of effecting a kind of self-escape: an escape from tiring, uncomfortable or painful awareness of ourselves. This kind of self-escape is put in a number of ways in reports on the effects of alcohol or other drugs. We drink too much and 'get out of our heads', 'get off our heads', get 'taken out of ourselves', we no longer quite know 'who we are, and what we are doing'. The question of what it might mean to be 'taken out of ourselves', and what role the ability of an addictive substance to take us out of ourselves plays in its capacity to be addictive, will be a central concern of the paper. In particular, the suggestion we want to make in this paper is that the hypothesis that a motivation for self-escape is a systematic cause of addictive desires for *alcohol* is one to be taken seriously.

¹ This is a jointly conceived, jointly written paper, the whole of which we are both fully responsible for. We would like to thank the editors for inviting a contribution to the *Mind & Language* workshop on Addiction, Self, and Self-Knowledge, in March 2015, out of which this paper developed. We would also like to thank Hanna Pickard, Léa Salje, and a referee for *Mind and Language* for very helpful comments. Emails: danielmorgan@ub.edu, l.o'brien@ucl.ac.uk.

We cannot appeal to a person level explanation like the motive of self-escape in explaining addiction without considering how such an explanation sits with competing explanations coming from neuroscience to the effect that addiction is caused by the effect that addictive substances have on the dopamine system. But, we think, the current work on the dopamine system does not rule out *pluralism* in relation to the explanations of addiction. We can, if pluralism is right, offer differing explanations of addictive desires – where such explanations involve offering different systematic causes. We will further suggest that pluralism with respect to such explanations is plausible both in relation to the types of cause of addiction, and with respect to levels – chemical, neurological, and personal level psychological. In his work on causation in psychology John Campbell points out that investigation into the causes of depression has thrown up two significant variables: earlier humiliation and genetically acquired serotonin deficiencies. He surmises that these work together: humiliating life events along with serotonin deficiencies tend to produce depression. He concludes that we have two causal variables ‘one psychological and one biological, which are jointly causes of later depression’ (2006: 65) and that that might be the most fundamental causal explanation of depression we can give. One way to take the suggestion to be made here, is as offering the hypothesis that we have at least two causal variables, one psychological, one neurobiological, that are jointly causes of addiction.

In Section 1 we will suggest two *prima facie* conditions we take addictive desires – as opposed to other ordinary non-addictive desires – to meet. One is that they can be very strong, and because of this, can lead us to act on them even when the consequences of doing so are terrible. The second condition is that addictive desires resemble each other in the causes that underpin them.

Section 2 looks at the neurophysiological causes behind addiction and argues that it explains at most *some* of the strength of addictive desires. Section 3 elaborates an alternative, personal level cause that seems to explain some of their strength: we argue that the motive of self-escape is, often, at least part of what makes addictive desires as strong as they are. In providing the empirical evidence to back this claim up we will concentrate on the case of alcohol.

Before discussing these conditions, it is worth thinking about the background against which they are formulated. It is natural to think that the concept of addiction is something of a natural kind concept in at least the following sense. We apply the concept to certain desires, and to certain behaviours, on the basis of certain relatively easy access markers. However, the referent of the concept is assumed to have a nature that is not exhausted by these markers. It is, for example, not simply obvious what the range of things to which one can be addicted is, and one can imagine trying to settle this question by arguing that a putative example of an addictive substance, or perhaps process, resembles, in its deeper nature, some already accepted paradigm.

The natural kind model raises the question of what level, or levels, the nature of addiction is describable at. On one possible view the nature of addiction is uniquely describable at the neurophysiological level. So, the way of pursuing the question of whether desires for a certain thing could be addictive is to ask whether that thing resembles, *in its neurophysiological action*, paradigm addictive things. A different view is that addiction has a neurophysiological nature, but that it also has a person level nature. A third view in logical space is that addiction just has a person level nature. We will be making the case for a version of the second view that emphasizes the capacity of engaging in the addictive project to facilitate self-escape. Hence our title.

1 Personal Level Addictive Dysfunction

We alluded above to ‘relatively easy access to markers’ for addiction. It will be helpful to flesh this out a bit more.

It can be tempting to act on an addictive desire even if one judges that doing so is an all things considered bad option. That, however, does not distinguish addictive desires from non-addictive desires. It can be tempting to eat a cupcake even if one judges that doing so is an all things considered bad option, e.g. because it means breaking one’s diet. The most immediately striking thing about addiction-involving cases is the extremity of the cases in which addictive desire can trump judgments about what courses of action are best.

If one knows that eating the cupcake will bring disastrous consequences to oneself and one’s loved ones – one knows that eating it not just a bad but a terrible option – eating the cupcake is no longer a tempting option. By contrast, in addictive cases, even if one judges that the option of taking an addictive substance is terrible the option of taking it can remain tempting. Let us consider the following claim about addictive desires:

Personal Level Addictive Dysfunction: It can be tempting to act on an additive desire even if one judges that acting on that desire is a terrible option.

Addictive behaviour has often been regarded as compelled or unfree. This has recently been a less popular theoretical option. But something about addictive behaviour must explain why it is such as to be easily mistaken for unfree behaviour. The *Personal Level Addictive Dysfunction* claim provides one such explanation. If one is looking on at addictive behaviour from the outside, one may reason that it *must* be unfree since no one could freely choose such terrible options. Jumping to this conclusion is unjustified,

since it neglects quite how strong addictive desires can be. But neglecting the strength of addictive desires is natural if one is only introspectively familiar with less strong desires – which fortunately most people are. (Of course, a different explanation will be required of why some addicts regard their own behaviour as unfree).

Personal Level Addictive Dysfunction does really seem to point to a genuine dysfunction. The evolutionary point of desire, as of mental states generally, is to get people to pursue options that are good for them. Addictive desires often get people to pursue options that they judge to be bad for them, and that indeed are bad for them. That qualifies them as dysfunctional. (See Levy, 2013). *Personal Level Addictive Dysfunction* bears comparison to a similar thesis that Richard Holton endorses, and regards as capturing something central about addictive desire: ‘Addictive desire is uncoupled from judgment about what is best, whereas non-addictive desire is not’ (2009: 109). Holton’s thesis seems false if uncoupled means has a life independently of what is judged, for reasons that Holton’s own very helpful framework make clear. Some desires are *cognitive* – in the sense that judging that pursuing them is not a good option causes them to disappear without residue. But plenty of desires, including plenty of non-addictive desires, are not cognitive in this sense. Judging that it would be a bad idea to eat the cupcake because it brings terrible consequences may not cause the desire for the cupcake to disappear without residue. The cupcake itself remains a delicious looking cupcake, and some of the force of the desire for it may remain. It is just that the desire is massively outweighed by the judgement on the opposite side of the scales. The non-addictive desire remains but we are not tempted to act on it.

The notion of an option’s being *tempting* is a notion of common sense psychology. One might try to cash the notion out in other terms. One might, for example, say that an option is tempting if, and only if, refraining from taking it takes will power.

But this does not seem right. Refraining from taking the cupcake might take some small degree of will power. It is just that, since the consequences of taking it are so terrible, and the desire for the cupcake is not so strong, most people will be certain to exert the required will power. There are two platitudes about temptation that do not come close to an analysis, but are enough for our purposes. The first is that the worse one judges an option to be the less tempting it is. The second is that the more one desires an option the more tempting it is – perhaps because what confers on something the status of being a temptation is the fact one desires it. So options are tempting in two cases – either one does not judge the option to be very bad; or one does but one also desires the option very much. *Personal Level Addictive Dysfunction* says that acting on addictive desires can be tempting even if one judges that doing so is a very bad option. The shallow explanation of how this is possible is that addictive desires can be very strong desires. Deeper explanations will target the question of *why* addictive desires are so strong. Indirectly, an explanation of why addictive desires are so strong is an explanation of *Personal Level Addictive Dysfunction*. That is why the question of why addictive desires are so strong is our focus.

Nothing in the previous paragraph implies any simple connection between a desire's status as addictive and its degree of strength. Being a *strong* desire is not a sufficient condition for being an addictive desire, even in conjunction with the condition that the desire is a desire for something for which one can have addictive desires. Someone might offer you a million pounds to consume an addictive substance and you might as a result form an exceptionally strong desire to do that. But that desire would not be addictive. Plausibly, that is because it is not caused in the same way as paradigm addictive desires. Neither is being a *strong* desire a necessary condition for being an addictive desire. It is perfectly possible for someone to be a recovered addict for whom

there is a barely murmur of temptation in their daily life, even if they regularly encounter addictive cues. Whatever temptation there is, is nevertheless addictive in character. Plausibly, that is because it is caused in the same way as paradigm addictive desires. This leads us to our second claim about addictive desires:

Systematic Cause Claim: A desire is an addictive desire if and only if it is systematically sustained by the same kind of causes that are involved in paradigm addictive desires.

That, crucially, leaves it open what the relevant causes are. It leaves it open, in particular, whether they are solely neurophysiological.

1. Why are Addictive Desires so Strong? The Neurophysiology of Addiction.

Many addictive substances (e.g. amphetamine, nicotine, cannabis, cocaine, alcohol) all either stimulate dopamine release or decrease dopamine re-uptake. That would be quite a coincidence if their having this chemical profile were not at least part of the explanation of why they are so strong. Quite understandably, this thought has been the stimulus to much work on the neurophysiology of addiction.

We want to argue for two independent points about this work. The first point (2.1) concerns what it does say, in particular what it says about the interpretation of the dopamine system. On the two most prominent interpretative frameworks, addictive substances are seen as causing a certain kind of state – either a desire, or a learning state that represents the addictive substance as having a certain reward-value – that is *disproportionate* to the amount of pleasure the addictive substance produces. This disproportion thesis is intended to capture a sense in which addictive substances cause dysfunction in the dopamine system. We argue that the empirical evidence offered in

favour of this thesis, which chiefly consists in experiments involving rats, is inconclusive. The possibility that addictive substances cause strong desires/representation of high reward-value simply by causing a lot of pleasure has not been ruled out. The second point (2.2) concerns what the work on the neurophysiology of addiction does *not* say. In particular, even if it is agreed that states of the dopamine system are a kind of desire, there are reasons to think that these desires cannot be identical to the desires that are the proximate causes of addictive behaviour. So there is a need for a story about how the two kinds of state relate. We argue that the most plausible story is that states of the dopamine system are *inputs* to the desires that are the proximate causes of addictive behaviour. Accepting this view naturally leads one to ask whether there might be a different kind of input, one that has little to do with the dopamine system. This clears the way for us to argue, in Section 3, that the motive of self-escape is one of the causes of addictive desire.

Inspired by Robinson and Berridge's work on the dopamine system (see, for example, 1993), Holton (2009), Holton and Berridge (2013) and Dill and Holton (2014) defend a view about addiction that has two key features. First, states of the dopamine system are states of wanting/desiring (albeit not conscious states of wanting/desiring). Second, the peculiarity of a way in which addictive substances impact the dopamine system, and the state of wanting/desiring that the system can be in, is what accounts for the strength of addictive desires.

What is it then that makes it so hard to avoid succumbing to addictive desires?

To understand this we will need to know more about the mechanism of addiction. Work by Robinson and Berridge indicates that addiction involves, very roughly, a decoupling of wanting and liking. Standardly, once we like something, or believe we will like it, we want it; and conversely, once we do not like it, or

believe we will not, we do not want it. Addicts are different. They need not like the substance to which they are addicted: they need take no pleasure in getting them, not in the prospect of getting them. (Holton, 2009:104)

The general picture of the dopamine system underpinning this account is Robinson and Berridge's 'incentive salience' model. The incentive salience system involves desires for things that, in general, contribute to the health of the organism. Its succeeding in doing this depends on two causal connections. There is a connection between how likely a substance is to contribute to health, and how much pleasure it gives the subject. There is a connection between how much pleasure the substance has given in the past and how much the subjects wants, or desires, that substance. These desires are states of the dopamine system. They are intrinsic desires in that they are desires for particular substances rather than the upshot of a general desire for pleasure, and beliefs that certain substances will give pleasure. They are 'salience' desires in that they are highly cue driven. Addictive substances 'hijack' the incentive salience system, causing it to malfunction, in that they disrupt the second correlation, causing desires for their consumption that are out of proportion to the amount of pleasure they produce. They do this by impacting on the desires *directly*. That is, the impact is purely via the substance's chemical properties, rather than by the causation of a state of pleasure which in turn leads to an incentive salience desire of a normal, proportionate degree of strength.

2.1 On the Evidence for the Side-lining of Pleasure

Many readers of this paper will consume alcohol. They will probably believe that the alcohol they consume contributes to states of pleasure they experience. It is unlikely that familiarity with what the incentive salience account says about addictive

substances will lead them to regard their desires for alcohol as basically misguided states, ones that they ought to rid themselves of if only they could, even if their sole concern were to maximize their own pleasure (call this the error-theory of drinking). Of course, not everyone who drinks alcohol is an alcoholic. But the basic mechanism the incentive salience account invokes applies to addictive substances as such, not to addictive substances as these impact on a particular subset of the population. (This is why the rats under examination in the experiments that are intended to provide the crucial evidence for Berridge's account are not picked for being particularly addiction-susceptible rats).

Is any of this a problem for the incentive salience account? It might be said that it is not on the grounds that the account is not concerned to deny that addictive substances cause pleasure. Rather, the claim is that, even if they do, it is not thereby, or not solely thereby, that they cause desires for their own consumption. Addictive substances also cause such desires directly and chemically by impacting on the dopamine system.

This contrast – between causing desires indirectly by causing states of pleasure vs causing them directly and chemically – only holds however if it is accepted that the dopamine system involves states of desire, but not states of pleasure. (If it involves both kinds of states, then 'causing chemically' does not entail 'causing directly rather than via causing a state of pleasure'). That is something that Berridge does accept. He is led to do so chiefly by experiments with rats, in which the rats are interpreted as wanting addictive substances they do not particularly like. We need now to say why we are not convinced by his interpretation of those experiments.

One might be interested in addiction in rats for its own sake. But one might also be interested in addiction in rats because, although one's ultimate interest is addiction

in general or addiction in humans, some feature of rats makes them an especially easy to deal with object of the study. Apart from the fact that it is legal to lock rats up in specially designed cages and administer addictive substances to them for experimental purposes, the crucial cited feature of rats is that there are independent measures of their *wanting* something (their pursuing it, e.g. by pressing a lever that releases it) and their *liking* it (exhibiting certain facial expressions when they consume it). It is these measures that are relied on in arguing that rats wanting for addictive substances can be disproportionate to their liking for them. And the crucial experiments are in fact supposed to involve wanting and liking coming apart – not just their both being present to similar degrees, but the latter playing no causal role in producing the former.

Two questions at this point. (i) Isn't it fairly plausible that, *with humans*, we have independent measures of wanting and liking. Wanting can be measured by tendencies to pursue, or by introspective report. Liking can be measured by introspective report. So, shouldn't we be able to replicate the wanting-without-liking experiments in humans? Can we? Rookie users of addictive substances typically report liking them an awful lot. Admittedly, addicts who have been using a long time often say that they don't enjoy the substances they use as much as they did when they began using. But it isn't clear that this is relevant. The choice the long-term addict faces is between using *now*, say on a particular day in 2016, and not using *now*. So, the relevant comparative question about pleasure isn't about how the level of pleasure consequent on using now compares with the level of pleasure achieved by using *then*, say on a particular day in 2000 when they first tried the drug. The relevant question is about how the level of pleasure consequent on using now compares with the level of pleasure consequent on *not* using now. If that second difference is large – which it very plausibly will be, especially given how unpleasant it is to do without a substance to which one has

become addicted over an extended period – then the fact that the addict strongly desires to use now isn't that surprising. It isn't clear why the wanting should count as 'disproportionate'.

(ii) Do we really have a measure of liking/pleasure *in rats*? Aren't the facial expressions that are treated as that measure in fact a measure of pleasure *in taste* more specifically?² Since the pleasure associated with addictive substances is not usually pleasure in taste (except incidentally, as in the case of alcohol), relying on this measure risks *understating* how much rats like addictive substances and consequently *overstating* how much the degree to which they want addictive substances is disproportionate to the amount they like them. The evidence seems clearly to show that the rats like the taste of sugar – measured by facial expressions – no more when they have amphetamines in their brains than when they don't. Nevertheless, they are much more motivated to pursue the sugar – they want it much more. But that is not enough to rule out the possibility that they get more pleasure from sugar when they have amphetamines in their brain, but the pleasure is not pleasure in taste, and therefore is not captured by the facial expressions that are a measure of pleasure in taste.³

It was pointed out above that much work on addiction is supported by the thought that it would be an odd coincidence if addictive substances all impacted on the dopamine system, but that fact about them played no role in the formation of addictive desires. We find this thought entirely compelling. In exactly the same way though, it would be an odd coincidence if addictive substances all typically caused pleasure, but

² Berridge (2000: 175-6) points out that when first studied they were thought to be responses to particular tastes – e.g. saltiness – until it emerged that, e.g., a salt-deprived rat will produce the expressions for salty substances whereas a salt-sated rat will not.

³ Notice that we have not claimed that there isn't anything metaphysically or conceptually impossible about a substance that causes strong desires for its own consumption without causing intense pleasure states. We *have* claimed that the evidence that paradigm addictive substances are such substances is inconclusive.

that fact played no role in the formation of addictive desires. What is it about the neurophysiology of addictive substances that accounts for the fact that they typically cause pleasure, even if they don't always? Suppose that there is some system of the brain, call it system P, other than the dopamine system that addictive substances impact on, and that involves states of pleasure. In general, since pleasure states correlate with desires (according to common sense as well as the incentive salience account) system P must impact on the dopamine system. So how is it known that addictive substances do not impact on the dopamine system by impacting on system P? What stops the kind of interaction between system P and the dopamine system – the kind of interaction that occurs when the pleasure of cupcake causes a desire for cupcake say – from occurring when addictive substances cause pleasure?

So far, we have focused on Berridge's interpretation of the dopamine system, and argued that his view that addictive substances cause desires for their own consumption whose strength is out of proportion to the amount of pleasure that they cause is unsupported by the evidence. We think that much the same point applies to a different interpretation of the dopamine system. On the alternative account we have in mind, the dopamine system is a learning system (See Schultz et al. 1997, Tobler et al. 2003, Schultz 2006).⁴ States of the system represent things as being better or worse – or more or less pleasurable – than expected. Adherents of this view typically hold that addictive substances cause dysfunction in the learning system in that they cause states that represent the substances as being better than expected. As Levy very helpfully puts it:

⁴ The difference between these two accounts disappears if one thinks that desires themselves are states of the learning system. This view of desires is defended in Schroeder (2004). Note also that a model developed in McClure, Dew and Montague (2003) suggests that these two accounts of dopamine function are computationally reconcilable. The so called 'actor-critic' model offers a learning rule that suggests that dopamine could code both for the value of predicted rewards, and for assigning motivational attention to actions most likely to lead to reward.

In effect, the dopaminergic system responds to drugs with the signal that consumption is better than expected. It does so every time the drug is consumed. The addict cannot learn the reward value of the drug, because the system for reward value learning is dysfunctional. On every occasion the drug is consumed, the dopaminergic system reports that the drug is more rewarded than expected. The result is pathological learning; the system treats the drug as of ever increasing value. (Levy, 2013:3)

Just like Berridge's view, the most striking part of the learning-system based view about addiction is a thesis about disproportionality – the difference is that the disproportionality is between a learning state and a pleasure state rather than between a desiring/wanting state and a pleasure state. The criticism we made above about the evidence for Berridge's view applies equally to the learning-state based view, since the key point of the criticism is that we lack a measure of pleasure/liking in rats. The difference between the desiring-system interpretation of the dopamine system and the learning-system interpretation is not critical for our purposes. For convenience we will speak as though states of the dopamine system are desires, rather than representations of reward-value.

We have argued that the evidence offered for the claim that addictive states cause desires for their consumption that are disproportionate to the amount of pleasure they cause is not supported by the evidence. So, we have not accepted the part of Berridge's view about addictive substances that is (i) most distinctive or that (ii) is intended to capture the sense in which addictive substances involve *dysfunction* in the dopamine system. This turns out to be a less drastic disagreement than one might expect. Here is why. Even if one is unconvinced by the disproportion thesis, one can still accept Berridge's claim that dopamine spikes are a kind of desire that can be very

strong. We are interested in explaining why it is that addictive desires can be so strong as to cause one to act on them even if one judges that the consequences of doing so are terrible. That is, we are interested in explaining the personal level dysfunction that addiction involves. It is clearly a contribution to explaining this to say that addictive substances cause dopamine spikes, and that dopamine spikes are a kind of strong desire. But it doesn't make an additional contribution to add that the dopamine spikes that addictive substances cause a kind of desire that is *disproportionately* strong. The effect of adding that adjective is to make a claim about the extrinsic properties of the desire – about how the strength of the desire compares with the strength of certain pleasure states. But it is the desires' intrinsic strength that gives them the capacity to overwhelm judgment. The disproportion claim is an interesting claim about the dopamine system. But even if it is true it doesn't explain personal level addictive dysfunction. There are metaphysically possible substances that cause strong desires for their consumption by causing states of intense pleasure (whether or not the addictive substances we are familiar with are examples of such substances). And it would certainly be possible to be addicted to such substances, even though they wouldn't involve dysfunction in the dopamine system in Berridge's sense. A way of summarizing this point is to say that the dysfunction in the dopamine system that Berridge posits is not explanatory of the personal level dysfunction that addiction involves.⁵

Suppose it agreed that dopamine spikes are strong desires. That is clearly explanatorily relevant to the question of why desires for addictive substances can be so strong. The task of the next section is to say a bit more about how the explanation goes,

⁵ One could think of the function of the dopamine system less narrowly, as being to produce desires for things that are good for us, rather than to produce desires for things that are pleasurable. And then the two dysfunctions, the personal level one and that involving the dopamine system, would be less far apart.

and to argue that the explanation ought to be supplemented by factors that are not neurophysiological.

1.2 Monism and Pluralism

The desire/wanting states attributed to the dopamine system are *not* conscious. The desires that are the proximate causes of addictive-substance pursuing behaviour in humans *are* conscious – indeed addicts frequently attest that they are phenomenologically irrepressible, as conscious and subjectively inescapable as mental states get. So, by Leibniz’s law, the two kinds of state cannot be identical. But, consistent with Leibniz’s law, it could still be that the desire states of the dopamine system are the *sole inputs* to the conscious desires that are the proximate causes of human addictive substance-pursuing behaviour. If that were so, the whole explanation of the strength of the conscious desires would be neurophysiological. After all, the whole explanation of the strength of the states of the dopamine system is clearly neurophysiological. And, on the proposed view, the conscious desires have no other kind of input – the states of the dopamine system are their *sole* input.⁶

Are desire states of the dopamine system the sole inputs to conscious desires for addictive substances? The more general thesis that states of the dopamine system are the sole inputs to conscious desires *for the kinds of things that impact on the dopamine system* – addictive substances but also, e.g., food and sex but not metal gizmos or world

⁶ To belabor the point, the key word here is ‘sole’. ‘Input’ is a placeholder. It refers to a very undemanding relationship. How that relationship should ultimately be precisified – whether for example we should think in terms of the notion of constitution or realization – is a difficult question in the metaphysics of the relationship between conscious and non-conscious mental states. This general question has nothing to do with addictive desires in particular – we can equally ask, for example, about the relationship between our conscious visual experiences and the non-conscious states of our visual processing systems – and we do not commit to any particular way of answering it.

peace – does not, on reflection, seem right. Take food. One could experience a very pure kind of desire for a particular food-item the sole input to which was a state of the dopamine system. More usually though, one's desire for food will have other inputs beside. Suppose one wants a cupcake. How much one wants it will depend in part on how big the associated dopamine spike is. But other factors are very likely also relevant. For example, cognitive factors are. Is it Ramadan or not? Is the cupcake healthy or not? If the opportunity to eat now is missed, how far into the future will the next opportunity for eating be? At least dispositionally, a subject will have beliefs on such matters, and those beliefs will impact on the strength of their desire. In cognitively sophisticated creatures, even our baser desires will be cognitively adulterated. So, at least in general, it seems more plausible to think of states of the dopamine system as *one kind* of very important input into conscious desires for things that impact on the dopamine system. The strength of the resultant conscious desires will depend on other factors as well. These other factors may not be neurophysiological.

There is a way of accepting this general line of argument while denying its importance in relation to addictive desires. It can be summarized in the following equation:

$$\text{Addiction-in-humans} = \text{addiction in rats} + \text{human self-control}$$

Here is the basic thought. With addictive substances, precisely because their effects are known to be so harmful, a typical subject's cognitive desires will tend to weigh *against* using them. When humans end up with desires for addictive substances, those desires are not plausibly thought of as the sum of a push from the dopamine system *plus* a separate push from cognitive desire. They should rather be thought of as the sum of a

push from the dopamine system *minus* a push in the opposite direction from cognitive desire. In so far as the *resultant* desire for the addictive substance is strong, its strength will be derived from the dopamine system. So, one might have thought, the extra that needs to be added to a story of addiction in rats (if we assume that the dopamine story is indeed the full story of addiction in rats)⁷ in order to understand addiction in humans will *all* be on the against-using side of the ledger: it will all be about humans having self-control in a way that rats do not and therefore having a better chance to resist the extremely powerful incentive salience desires that addictive substances cause in both species.

We'll first give a general reason for thinking this must be wrong: that there must be other factors to add to the pro-using side of the equation. We'll then aim to identify and explore some specific non-neurophysiological factors.

The general point is this. One of the most puzzling facts about addictive substances

- given what we know of their capacity to affect the dopamine system in such a way that addictive substances come to be marked on each ingestion as particularly 'pursuable' – is that we are not all addicts. Perhaps some of this variation is due to variation in different individual's brains, in particular in different individuals' dopamine systems. However, the data is pretty clear that neurophysiological make-up is not the only predictor of whether one is susceptible to addiction. The result of epidemiological studies of alcoholism in identical compared to fraternal twins seems to suggest approx. 50-60% of addiction is due to genetic factors. (Prescott and Kendel, 1999). Moreover, it

⁷Arguably, the rat park experiments (see Alexander *et al* 1981) give us a reason to not assume that. The rats who hang out with other rats and feel less inclined to use drugs seem to have desires of a kind that counteract the effects of the dopaminergic motivational states for the drugs.

is also well established that personal level social factors that seem to cut across such differences in genetic make-up, such as stress and shame, *are* predictors of addictive behaviours. These are some *prima facie* considerations in favour of explanatory pluralism.

To back this reply up, we need to point to some additional causal variables that might give rise to addictive desires in humans.

A first candidate mechanism involves the *desires for pleasure*. Someone might know that taking a particular substance will cause them pleasure, and they might, for that very reason, form an (instrumental) desire to take that substance. (It is not just a datum that addictive desires are non-instrumental, as they turn out to be on the incentive salience account). This would be a *cognitive* desire since if the person came to the view that the substance would not cause them pleasure, the desire, or the portion of it that was cognitive as opposed to rooted in the incentive salience system, would disappear without residue. If this is a mechanism behind addictive desire, that explains a datum that seems very hard to explain on Berridge's incentive salience story: the fact that addictive substances typically cause pleasure. The pleasure states would be playing a causal role not via proportionate states of desire in the dopamine system – this is what Berridge says does *not* happen – but by causing subjects to form expectations of pleasure on the basis of which they form cognitive desires for addictive substances. The fact that addictive substances typically cause pleasure would thus no longer have the look of being such a surprising coincidence.

The second mechanism, and the one we think is harder to describe but likely to be more important, is better conceived as not involving the incentive salience system but *also* as not involving cognitive desires. Dill and Holton point out in passing, indeed in a footnote (2014:7), that there is no reason to take their distinction between

incentive salience desires and cognitive desires to be exhaustive. The distinction between cognitive and non-cognitive desires is clearly exhaustive. But there is no reason to think that the only kind of non-cognitive desires are incentive salience desires.⁸ Dill and Holton suggest desires ‘involved in emotional reactions such as fear or guilt’ as examples that do not fall neatly into either the category of cognitive desire or the category of incentive salience desire. Without seeking to put forward a general theory of the kind of desires involved in appetites or emotions, we want to take seriously the idea that a particular example of these *third* kind of desire might be a crucial cause behind addiction. In fact, we think that for those who are not already immersed in monistic neurophysiology-oriented discussions of addiction, the suggestion that addiction might systematically be a strategy for dealing with negative emotion, or some other kind of response to negative emotion, will seem almost prosaic.

3. Intoxication and Self-Consciousness

As we said at the start of this paper, addictive substances tend to be intoxicating. They are intoxicating in the sense that they give intense pleasure, but they are also intoxicating in the way that they take their consumers ‘out of themselves’. What do we mean by that? One very obvious and common use of alcohol is to offer subjects relief from the self-consciousness and shyness which often pervades social situations. Social self-consciousness puts us in the position of monitoring ourselves, our actions, our appearances, our social standing, and of monitoring the responses others have to us.

⁸ As our discussion of craving and desires suggested, we do not in fact think of appetitive desires – whether for water, cupcakes or gin – as plausibly construed as necessarily involving, or being only as resilient as, the beliefs that that which is desired will cause pleasure. We might explain an appetitive desire for something in terms of the pleasure it brings without assuming that the subject has beliefs about its pleasure bearing properties.

This can be a painful and exhausting business that can be ameliorated by changing our states of consciousness with the use of drugs. Anyone who drinks alcohol will recognize the way in which a glass of wine or two will make one less self-conscious. One becomes less sensitive to self-relevant information. One does less to monitor and control one's impulsive desires and behaviours: one rather focuses one's attention outwards to others. To those for whom the attention of others is uncomfortable the release from such social self-consciousness is highly desirable, and provides a motivation to drink. Of course, the kind of disinhibition that comes from assiduously seeking to secure a kind of self-escape can mean that a subject does things and appears in ways that when they come back to think about them, either on the basis of memory or report, lead to painful self-reflection. But, in the moment, drinking helps with self-escape.

Social self-consciousness is not the only kind of self-consciousness a subject may try to escape. In non-social settings as well some subjects will be particularly given – either by circumstance, cognitive disposition or character – to reflect on and monitor themselves and their own thoughts. Painful, repetitive or obsessive self-focused thinking can be occasioned by interaction with others, but are often also associated with stress, anxiety, depression and shame and lack of self-esteem.

Our suggestion is that negative self-consciousness involves the subject being in a costly emotional state from which she is likely to seek relief through self-escape. Addictive drugs tend to facilitate such escape. There is some evidence (Stewig and Tangney) and we hypothesize, that highly self-conscious individuals will be more likely to have addictive desires than low self-conscious individuals. In talking about highly self-conscious individuals, psychologists usually mean individuals who measure as self-conscious using the Self-Consciousness Scale of Fenigstein, Schemer & Buss 1975. It is designed to pick out the extent to which a subject has a tendency to direct attention

toward herself. The tendency is taken to manifest in the following ways: (i) an intense focus on behaviour past, present and future, (ii) a heightened sensitivity to privately experienced feelings, (iii) recognition of positive and negative characteristics in oneself, (iv) a tendency to introspect (v) imagining oneself (vi) an awareness of how one appears to others, (vii) concerns about others' concerns (Buss 2001). A distinction is drawn between private and public self-consciousness. Private self-consciousness is taken to be a tendency to focus on aspects of oneself not obviously accessible to others. Public self-consciousness is taken to be the tendency to focus on aspects of oneself that are observable or public, for example one's appearance and behaviour.

We think that high levels of both private self-consciousness and public self-consciousness, when associated with negative circumstances and appraisals, can be painful and can motivate a tendency to seek self-escape. It seems common sense to say that, whatever the basis for the need for self-escape, addictive substances seem to enable a subject to secure it to some degree or other. It is our suggestion that the motivation for self-escape in individuals is part of what makes addictive desires so strong. This suggestion rests on making plausible at least the following two claims. One, that addictive substances *do* tend to secure self-escape, and break a tendency for self-focussed consciousness. Two, that subjects are motivated in their use of these substances by their power to secure self-escape.

Let us start with the first claim. Although the suggestion that addictive substances facilitate self-escape is, it seems to us, a clearly robust one, the defence of it offered so far has drawn on little more than an invitation to phenomenological reflection. There are, however, studies that suggest that this is not an illusion. Hull, for example, in his defence of what he calls 'The self-awareness model' of alcohol consumption argues, on the basis of a series of experiments, that 'alcohol interferes with

an encoding process to a state of self-awareness' (Hull and Reilly, 1983: 514). In the first experiment subjects, who were told that they were drinking alcohol, gave short speeches after they drank either an alcoholic beverage or a placebo. Hull found that 'Subjects who consumed alcohol made significantly fewer self-focussed statements, and used significantly fewer first person pronouns than did subjects who consumed the placebo' (Hull and Reilly, 1983: 514). He confirmed in a follow up that the effect was independent of expectation. Further Hull found that in individuals with 'high private self-consciousness' alcohol consumption reduced self-relevant coding to the level of individuals with 'low private self-consciousness (measured by Fenigstein, Schemer & Buss 1975). Taking self-focussed judgments and use of the first person as markers of self-awareness Hull's data suggests that alcohol consumption leads to a reduction of self-awareness.

Moeller and Goldstein's review of impaired self-awareness in human addiction also suggests a connection between drug use and consequent reduced self-awareness. The view mainly concerns data on activation of the area of the brain referred to as rACC/vmPFCt – the rostral anterior cingulate cortex, and adjacent anterior medial prefrontal cortex. According to Moeller and Goldstein, the rACC/vmPFC

forms part of the network that supports self-related processing, with the specific hypothesized function of tagging stimuli as being personally relevant (van der Meer et al. 2010, Abraham 2013). This region is involved in generating the affective response that may occur shortly after error commission (Edwards et al. 2012) and in resolving emotional conflict (Etkin et al. 2006). Relatedly, the rACC activates during the experience of negative self-conscious emotions, including embarrassment (Sturm et al. 2013), and after receiving emotion-laden feedback (Mies et al. 2011). (Moeller and Goldstein, 2004:635)

Moeller and Goldstein report lower rACC/vmPFC activation in drug users, when related to non-drug related cues. They summarize the findings as being able to 'collectively indicate that errors fail to engender the appropriate tagging of personal relevance or emotional significance in addicted individuals' (p.638). One might worry that the correlation between lower activation in the part of the brain responsible for processing self-relevant information and state suggests that addicts, far from being highly self-conscious individuals seeking self-escape, are in fact individuals who lack capacity for self-relevant processing. However, the data in fact suggests that the effects on the capacity to self-process are not permanent and that when addicts stop taking their drug of choice their frontal cortex activity begins to increase again. This suggests room for a cycle of drug use based on self-escape and the return of previous levels of self-consciousness. Further, the data shows that abstinent addicts show a *higher* activation of the fACC/vmPCC than either active addicts, recent abstinent, or controls. This does give some data in favour of the first element of the self-awareness theory – the element that suggested that addictive substances attenuate self-awareness: measured by rACC/vmPFC activation there seems to be a rise in self-awareness when drugs are no longer used, and a suppression while they are. Moeller and Goldstein hypothesize that the high activation of rACC/vmPFC in abstinent addicts may be due to increased self-awareness as a result of the subject working to avoid cues. However, that seems only a guess – and why would the long term abstinent be working harder to avoid processing cues as self-relevant than the short term abstinent? Another possibility would be that the high levels of self-awareness amount the group of abstinent addicts marks a return to 'normal' for them, and stands as a disposition from which they sought escape.

What of the second claim, that addicts use self-awareness disrupting substances *in order to* reduce self-awareness? Even if alcohol – and perhaps other addictive substances – do lead to reduced self-awareness have we any evidence that they are consumed due to an appetite for this effect?

A first naïve thought one might have is this: in general, if something has a particular effect, and people have experience of the effect, and the effect is desirable, people will use the thing as a means of getting the effect. And, in the case of drinking, all three conditions mentioned in the antecedent of the general claim seem to be satisfied. It does not seem contentious to claim that most people, not only addicts, who enjoy drinking alcohol have an appetite for the release from self-consciousness that it brings.

We also have evidence of the simple correlative claim that highly self-conscious individuals, suffering painful forms of self-consciousness, seem to be at high risk of substance abuse. (See Hull, Randles and Tracey, Steuwig and Tangney). And the claim should come as little surprise: highly self-conscious individuals will, in focussing on themselves, have more occasion to negatively evaluate themselves, and be sensitive to negative evaluations of themselves by others. That will lead to a painful self-consciousness from which they might seek escape, either through trying to change themselves in the relevant ways or, as Hull puts it, seek 'avoidance of the self-aware state' (1981: 588).

Also relevant are Hull's studies in which known negative evaluations of the self are related to alcohol consumption. Hull's data suggests that while relapse rates of alcoholics with low self-consciousness ratings are *not affected* by negative evaluation from life events associated with failure, relapse rates of alcoholics with high self-consciousness ratings are *highly affected*. This suggests that it is not the negative evaluation itself that causes a subject to seek out alcohol but the reminder of it in self-

focussed thinking. In a couple of further experiments Hull found that individuals with 'high private self-consciousness' and a negative evaluation in a task just completed, drank considerably more than individuals with 'high private self-consciousness' and a positive evaluation. This suggests that, at least for a significant group of addicts, their high self-consciousness is a factor in their seeking out their chosen drug. Given that we know the chosen drug reduces self-consciousness it is plausible to hypothesize that such addicts are motivated by a desire for self-escape.

One final area to look for evidence is in what has proven effective in treatment. The evidence for effective treatment is difficult to gather, and given the lack of approved neuropharmacological treatments for addiction, there is not consensus on what kinds of interventions best help reduce the strength of, or increase the efficacy of the subject in resisting, addictive desires. What, however, seems undeniable is that treatment options that work to reduce painful self-focussed thinking, either by focussing on others, or by improving self-esteem, have had some efficacy in reducing addictive behaviour. Furthermore, heavy use of addictive substances often disappears spontaneously when individuals change their circumstances, in particular their social circumstances.

There are two aspects to the suggestion being offered in this paper – one completely general and one specific to the form of motivation we think is operative in many cases of addiction.

The completely general point is that the whole story of human addiction cannot lie in the capacity to certain substances to 'hijack' the human brain. This is absolutely obvious on the assumption that our brains are similar to each other: the incidence of addiction is too unevenly spread in the human population for the action of commonly available substances to be its sole cause. However, even if human brains vary considerably, we have excellent evidence that the 'hijack' story cannot be the whole

story. The same human with the same brain can cease to exhibit addictive behaviour just through changing circumstances. And humans with very similar brains – identical twins – are not even close to being determined to either be an addict or not. So, if we are seeking to understand addiction it is clear that other psychological or circumstantial causal variables must be involved.

The more specific suggestion is that an appetite for self-escape is a further important causal variable for addictive desires. It may be – for all we have said – that the appetite for self-escape is only one among other causal variables, and may explain the strength of addictive desires only for a sub-group of addicts. Although we think that the appetite for self-escape gives us a means of explaining why addictive desires for substances that attenuate self-consciousness are so strong, there may be other ways. We might wonder, for example, whether addictive desires are not sometimes as strong as they are due to an appetite to escape tension, anxiety or pain in general, as opposed to painful self-consciousness in particular. Hull argues against the tension hypothesis on the grounds that there is no clear evidence that alcohol, at least, acts to alleviate tension directly – in the way we do have evidence that it acts to disrupt the exercise of self-conscious cognitive capacities. However, Hull's data also suggests that high self-conscious individuals are subject to different patterns of alcohol use to low self-conscious individuals: the high self-conscious group relapse under the pressure of negative self-evaluation to a significantly greater extent than the low self-conscious. That leaves open the possibility that low self-conscious subgroups are drinking alcohol for other appetites, or due to other mechanisms. Our pluralism does not rule out further systematic causes for different sub-groups. Furthermore, even if the phenomenological and social scientific evidence that we have offered for the hypothesis that self-consciousness is relevant to addictive desires does not convince, the point

about the insufficiency of the neurophysiological explanation is unaffected: there is a need for supplementary causal variables.

We may wonder at this point whether defenders of the dominance of the 'hijack' view might say that while self-escape – or other possible causes of addiction – may be needed in *some* cases to explain why addictive desires are so strong, nevertheless all addictive desires will involve a dysfunction of the dopamine system. In so far as it is a necessary condition of an addictive desire that it involve a dysfunction in the dopamine system that explanation has a claim to being the dominant explanation.

The first thing to say is that even if the action of the dopamine system were to be established to be a necessary condition, or in fact involved in *every* case of an addictive desire, that would not be sufficient to show it to be the dominant explanation. There is the possibility that in those cases where self-escape is operative it is a more important cause than dopamine spikes, because although both are present, the desire for self-escape in doing more of the work.

The way we have set things up here has been not to take issue with the claim that a significant part of the explanation of addictive desires lies in the functioning of the dopamine system. Indeed, given that the actual addictive substances we use both intoxicate *and* act on the dopamine in system in systematic ways – they make more dopamine available – we do not deny that explanations of addictive desires will involve appeal to the activity of the dopamine system. We have sought to raise a question about whether it has been established that addictive desires were *disproportionately* strong, i.e. strong relative to the amount of pleasure addictive substances cause, but we have not denied that dopamine spikes play a role in addiction. However, we might ask whether they must. Suppose we (or most likely David Nutt) manage to develop and market two substances: Balcohol and Calcohol. Balcohol has exactly the same effect on

our dopamine system as alcohol – those who drink it get the same dopamine spikes and taste pleasure without *any* effects on our self-consciousness. Calcohol gives us the intoxicating effects and the release from self-consciousness without any dopamine spikes.⁹

It is an open question whether humans would develop addictive desires for Balcohol or Calcohol. The neurophysiological monist must take Balcohol to be addictive in the way alcohol is, and deny that Calcohol is addictive at all. The pluralist can allow that both, or neither, might turn out to be addictive. We think it very likely that those who find it hard being ‘in their own heads’ would over use Calcohol with bad consequences. We also think it very likely that fewer individuals will get hooked on Balcohol than alcohol.

Thinking about the distinction between Balcohol and Calcohol raises the question: which would it be worse to be addicted to? Which would be more problematic in itself and which would lead to worse consequences, assuming each were the same cost and had the same availability? Answering these questions brings out advantages in the idea that self-escape is a primary motivation causing addictive desires. First, it goes some way to explaining an element of the pejorative aspect of the adage ‘addict’, and, second, it brings out a personal level psychological reasons for why overuse of a self-consciousness suppressing substance might lead to further overuse. The motivation of self-escape is not any old motive of pain avoidance. To seek to get away from one’s own thought and evaluative feelings, to get away from others’ view of us, to shield ourselves from self-relevant information, is in some way to seek to get away from ourselves. And there is a creeping worry about the person who gives themselves away to the drugs

⁹ David Nutt is a neuropharmacologist at Imperial College London who is developing synthetic alcohol substitutes which aim to give the drinker the desired effects of drunkenness without many of the risks of alcohol. The substance he is seeking to produce is close to what we have called Calcohol.

they use. The worry is both that to do so is to lose the capacity to self-rectify, to self-improve and to self-control, but it is also that to do so is to cede the standing from which one is expected to take ethical responsibility for living well. Our concern is not just that the person themselves will suffer, and that they will hurt others, as a result of their drug use, it is also that one loses a core aspect of the person one is concerned for. Moreover, acting on the motivation for self-escape is to act on a motivation that, if it is too often acted on, will undermine the means one has of escaping the addiction. If we were to develop Balcohol and Calcohol and it turned out that one *could* be addicted to both, our betting is that the Calcohol addict is going to worry us more, and is going to get a rougher time of it.

Perhaps it will come as no surprise that, tasked with thinking about addiction, we have identified an explanation that we as philosophers are equipped to help in the delineation and description of. One might worry that we are like the man who looks for his lost keys only under the street lamp, because that is where the light is. But note, it would be equally surprising if philosophy were no more able to help with understanding human addictive desire than able to help understand tuberculosis. We are pretty sure that the keys – or some of them – are under the street lamp.

Daniel Morgan, University of Barcelona

danielmorgan@ub.edu

Lucy O' Brien, University College London

l.o'brien@ucl.ac.uk

References:

Abraham, A. 2013: The world according to me: personal relevance and the medial prefrontal cortex. *Frontiers in Human Neuroscience* 7, 341.

Alexander, B.K., Beyerstein, B.L., Hadaway, P.F., and Coombs.R.B. 1981: Effect of early and later colony housing on oral ingestion of morphine in rats. *Pharmacology, Biochemistry and Behavior*, Vol 15, 4:571-576.

Berridge, K. 2000: Measuring hedonic impact in animals and infants: microstructure of affective taste reactivity patterns. *Neuroscience and Biobehavioral Reviews* 24 (2000) 173-198.

Buss, A. 2001: *Psychological Dimensions of the Self*. Thousand Oaks, CA: Sage Publications.

Berridge, KC, Robinson TE. 1998: What is the role of dopamine in reward: hedonic impact, reward learning, or incentive salience? *Brain Research, Brain Research Reviews*. Dec 28(3): 309-69.

Campbell, J. 2006: An interventionist approach to causation in psychology, in A. Gopnik & L.J.Schulz (eds.), *Causal Learning: Psychology, Philosophy and Computation*. Oxford: OUP, 58-66.

Dill, B. and Holton, R. 2014: The addict in us all, *Frontiers in Psychiatry* 5, 139: 01-20.

Edwards, B.G. et al, 2012. Joint ICA of ERP and fMRI during error-monitoring, *Neuroimage* 59, 1896-1903.

Enoch, M. A., & Goldman, D., 2001: The genetics of alcoholism and alcohol abuse, *Current Psychiatry Report*, 3, 2: 144-51.

Etkin, . et al. 2006: Resolving emotional conflict: a role for the rostral anterior cingulate cortex in modulating activity in the amygdala. *Neuron* 51: 871-882.

Fenigstein, A., Scheier, M., and Buss, A. 1975: Public and Private Self-Consciousness: Assessment and Theory. *Journal of Consulting and Clinical Psychology* 43 (4): 522-527.

Holton, R. 2009: *Willing, Wanting, Waiting*, Oxford: OUP.

Holton, R. and Berridge, K. 2013: Addiction between Compulsion and Choice, in Neil Levy (ed.), *Addiction and Self-Control*. Oxford: OUP.

Hull, J 1981. A self-awareness model of the causes and effects of alcohol consumption, *Journal of Abnormal Psychology*, Vol. 90, No. 6, 586-600.

Hull, J. and Reilly, N. 1983: Self-Awareness, Self-Regulation, and Alcohol Consumption: A Reply to Wilson. *Journal of Abnormal Psychology* Vol 92, No 4, 514-519.

Levy, N. 2014: Addiction is not a brain disease (and it matters), *Frontiers in Psychiatry* 4: 24.

McClure SM., Daw ND., Montague PR. 2003: A computational substrate for incentive salience, *Trends in Neuroscience* 26.8: 423-8.

Mies, G.W. et al. 2011: The anterior cingulate cortex responds differently to the validity and valence of feedback in a time-estimation task. *Neuroimage* 56, 2321-2328.

Moeller SJ, and Goldstein RZ 2014: Impaired self-awareness in human addiction: deficient attribution of personal relevance. *Trends in Cognitive Science* . 18, 12: 635-41.

Potter-Efron, R. and Carruth, B. 1989: *Shame, Guilt, and Alcoholism: Treatment Issues in Clinical Practice*. Haworth Addictions Treatment.

Prescott, C.A. and Kendler, K.S. 1999: Genetic and environmental contributions to alcohol abuse and dependence in a population-based sample of male twins. *American Journal of Psychiatry*, 156, 34-40

Randles, D. and Tracy, J.L. 2013: Nonverbal displays of shame predict relapse and declining health in recovering alcoholics. *Clinical Psychological Science*, 1, 149-155.

Robinson, T.E. and Berridge K. C. 1993: The neural basis of drug craving: an incentive salience sensitization theory of addiction. *Brain Research, Brain Research Reviews*, 18, 247-91.

Schroeder, T. 2004: *Three Faces of Desire*. New York: Oxford University Press.

Schultz, W., Dayan, P. and Montague, R.R. 1997: A neural substrate of prediction and reward. *Science* 275:1593-1599.

Schultz, W. 2006: Behavioural theories and neurophysiology of reward. *Annual Review of Psychology*, 57, 87-115.

Stuewig, J. and Tangney J.P. 2007: Shame and guilt in antisocial and risky behaviours. In J.L. Tracey, R.W. Robins and J.P. Tangney (eds.), *The Self-Conscious Emotions: Theory and Research*.

Sturm, V.E., Sollberger, M., Seeley, W.W., Rankin, K.P., Ascher, E.A., Rosen, H.J., Miller, B.L., and Levenson, R.W. 2013: Role of right pregenual anterior cingulate cortex in self-conscious emotional reactivity. *Social Cognitive and Affective Neuroscience*, 8, 468-474.

Tindell, A.J., Berridge, K.C., Zhang, J., Peciña, S. and Aldridge, W.A. 2005: Central pallidal neurons code incentive motivation: amplification by mesolimbic sensitization and amphetamine. *European Journal of Neuroscience*, 22, 2617-2634.

Tobler, P.N., Dickinson, A. and Schultz, W. 2003: Coding of predicted reward omission by dopamine neurons in a conditioned inhibition paradigm. *Journal of Neuroscience* 23, 10402-10410.

Treeby, M.S. 2011: *Shame and Guilt: Implications for the Regulation of Alcohol Use*. Phd Thesis, University of Tasmania.

van der Meer, L. 2010: Self-reflection and the brain: a theoretical review and analysis of neuroimaging studies with implications for schizophrenia. *Neuroscience and Biobehavioural Reviews* 34, 935-946.

Wilson, T. 1983: Self-awareness, self-regulation, and alcohol consumption: an analysis of J. Hull's Model. *Journal of Abnormal Psychology*, 92, 505-513.

Zhuang X., Oosting R.S., Jones S.R., Gainetdinov, R.R., Miller, G.W., Caron M.G., and Hen, R. 2001: Hyperactivity impaired response habituation in hyperdopaminergic mice. *Proceedings of the National Academy of Sciences of the United States of America*, 98, 1982-1987.