DECONSTRUCTING AND RECONSTRUCTING BEHAVIOUR RELEVANT TO MENTAL HEALTH DISORDERS: THE BENEFITS OF A PSYCHOLOGICAL APPROACH, WITH A FOCUS ON ADDICTION

Lydia G Rutherford & Amy L Milton

Department of Psychology, University of Cambridge

Abstract: 170 words

Main text: 7311 words

Figures and tables: 2

References: 158

Corresponding author:

Dr Amy L Milton

Department of Psychology

University of Cambridge

Downing Site

Cambridge

CB2 3EB

UK

Tel: +44 1223 333593

Email: alm46@cam.ac.uk

ABSTRACT

RUTHERFORD, L.G. and Milton, A.L. Deconstructing and reconstructing behaviour relevant to mental health disorders: what can psychology offer? NEUROSCI BIOBEHAV REV XX(X)XXX-XXX, 2021. – Current treatments for mental health disorders are successful only for some patients, and there is an unmet clinical need for new treatment development. One challenge for treatment development has been how best to model complex human conditions in animals, where mechanism can be more readily studied with a range of neuroscientific techniques. We suggest that an approach to modelling based on associative animal learning theory provides a good framework for deconstructing complex mental health disorders such that they can be studied in animals. These individual simple models can subsequently be used in combination to 'reconstruct' a more complex model of the mental health disorder of interest. Using examples primarily from the field of drug addiction, we explore the 'psychological approach' and suggest that in addition to facilitating translation and backtranslation of tasks between animal models and patients, it is also highly concordant with the concept of triangulation.

HIGHLIGHTS
-   There is an unmet clinical need for new treatments for mental health disorders
-   Psychology may provide a framework for understanding mental health disorders
-   The psychological approach facilitates translation and backtranslation
-   A framework for studying mental health will facilitate triangulation of findings

1. INTRODUCTION

Mental health disorders affect a substantial proportion of the population worldwide (Ritchie and Roser 2018) and are responsible for profound economic costs and reductions in quality of life (Olesen et al. 2012). Although both psychological and pharmacological treatments are available for many mental health disorders, it is estimated that 35-50% of those affected by mental health disorders do not receive any form of treatment (Wittchen et al. 2011). For those who do receive treatment, many will relapse upon cessation of therapy (Holmes et al. 2014). Even for treatments that do have a good level of success in the clinic, the underlying mechanisms – and how these might be optimised – are often unclear. Therefore, there is a clear clinical need for new and better treatments for mental health disorders, as well as more efficient procedures for identifying symptoms.

Many of the challenges in developing treatments for mental health disorders arises because these disorders are associated with differences in brain structure and function, mediated by changes at many different levels of analysis. Genetic differences can predispose individuals towards developing, or being resilient to, specific mental health disorders (Ducci and Goldman 2012; Ask et al. 2021). The environment experienced by an individual also contributes, including both an individual's current environment and their early life experience (Daskalakis et al. 2013). Evidence also suggests that the environment may influence the risk of developing mental health disorders prior to birth, both through the intrauterine environment (Schlotz and Phillips 2009; though see also O'Donnell and Meaney 2017) and epigenetic changes (Babenko et al. 2015). These genetic, epigenetic, and environmental factors serve to modify the structure and function of neurons, neural circuits and, ultimately, behaviour. Thus, mental health disorders can be considered to be caused by factors ranging from genes to behaviour.

The complexity of mental health disorders has led to a reliance on models, and especially animal models. By definition, a model simplifies a complex real-world phenomenon such that it becomes more tractable to investigation; from the perspective of understanding mental health disorders and

the mechanisms underlying them, this is clearly a strength. However, this strength also gives rise to the limitation that a model will never fully encapsulate the mental health disorder of interest, potentially giving rise to a 'translation gap'. Ultimately, what is needed is a 'triangulation' or 'convergence' approach, in which different experimental approaches at multiple levels of analysis are used to address a single hypothesis (Munafò and Davey Smith 2018), and in which the limitations of any individual model can be considered in the context of a wider body of experimental evidence.

Our research focuses on the use of animal models to capture psychological processes and behaviours that are relevant to mental health disorders. We have been particularly influenced by the approach of animal associative learning theory (Mackintosh 1974; Dickinson 1994; Everitt et al. 1999), which allows the assessment of specific psychological processes through the use of carefully designed behavioural tasks. Although any individual task is only likely to capture one small aspect of behaviour relevant to a mental health disorder, here we advocate an approach in which these behaviours are deconstructed into their underlying psychological processes, and subsequently reconstructed by considering how manipulations affect each of these processes individually. This 'psychological approach' is already used by a number of labs worldwide (Venniro et al. 2020), and below we consider more specific examples from our own research and other investigators. Due to the broad range of literature that could be considered under the umbrella of 'mental health disorders', we have necessarily been selective in our discussion of the literature and have chosen to focus primarily upon rodent models of drug and alcohol addiction in our illustrative examples. We should emphasise that our choice of this literature is not intended to diminish the importance of other species, including non-human primate and invertebrate models, both of which have contributed markedly to understanding of mental health disorders and their biological mechanisms (Murray et al. 2011; Narayanan and Rothenfluh 2016; Dwyer 2017; Oikonomidis et al. 2017; Cognigni et al. 2018; (Banks and Negus 2017)). Furthermore, we do not intend to appear dismissive of ethological approaches to animal behaviour (Mobbs and Kim 2015; Peters et al. 2015); these approaches are also critical to triangulate the biological and psychological mechanisms that underlie mental health disorders. It is most likely that the most progress into understanding mental health

disorders will be made when an integrated approach, including scientists studying mental health disorders at all levels of analysis, is taken (Milton and Holmes 2018).

2. USING PSYCHOLOGY AS A FRAMEWORK FOR UNDERSTANDING MENTAL HEALTH DISORDERS

As noted above, mental health disorders are undeniably complex, being produced by a combination of specific vulnerability factors and genetic and epigenetic predispositions, interacting with the individual's past and current environment. There is debate within the clinical literature about how to best understand and classify mental health disorders (Casey et al. 2013). Both the Diagnostic and Statistical Manual of Mental Disorders (DSM-5) and International Classification of Diseases (ICD-11) take a **categorical approach** (see **Table 1** for a glossary of psychological terms, which are emboldened in the text at first use) to the diagnosis of mental health disorders, requiring a specific number and combination of symptoms to persist for a defined length of time for a specific mental health disorder diagnosis to be made. However, a number of investigators have begun to question this categorical approach (Casey et al. 2013; Gillan et al. 2017), arguing that a **dimensional approach** may be more appropriate. (Indeed, this is formalised by the US National Institute of Mental Health's Research Domain Criteria, or RDoC - Insel et al. 2010.) Many mental health disorders are characterised by changes in functioning of patients along specific dimensions – for example, **compulsive behaviour** is observed in both obsessive-compulsive disorder (OCD) and drug addiction – so it may be possible to consider a construct or dimension of 'compulsivity' as being similarly affected in both conditions (Figee et al. 2016; Gillan et al. 2017). According to this 'dimensional psychiatry' view (Robbins et al. 2012; Heinz et al. 2016), each dimension represents a continuum, and the differences between disorders would depend upon how functioning along other dimensions (e.g. impulsivity, anxiety) are affected, giving rise to individual patient profiles. Ultimately, the functioning of an individual along these dimensions would depend on both innate and environmental factors as outlined above.

In the face of such complexity, how can progress be made in understanding mental health disorders? As in all branches of science, one prominent approach is to simplify the phenomenon of interest by using models. Our favoured approach – and one that we argue for here – is to use psychological processes as a framework to understand mental health disorders. Specifically, we argue that using behaviourist approaches, with a focus on **pavlovian** and **instrumental** associative learning, provides a method for analysing specific behaviours, allowing them to be deconstructed and ultimately reconstructed to provide a fuller understanding of the model, and the phenomenon of interest.

2.1. Modelling mental health disorders as disorders of pavlovian and instrumental learning: the example of drug addiction

Drug addiction is a chronic, relapsing disorder in which drug-seeking and drug-taking behaviours come to dominate the behavioural repertoire of the individual, at the expense of other activities and rewards, and continue despite adverse consequences (American Psychiatric Association 2013). There are many different theories of drug addiction (Jentsch and Taylor 1999; Robinson and Berridge 2000; Koob and Le Moal 2008) but one prominent class of theory has postulated that drug addiction can be conceptualised as a disorder of learning and memory (Everitt et al. 2001; Hyman et al. 2006; Torregrossa et al. 2011; Tronson and Taylor 2013). Specifically, addiction can be considered to involve the aberrant engagement of pavlovian and instrumental learning processes normally engaged in adaptive foraging behaviour (Everitt et al. 2001). The acts of **drug-seeking** and **drug-taking** are instrumental behaviours, in which an action is reinforced by an outcome. Thus, in the early stages of drug-seeking and drug-taking, these behaviours are **goal-directed** and under the control of the individual. Furthermore, drug-seeking and drug-taking do not happen in isolation but are associated with the reinforcer of the drug high in specific places (contexts) and in the presence of specific environmental cues (such as people and drug paraphernalia) that come to be associated with the drug high in a pavlovian manner. These pavlovian and instrumental memories can interact,

such that the pavlovian cues can increase the likelihood of drug-seeking and drug-taking. This occurs through at least three psychologically and neurobiologically dissociable processes: **pavlovian conditioned approach**, **conditioned motivation**, and **conditioned reinforcement** (Milton and Everitt 2010; Milton and Everitt 2012). These 'positive' associations between the drug high and contexts and conditioned stimuli can invigorate ongoing drug-seeking and drug-taking behaviour. Furthermore, cues associated with psychological or physiological drug withdrawal can also act to support **conditioned withdrawal** (Wikler 1948; McAuliffe 1982), with drug-seeking and drug-taking alleviating withdrawal through **negative reinforcement**. Thus, both positive and negative pavlovian cues can become potent precipitators of relapse when an individual is trying to maintain abstinence.

Pavlovian and instrumental drug-associated cues can interact to influence behaviour during the acquisition of drug-seeking and drug-taking behaviour and can continue to act if an individual subsequently develops compulsive drug use. Controlled, recreational drug use, which is 'model-based' or 'goal-directed' can transition to 'model-free', **habitual** drug-seeking behaviour (Everitt and Robbins 2005; Everitt and Robbins 2013; Everitt and Robbins 2016), although it should be noted that model-free systems still incorporate a representation of reward value, and therefore stimulus-response systems more likely mediate habitual behaviour (Miller et al. 2019). This shift to habitual drug-seeking can occur in individuals who are not addicted; the final determinant of *addictive* behaviour is a loss of control over the drug-seeking behaviour, which leads to *compulsive* habitual drug-seeking, and which only occurs in a subset of individuals who may otherwise show high levels of drug intake and high motivation for drug (Deroche-Gamonet et al. 2004; Pelloux et al. 2007; Giuliano et al. 2018). Thus, although the vast majority of animals will acquire drug self-administration behaviour when given the opportunity to do so, only a minority of animals go on to develop behaviour consistent with addiction in human patients. This is consistent with findings in humans, where a relatively small percentage (approximately 20%) of those experienced with drugs of abuse subsequently become addicted (Anthony et al. 1994). Thus, understanding the influence that pavlovian and instrumental associations have on drug-seeking behaviour can be studied separately

from the transition of these 'adaptive' processes into the 'maladaptive' memories that subsequently support addictive-like behaviour.

As is the case for drug addiction in patients, the likelihood of any individual in the model population showing compulsive behaviour is partly due to genetic predispositions (such as differences in striatal structure and function leading to increased **impulsivity** - Dalley et al. 2007; Belin et al. 2008; Caprioli et al. 2014) and environmental factors, such as extended access to the drug of abuse (Vanderschuren and Everitt 2004; Pelloux et al. 2007; Giuliano et al. 2018). However, regardless of whether the development of addiction in a specific individual is more dependent on genetic or environmental risk factors, these examples should illustrate the explanatory value of taking a psychological approach in understanding a mental health disorder as complex as addiction.

2.2. Advantages of using a psychological approach informed by animal learning theory: deconstructing and reconstructing behaviour

With this psychological approach underpinning our understanding of addiction, it becomes possible to interrogate specific aspects of addiction as a disorder using different behavioural tasks. If the reinforcing properties of drugs are the issue of interest, then **self-administration** procedures can be used in rodents, but may not require the extended access to drug that it is necessary for studies of habitual behaviour. If drug-seeking habits are of interest, then behavioural tasks can be designed that promote habitual over goal-directed behaviour (for example, extensive training and the use of **interval schedules**). If compulsive drug-seeking is required, then it is necessary to introduce some form of **punishment** for drug-seeking that will identify the subpopulation who are resistant to punishment. Here we consider an illustrative example drawn from our own research, but similar examples of deconstructing and reconstructing complex behaviour into translational psychological constructs relevant to addiction can be found for stress (Sinha et al. 2011; Spanagel et al. 2014), social context (Heilig et al. 2016), behavioural allocation (Ahmed et al. 2013; Banks and Negus 2017) and the resistance of drug-seeking to punishment (Jean-Richard-Dit-Bressel et al. 2018).

Furthermore, while we focus on addiction as an example, the general principles of our argument are more widely applicable to mental health disorders including post-traumatic stress disorder (Yehuda and LeDoux 2007; Richter-Levin et al. 2019), obsessive-compulsive disorder (Szechtman et al. 2017) and mood disorders (Robinson 2018).

2.2.1. An illustrative example of deconstructing and reconstructing behaviour: understanding cue-induced reinstatement following enforced abstinence

An advantage of the psychological approach is that it allows existing behavioural tasks to be analysed from different perspectives: an illustrative example comes from our own research investigating cue-induced reinstatement following enforced abstinence from cocaine-seeking in rats. We had found previously that the capacity of a drug-associated cue to support learning of a new instrumental response could be disrupted by targeting its **reconsolidation** with administration of either an NMDA receptor antagonist (Milton et al. 2008a) or the β-adrenergic receptor antagonist propranolol (Milton et al. 2008b). However, while our previous work had shown that we could disrupt the pavlovian cue-drug memory association, we had not established whether this would reduce the capacity of the cue to support previously learned instrumental drug-seeking. To address this question, we used a procedure in which rats were trained to self-administer cocaine with an instrumental lever press response, which was also associated with the presentation of a light cue acting as a **pavlovian conditioned stimulus** (CS). Subsequently, after a short period of abstinence in the home cage and inducing the cue-drug memory to destabilise and attempting to block its reconsolidation, we assessed whether the cue was capable of enhancing drug-seeking behaviour in an extinction test. We found that when the NMDA receptor antagonist MK-801 was administered in conjunction with memory reactivation (but not when administered alone in 'non-reactivated' controls), the levels of cue-induced reinstatement in the MK-801-treated group were lower than in control rats (Milton and Everitt 2009). Thus, our data were consistent with the hypothesis that the reconsolidation of the cue-drug memory had been disrupted by NMDA receptor antagonism, and

that the CS was no longer capable of enhancing drug-seeking behaviour (which had not been affected by our specific experimental manipulation, but can be disrupted - see Exton-McGuinness and Lee 2015).

While these data were promising, the prospects of translating this reconsolidation-based approach seemed more likely using an amnestic agent with a more straightforward route to the clinic than the NMDA receptor antagonist MK-801. We thus repeated the experiment using the β-adrenergic receptor antagonist propranolol. However, unlike MK-801, and unlike our previous findings (Milton et al. 2008b), propranolol given in conjunction with memory reactivation did not reduce the capacity of the cue to enhance previously acquired drug-seeking behaviour.

The apparent discrepancy in the data generated by these two experiments led us to analyse the **cue-induced reinstatement** procedure in more detail. In the cue-induced reinstatement procedure, the pavlovian cue was presented in close spatial proximity to where the instrumental drug-seeking response was made. This influences the manner in which the pavlovian cue can interact with instrumental behaviour. As we have previously reviewed in more detail (Milton and Everitt 2010; Milton and Everitt 2012), pavlovian cues can influence instrumental responding through three psychologically and neurobiologically dissociable processes: conditioned approach, conditioned motivation and conditioned reinforcement. The presentation of the pavlovian light cue above the active lever would allow the light, acting to support pavlovian conditioned approach (akin to the 'sign-tracking' phenomenon described elsewhere - Tomie et al. 1989; Flagel et al. 2007) to bring the rat into the vicinity of the lever instrumentally associated with cocaine self-administration. This, in itself, would facilitate instrumental responding. Furthermore, because the pavlovian light cue and the instrumental response had been associated with same reinforcer, the light would be able to invigorate instrumental responding through the process of pavlovian conditioned motivation (sometimes referred to as 'pavlovian-instrumental transfer'). Finally, because the memory was tested in the absence of reinforcer delivery, instrumental responding throughout the session would be supported by the conditioned reinforcing properties of the pavlovian light cue, which are highly persistent and resistant to extinction (Di Ciano and Everitt 2004). Thus, the 'cue-induced

reinstatement following enforced abstinence' procedure we had used (Milton and Everitt 2009) had the potential to engage all three pavlovian 'routes to relapse' (Milton and Everitt 2010).

Although we had shown that the memory underlying pavlovian conditioned reinforcement could be disrupted by β-adrenergic receptor antagonism (Milton et al. 2008b), we had not investigated whether drug memories supporting pavlovian conditioned motivation and conditioned approach could similarly have their reconsolidation disrupted with the administration of propranolol. Certainly, cue-sucrose memories supporting conditioned approach and conditioned motivation appeared not to require adrenergic signalling for their reconsolidation (Lee and Everitt 2008). Thus, we investigated the requirement for NMDA receptors and β-adrenergic receptors in the reconsolidation of cue-drug memories supporting conditioned approach and conditioned motivation (Milton et al. 2012). To facilitate comparison with cue-sucrose memories (Lee and Everitt 2008), we investigated the reconsolidation of cue-alcohol memories rather than cue-cocaine memories, as the **consummatory response** of drinking alcohol is more comparable to the consummatory eating of sucrose than the delivery of an intravenous cocaine infusions. Consistent with the findings on cue-sucrose memories, and the cue-induced reinstatement data, we found that while MK-801 was capable of disrupting the reconsolidation of all three 'routes to relapse', the β-adrenergic receptor antagonist propranolol only disrupted the reconsolidation of the memory underlying conditioned reinforcement. Importantly, this did not appear to be the result of changing the reinforcer from cocaine to alcohol, as subsequent work showed that propranolol could disrupt the reconsolidation of a cue-alcohol memory supporting conditioned reinforcement (Schramm et al. 2016). Thus, by deconstructing the cue-induced reinstatement task into its underlying psychological processes – instrumental drug-seeking, conditioned approach, conditioned motivation, and conditioned reinforcement – it was possible to interrogate these processes separately using specific behavioural procedures. This ultimately provided insight into why we observed such different effects in behavioural tasks which, on the face of it, appeared very similar in measuring the capacity of a pavlovian cue to influence instrumental behaviour.

2.2.2. Relating psychology to neuroscience: learning theory as a framework for guiding neural circuit interrogations

A further advantage of using a psychological approach to understand mental health disorders comes from the rich literature understanding the principles of these learning processes and relating them to neural regions, neurochemistry and neural circuits. Lesion studies, for example, have shown that the pavlovian and instrumental learning processes described above depend upon different regions of the amygdala and striatum, respectively (see Cardinal et al. 2002; Wassum and Izquierdo 2015, for review). Careful and systematic design of behavioural tasks to isolate specific psychological processes has revealed how impairments in responding can be caused by different deficits in psychological processes. One such example showing the power of this approach, even with 'crude' manipulations such as excitotoxic lesions, is the dissociation of different prefrontal regions in supporting cognitive flexibility. The Intra-Dimensional/Extra-Dimensional Set-Shift Task (IDED) was originally developed to further characterise cognitive deficits usually measured in patients using the Wisconsin Card Sorting Task (Grant and Berg 1948), allowing the independent assessment of different psychological processes underlying cognitive flexibility, including basic stimulus-reward learning, reversal learning, intra-dimensional attentional set-shifting and extra-dimensional set-shifting. As part of the Cambridge Neuropsychological Test Automated Battery (CANTAB), the test was specifically designed to be independent of language, making it well-suited for healthy human volunteers and patients across the lifespan. The independence of the task from language-based instructions made it readily backtranslatable to animal studies (Roberts et al. 1988; Dias et al. 1996a; Dias et al. 1996b; Dias et al. 1997). The ability to backtranslate the task to non-human primates allowed for the necessity of defined prefrontal regions to be investigated using controlled and defined excitotoxic lesions, giving further insight into the deficits shown by patients, whose brain damage and/or dysfunction typically do not respect anatomical boundaries. In turn, this has prompted observations of regional dysfunction in other patient groups – for example, changes in prefrontal function in addiction patients (see Goldstein and Volkow 2011, for review) – to be examined through the lens of specific deficits in dissociable psychological processes using the IDED (Banca et al. 2016).

A further example of the power of the psychological approach to inform understanding of neurochemistry relevant to mental health disorders arises from the study of affective biases, as relevant to major depressive disorder (MDD). As a disorder diagnosed partially on self-reported feelings of low mood, MDD would appear to pose a particular challenge for modelling in animals. However, by focusing on the effect of affective biases on decision-making, it has been possible to develop translational measures of affective bias for use in both humans (Aylward et al. 2020) and rodents (Stuart et al. 2013; Hales et al. 2016). The 'affective bias task' (ABT) measures the influence of affective state on the memory of a rewarding experience, with individuals biasing their behaviour on a choice test towards cues associated with a more rewarding experience (Stuart et al. 2013). The ABT captures the difference in temporal effect between traditional antidepressants and novel antidepressants such as ketamine (Stuart et al. 2015), and has provided insight into potential psychological mechanisms for this difference. As argued by Robinson (2018), it is plausible that rapid-acting antidepressants attenuate negative affective biases, allowing individuals to return to a neutral emotional state. By contrast, traditional antidepressants put individuals into a positive affective state, but are delayed in their antidepressant action because new incentive learning is required to overcome previously learned negative affective biases, before any effects can be seen on behaviour. Thus, the psychological approach and development of translational behavioural tasks has provided the opportunity to test novel hypotheses about the mechanisms of action of drugs with clinical efficacy.

Beyond assessing the contribution of specific brain regions to dissociable psychological processes, the psychological approach has also furthered understanding of neural circuits supporting specific behaviours. One such example concerns the hypothesised shift in dependence of drug-seeking behaviour from the ventral to the dorsal striatum as drug-seeking becomes habitual (Everitt and Robbins 2005; Everitt and Robbins 2016). Informed by anatomical evidence of the 'spiralling' dopaminergic circuitry connecting the dopaminergic midbrain to striatal subregions in both primates (Haber et al. 2000) and later in rodents (Ikemoto 2007), it was first established that extended experience of drug self-administration led to increased dopaminergic signalling within the

dorsolateral striatum (Ito et al. 2002), and that drug-seeking became dependent upon this dorsolateral striatal dopaminergic signalling (Vanderschuren et al. 2005). Subsequent studies began to interrogate the circuitry underlying this shift in the dependence of drug-seeking on different striatal regions, using a 'disconnection' approach. By lesioning the ventral striatum unilaterally, and antagonising dopaminergic signalling in the contralateral dorsal striatum, it was possible to show a serial ventral-to-dorsal serial circuit that supported well-established drug-seeking behaviour (Belin and Everitt 2008). Based on animal learning studies indicating that the dorsal striatum is necessary for habitual responding for non-drug reinforcers (Yin et al. 2004; Yin et al. 2005), it was subsequently shown that the drug-seeking behaviour supported by the dorsal striatum is habitual (Zapata et al. 2010).

Understanding this neural circuit has further contributed to understanding the psychological influence of cues on habitual behaviour, as similar disconnection approaches have been applied to determine how the amygdala contributes to goal-directed and habitual behaviour. It is known from studies of rodents responding for food reinforcers that projections from the basolateral amygdala (BLA) to the posterior dorsomedial striatum (pDMS) are necessary for goal-directed learning (Corbit et al. 2013), though the connection may be indirect, via the prelimbic cortex (Fisher et al. 2020). Similarly, goal-directed cocaine-seeking depends upon the pDMS, but the development of habitual cocaine-seeking depends upon the anterior dorsolateral striatum (aDLS; Murray et al. 2012). By disconnecting the amygdala and aDLS, it was subsequently shown that while the DLS control of drug-seeking is initially triggered by the BLA, its maintenance depends upon the central nucleus of the amygdala (Murray et al. 2015). The influence of drug-associated cues on behaviour increases as drug-seeking becomes habitual (Corbit and Janak 2016) and DLS-dependent drug-seeking is resistant to extinction of drug-associated pavlovian cues (Bender and Torregrossa 2021), consistent with the limited efficacy of prolonged cue exposure therapy for drug addiction (Mellentin et al. 2017). However, the DMS is not disengaged when drug-seeking becomes habitual (Vandaele et al. 2019) and inactivation of the DLS through antagonism of the AMPA subtype of glutamate receptor makes drug-seeking once again sensitive to cue extinction (Bender and Torregrossa 2021).

Electrophysiological recording of different neuronal populations has also provided further insight into correlations between particular associative learning processes and underlying patterns of neuronal activation, both at the population-level and the single-unit level. For example, electrophysiological recordings during a pavlovian-instrumental transfer task have corroborated lesion data regarding the contribution of the nucleus accumbens to pavlovian conditioned approach (Day et al. 2006) and showed subtle differences in the contribution of the nucleus accumbens core and shell to conditioned motivation, and the modulation of these responses by cocaine self-administration (Saddoris et al. 2011). They have also provided insight into differences in the responsivity of striatal subregions during instrumental tasks, with DMS neurons showing stronger phasic responses to reinforcement, consistent with the role of this region in goal-directed behaviour, and DLS showing stronger phasic responses to lever pressing, consistent with its role in habitual behaviour (Fanelli et al. 2013). In addition to providing insight into the processing undertaken by specific brain regions engaged by behavioural tasks, these correlative measures can also be used to guide subsequent causal manipulations.

New and specific approaches to modulating neuronal function, for example optogenetics (Boyden et al. 2005) and chemogenetics (Armbruster et al. 2007), have understandably generated excitement in the potential that they offer for understanding neural circuits. In particular, the cell-type specificity, projection-specificity and the fine temporal control of techniques allows more subtle manipulations than have previously been possible with lesions, temporary inactivation, or pharmacological manipulations. Furthermore, we would suggest that these genetic approaches have yielded particular insight when paired with specific behavioural tasks informed by a psychological approach.

Chemogenetic techniques have provided further insight into the circuit mechanisms supporting psychological processes, originally established through lesion and inactivation approaches. For example, the well-established role of the BLA in representing the affective value of pavlovian cues (as discussed above) has been further parcellated using chemogenetic techniques. Inactivating projections from the BLA to the orbitofrontal cortex impaired the capacity of a pavlovian cue to guide instrumental behaviour through conditioned motivation (Lichtenberg et al. 2017), and

chemogenetic and optogenetic manipulations showed that reciprocal projections from the lateral orbitofrontal cortex and medial orbitofrontal cortex supported the encoding and retrieval, respectively, of the memory associating specific rewards with instrumental responding (Malvaez et al. 2019). Similarly, building on previous findings indicating that projections from the BLA to the nucleus accumbens core are important for supporting conditioned reinforcement (Parkinson et al. 1999; Cardinal et al. 2002), chemogenetic inactivation of these projections was shown to acutely impair the capacity of a pavlovian cue to guide drug-seeking behaviour (Puaud et al. 2021).

The time-specificity afforded by optogenetic approaches, in addition to the projection-specificity of viral approaches, has also allowed further insight into the circuits supporting psychological processes, particularly those related to the learning and maintenance of pavlovian and instrumental memories. For example, it has long been known that drugs of abuse increase dopaminergic signalling from the ventral tegmental area to the nucleus accumbens (Wise and Bozarth 1987) and that limbic system dopaminergic responses increase to environmental cues as they come associated with motivationally relevant outcomes in a pavlovian manner (Schultz et al. 1993). Optogenetic activation of midbrain dopamine neurons, in close temporal proximity to cue presentation, was sufficient to transform neutral cues into pavlovian conditioned cues in terms of their effects on behaviour (Saunders et al. 2018), with activation of VTA neurons imbuing the cues with conditioned reinforcing properties, and activation of substantia nigra pars compacta neurons leading to general increases in cue-directed motivated behaviour. Consistently, optogenetic inhibition of VTA neurons during consumption of a food reward led to persistent learned decreases in responding to pavlovian cues associated with the same reward (Fischbach and Janak 2019). On instrumental tasks, optogenetic inhibition of VTA neurons prevented the initiation of instrumental responding and reduced ongoing instrumental behaviour (Fischbach-Weiss et al. 2018). However, although optogenetic stimulation of the VTA was capable of supporting instrumental responding, it could not support compulsive responding unless the orbitofrontal cortex was inhibited in a temporally-locked manner relative to punishment (Pascoli et al. 2018). Although this finding may appear to conflict with studies indicating that lesions of the orbitofrontal cortex did not affect compulsive drug-seeking (Pelloux et al. 2013), the effects of

orbitofrontal inactivation on compulsive responding were only temporary, and compulsive reward-seeking was in fact associated with *increased* plasticity between the orbitofrontal cortex and the striatum (Pascoli et al. 2018). The use of optogenetics, and particularly the temporal specificity afforded by this technique, allows further insight to be gained into the contribution of specific brain regions in supporting defined psychological processes.

2.2.3. Learning processes as computation: psychology as a structural framework for computational modelling

Theories with the ambition of describing animal learning in mathematical terms have a long and rich history in psychology. Theories, such as those proposed by Rescorla & Wagner (1972), Pearce & Hall (1980) aimed to describe learning in a formal mathematical way, which was readily translatable to subsequent computational modelling (e.g. Sutton and Barto 1981). There is a wealth of literature on computational modelling of learning, decision-making and mental health disorders that is beyond the scope of the current review (though the interested reader is referred to Mollick and Kober 2020, for a discussion of computational models relevant to drug addiction). However, it is clear that computational modelling, informed by psychology, allows for the development of specific predictions that can subsequently be tested in both humans and animals. Furthermore, some of the parameters required for such modelling – for example, **perseveration** quantified as stimulus or side 'stickiness' (Robbins and Cardinal 2019) – can be related to specific psychological processes and reveal new hypotheses to be tested experimentally.

2.2.4. Overcoming potential limitations of a psychological approach

As noted above, one of the major strengths of models – namely, their capacity to make the study of complex phenomena tractable by simplifying them – is necessary also a major limitation. Although the conceptualisation of mental health disorders as disorders of maladaptive learning and memory has produced major insights, modelling by definition discards 'unimportant' information which may,

nonetheless, be relevant to the patient situation. Examples of this include the 'pure' pharmacological effects of drugs of abuse in drug addiction, or sensitisation of stress systems following withdrawal from drugs – the 'dark side' of addiction (Logrip et al. 2012; Koob 2021).

Drugs of abuse are psychoactive and have effects on neuronal signalling and function that are separate from any effects on learning. These neuroadaptations are complex and depend upon a wide range of factors, including the pattern of administration (including whether the drug is actively or passively administered), the class of drug, duration of exposure, withdrawal periods, and the brain regions of interest (Marie et al. 2019). These can be studied using methods that do not rely on learning, such as the induction of alcohol dependence using ethanol vapour inhalation (Gilpin et al. 2008). However, it is worth noting that even these 'pure' pharmacological effects can interact with learning. Animals (and humans) learn to associate drug effects with specific environmental cues, giving rise to phenomena such as **conditioned tolerance** (Siegel 1977) and conditioned withdrawal (Wikler 1948; McAuliffe 1982; Kenny et al. 2006). Furthermore, these physiological and psychological withdrawal phenomena, including the **anhedonia** and **dysphoria** that constitute the 'dark side of addiction' (Koob 2021), can further interact with learning systems to support drug-seeking through negative reinforcement and the avoidance of both actual and conditioned withdrawal. Instrumental behaviour can also be modulated with the passive administration of psychostimulants, as experimenter-administered amphetamine (Nelson and Killcross 2006), methamphetamine (Furlong et al. 2018) and cocaine (LeBlanc et al. 2013) can all facilitate the development of habitual behaviour through neuroadaptations that are separable from habitual drug-seeking itself. Thus, understanding non-associative changes into the sensitisation of stress systems provides further insight into the behavioural changes found in addiction; similar types of analysis can be applied to other mental health disorders, such as post-traumatic stress disorder (Stam 2007). In both cases, it is clear that the use of a psychological framework can help to disentangle the 'physiological' from the 'psychological' aspects of mental health disorders, and that these provide further insight into the disorder of interest.

3. THE IMPORTANCE OF (BACK)TRANSLATION AND HOW PSYCHOLOGY CAN SUPPORT THIS

The aim of understanding the mechanisms that underlie mental health disorders is valid in its own right, but for many investigators there is a parallel interest in exploiting this knowledge to develop new and better treatments for mental health disorders. Where animal models are used to gain insights into mechanisms, these need to be translated to the human patient situation. To develop animal models, *backtranslation* is critical: using insights gained from the human clinical condition to inform and refine animal models. This can occur at many different levels, from genetic models to behavioural models, and ultimately translation and backtranslation should work together iteratively to facilitate basic insights into clinical innovation (Milton and Holmes 2018). Translation and backtranslation are not always straightforward, but we argue that again, focusing on psychological processes and behaviour can provide both a framework and a bridge between basic and clinical science.

3.1. Challenges for translation and backtranslation

There are both practical and theoretical challenges presented by translating and backtranslating (or 'reverse translating' - see Venniro et al. 2020) research relevant to mental health disorders. We will first consider the practical issues and theoretical issues before considering how a psychological approach may help to address these.

3.1.1. Practical challenges in translation and backtranslation

Practical challenges in translation often stem from the tension in understanding basic mechanisms in a highly refined manner while also adapting these insights into a treatment that could be readily administered non-invasively to humans. For example, many animal studies aiming to understand mental health disorders will use neural manipulation strategies – traditionally lesions or inactivations,

but now more commonly viral strategies such as optogenetics or chemogenetics of specific brain regions or projections – that cannot be readily applied to humans. However, we assert that these are practical issues. The effects of many manipulations to specific brain regions can also be seen when the same class of drug is administered systemically – for example, β-adrenergic receptor antagonism disrupts the reconsolidation of memories supporting cocaine conditioned place preference when administered directly to the basolateral amygdala (Bernardi et al. 2009) and when administered by systemic injection (Bernardi et al. 2006). In a similar manner, although the specific drugs used in animal studies may not be directly translatable to humans' patients, there are often other drugs with similar mechanisms of action that can be translated. This can also contribute to backtranslation; for example, anisomycin is a protein-synthesis inhibitor that is widely used in animals but has a similar mechanism of action to rapamycin (Sirolimus) which can be used in humans. Consequently, this has motivated the use of rapamycin in animal studies, where its effects have been consistent with those found using anisomycin (Blundell et al. 2008; Barak et al. 2013).

Another practical challenge, particularly for backtranslation, stems from differences in the types of tools used to assess human and animal behaviour. Many patient studies make use of validated questionnaires relevant to the mental health disorder under study, which undoubtedly provide valuable information about the severity and nature of the disorder in individual patients. While these questionnaires can also be administered to healthy human participant control groups, the reliance of language introduces an obvious difficulty in translating questionnaire-based findings to animal models. However, we would again assert that this is a practical issue. Many studies with patients use questionnaire measurements as part of a wider battery of cognitive and psychological assessments, many of which can be more directly backtranslated to animals. Test batteries such as CANTAB (Sahakian et al. 1988) and NEWMEDS (Hvoslef-Eide et al. 2015) are non-language-based and were designed to assess specific psychological processes via touchscreen tasks in both humans and animals. Other tasks, such as the Observing Response Task (Eagle et al. 2014; Morein-Zamir et al. 2018), a measure of compulsive-like checking relevant to obsessive-compulsive disorder, and the Response Bias Probabilistic Reward Task (Pizzagalli et al. 2008; Der-Avakian et al. 2013), which

measures psychological processing relevant to anhedonia in depression, have been developed specifically from a translational and backtranslation perspective, with parallel human and rodent versions of the task being developed through collaboration. Careful and inspired behavioural task design has also allowed other factors, such as social influences on drug-seeking behaviour, to be studied in animals (see Venniro et al. 2020, for review). Other investigators have developed human tasks that are analogous to tasks more traditionally used in rodents, including drug-conditioned place preference (Childs and de Wit 2009), pavlovian autoshaping to measure sign-tracking and goal-tracking (Wardle et al. 2018; Colaizzi et al. 2020) and pavlovian-instrumental transfer (Talmi et al. 2008).

Therefore, it is possible to develop translational and backtranslational tasks, although it should be acknowledged that this type of task development does raise some challenges. Even though tasks may be designed to target the same psychological processes in humans and animals, there are some potential but subtle confounds that can be overcome with careful task design. For example, tasks as simple as pavlovian fear conditioning with mild electric shock can be perceived very differently in humans and animals, as humans have shock values individually titrated to their own tolerance levels and are free to leave an experiment at any time, while rodents usually have no control over the magnitude and strength of shock, and no ability to escape the testing situation (as reviewed more extensively by Flores et al. 2018). Humans also more readily acquire tasks that require extensive training in experimental animals, often necessitating a more difficult version of the task to be developed for humans. (However, it should also be noted that human participants often receive task instructions, which is obviously not the case for animals performing analogous behavioural tasks.) Regardless, we would argue that these are again practical issues that can be overcome with careful experimental design.

Furthermore, great care needs to be taken to compare appropriate populations in the human and animal versions of behavioural tasks; for example, it has been argued recently that evidence from human populations is not consistent with the theory that addiction depends upon the development of habitual and compulsive drug-seeking behaviour, derived primarily from animal models (Hogarth

2020). However, many of the studies cited in support of the argument were from non-addicted drug users (e.g. non-daily smokers in Hogarth and Chase 2011) and even amongst dependent individuals, only a proportion will transition to compulsive drug use (as noted in §2.1). This may account for an apparent lack of difference in habit learning at the group level (Luijten et al. 2020), though it is worth noting that dependence severity correlates with a reliance on habits (Luijten et al. 2020). Thus, we agree with Epstein (2020) that it is critical to define what models encapsulate, and to whom they refer. 'Habit theories' of addiction may not apply to every drug user (see also Vandaele and Ahmed 2020), but they are relevant to the 20% of individuals who show persistence in drug use despite adverse consequences. If anything, these controversies within the addiction literature support the value of translational and backtranslational approaches, and the progress that can be made by close collaboration between preclinical and clinical investigators (see Venniro et al. 2020, for further evidence supporting this argument).

3.1.2. Theoretical challenges in translation and backtranslation

In addition to the practical challenges discussed above, there are also theoretical challenges to the aims of translation and backtranslation. One major challenge concerns whether the phenomenon of interest is directly observable or has a correlate in animals – for example, the intrusive thoughts that are a symptom of many different mental health disorders, or the hallucinations and delusions that characterise disorders such as schizophrenia. However, it is important to remember that the value of animal models comes from simplifying complex phenomena to more tractable questions, and in doing so these models serve a useful purpose in making underlying assumptions explicit. For example, what counts as an 'intrusive thought'? Rumination or delusion may be a step too far for animal models, but if we consider intrusive thoughts as similar to the flashbacks in post-traumatic stress disorder or craving for drugs in addiction – both elicited unconsciously by environmental 'trigger' conditioned stimuli – then the study of intrusive thought becomes more tractable. Furthermore, with developments in neuroscience techniques, such as repeated and time-separated

measures of neural activation, it becomes easier to see how these difficult phenomena could be studied in animals. It may not be possible to get a verbal self-report from animals, but it might be possible to observe behaviour normally associated with 'trigger' conditioned stimuli, or to measure reactivation of neuronal ensembles that encoded the trigger stimulus when it is presented subliminally.

A final theoretical problem presented by animal studies concerns the level of control and population variation in lab-based populations as compared to patients with mental health disorders. In the lab, subject histories and life experiences are highly standardised and very similar; this is not the case in clinical reality, where each individual instance of a mental health disorder occurs on a background of different life history, experiences, and individual predispositions. As discussed above, the issue of predispositions can be addressed in animals – indeed, some animals are bred with specific predispositions relevant to mental health disorders, such as alcohol-preferring 'P' rats (Bell et al. 2006) - with psychology providing a useful framework for considering predispositions such as trait anxiety or trait impulsivity (Voon and Dalley 2016). Once again, the value of models is in providing a level of abstraction that allows the noise of individual life experiences to be aggregated into more objective factors. Whether factors such as socioeconomic status – known to influence perceived control over the environment, and numerous mental health disorders (Moscarello and Hartley 2017) – are at their essence, the result of early life adversity, current stress, or an interaction of both is more readily addressable in a lab setting.

4. CONCLUSIONS: THE IMPORTANCE OF TRIANGULATION

Translational and backtranslational approaches are inextricably linked to reproducibility, which has been a challenge in psychological research in recent years (Open Science Collaboration 2015). A lack of replication could be thought to allude to false positives in the literature, blaming methodology apparently lacking the correct experimental controls (Shrout and Rodgers 2018). However, an alternative view is that science as a whole should be aiming for 'triangulation': that in order to obtain

reliable evidence for scientific hypotheses, multiple techniques and approaches should be used and the results from these different approaches integrated to inform whether a specific hypothesis is supported or not (Munafò and Davey Smith 2018). The concept of triangulation is highly concordant with our views, articulated above. Psychology provides a framework that allows common processes to be studied from animals to humans, allowing evidence from diverse approaches – for example, molecular biological techniques, genetic and epigenetic variation, single- and multiple-unit recordings, viral-based methods, pharmacology, functional imaging, connectivity analyses, computational modelling – to be integrated into tests of specific hypotheses. This combinatorial and collaborative approach may provide new insights into mental health disorders and help to both improve current therapeutic approaches and to develop new ones.

ACKNOWLEDGEMENTS

REFERENCES

Ahmed SH, Lenoir M, Guillem K. 2013. Neurobiology of addiction versus drug use driven by lack of choice. *Current Opinion in Neurobiology* **23**: 581-587.

American Psychiatric Association. 2013. *The Diagnostic and Statistical Manual of Mental Disorders (5th ed.).* American Psychiatric Publishing, Arlington, VA.

Armbruster BN, Li X, Pausch MH, Herlitze S, Roth BL. 2007. Evolving the lock to fit the key to create a family of G protein-coupled receptors potently activated by an inert ligand. *Proceedings of the National Academy of Sciences* **104**: 5163-5168.

Anthony JC, Warner LA, Kessler RC. 1994. Comparative epidemiology of dependence on tobacco, alcohol, controlled substances, and inhalants: findings from the National Comorbidity Survey. *Experimental & Clinical Psychopharmacology* **166**: 244-268.

Ask H, Cheesman R, Jami ES, Levey DF, Purves KL, Weber H. 2021. Genetic contributions to anxiety disorders: where we are and where we are heading. *Psychological Medicine*: 1-16.

Aylward J, Hales C, Robinson E, Robinson OJ. 2020. Translating a rodent measure of negative bias into humans: the impact of induced anxiety and unmedicated mood and anxiety disorders. *Psychological Medicine* **50**: 237-246.

Babenko O, Kovalchuk I, Metz GAS. 2015. Stress-induced perinatal and transgenerational epigenetic programming of brain development and mental health. *Neuroscience and Biobehavioral Reviews* **48**: 70-91.

Banca P, Harrison NA, Voon V. 2016. Compulsivity across the pathological misuse of drug and non-drug rewards. *Frontiers in Behavioral Neuroscience* **10**: 154.

Banks ML, Negus SS. 2017. Insights from preclinical choice models on treating drug addiction. *Trends in Pharmacological Sciences* **38**: 181-194.

Barak S, Liu F, Ben Hamida S, Yowell QV, Neasta J, Kharazia V, Janak PH, Ron D. 2013. Disruption of alcohol-related memories by mTORC1 inhibition prevents relapse. *Nature Neuroscience* **16**: 1111-1117.

Belin D, Everitt BJ. 2008. Cocaine seeking habits depend upon dopamine-dependent serial connectivity linking the ventral with the dorsal striatum. *Neuron* **57**: 432-441.

Belin D, Mar AC, Dalley JW, Robbins TW, Everitt BJ. 2008. High impulsivity predicts the switch to compulsive cocaine-taking. *Science* **320**: 1352-1355.

Bell RL, Rodd ZA, Lumeng L, Murphy JM, McBride WJ. 2006. The alcohol-preferring P rat and animal models of excessive alcohol drinking. *Addiction Biology* **11**: 270-288.

Bender BN, Torregrossa MM. 2021. Dorsolateral striatum dopamine-dependent cocaine seeking is resistant to pavlovian cue extinction in male and female rats. **Neuropharmacology**: 108403.

Bernardi RE, Lattal KM, Berger SP. 2006. Postretrieval propranolol disrupts a cocaine conditioned place preference. *Neuroreport* **17**: 1443-1447.

Bernardi RE, Ryabinin AE, Berger SP, Lattal KM. 2009. Post-retrieval disruption of a cocaine conditioned place preference by systemic and intrabasolateral amygdala b_2 and a_1-adrenergic antagonists. *Learning & Memory* **16**: 777-789.

Blundell J, Kouser M, Powell CM. 2008. Systemic inhibition of mammalian target of rapamycin inhibits fear memory reconsolidation. *Neurobiology of Learning and Memory* **90**: 28-35.

Boyden E, Zhang F, Bamberg E, Nagel G, Deisseroth K. 2005. Millisecond-timescale, genetically targeted optical control of neural activity. *Nature Neuroscience* **8**: 1263-1268.

Caprioli D, Sawiak S, Merlo E, Theobald DE, Spoelder M, Jupp B, Voon V, Carpenter TA, Everitt BJ, Robbins TW et al. 2014. Gamma aminobutyric acidergic and neuronal structural markers in the nucleus accumbens core underlie trait-like impulsive behavior. *Biological Psychiatry* **75**: 115-123.

Cardinal RN, Parkinson JA, Hall J, Everitt BJ. 2002. Emotion and motivation: the role of the amygdala, ventral striatum, and prefrontal cortex. *Neuroscience & Biobehavioral Reviews* **26**: 321-352.

Casey BJ, Craddock N, Cuthbert BN, Hyman SE, Lee FS, Ressler KJ. 2013. DSM-5 and RDoC: progress in psychiatry research? *Nature Reviews Neuroscience* **14**: 810-814.

Childs E, de Wit H. 2009. Amphetamine-induced place preference in humans. *Biological Psychiatry* **65**: 900-904.

Cognigni P, Felsenberg J, Waddell S. 2018. Do the right thing: neural network mechanisms of memory formation, expression and update in *Drosophila. Current Opinion in Neurobiology* **49**: 51-58.

Colaizzi JM, Flagel SB, Joyner MA, Gearhardt AN, Stewart JL, Paulus MP. 2020. Mapping sign-tracking and goal-tracking onto human behaviors. *Neuroscience and Biobehavioral Reviews* **111**: 84-94.

Corbit LH, Janak PH. 2016. Changes in the influence of alcohol-paired stimuli on alcohol seeking across extended training. *Frontiers in Psychiatry* **7**: 169.

Corbit LH, Leung BK, Balleine BW. 2013. The role of the amygdala-striatal pathway in the acquisition and performance of goal-directed instrumental actions. *The Journal of Neuroscience* **33**: 17682-17690.

Dalley JW, Fryer TD, Brichard L, Robinson ESJ, Theobald DEH, Lääne K, Peña Y, Murphy ER, Shah Y, Probst K et al. 2007. Nucleus accumbens D2/3 receptors predict trait impulsivity and cocaine reinforcement. *Science* **315**: 1267-1270.

Daskalakis NP, Bagot RC, Parker KJ, Vinkers CH, de Kloet ER. 2013. The three-hit concept of vulnerability and resilience: toward understanding adaptation to early-life adversity outcome. *Psychoneuroendocrinology* **38**: 1858-1873.

Day JJ, Wheeler RA, Roitman MF, Carelli RM. 2006. Nucleus accumbens neurons encode pavlovian approach behaviors: evidence from an autoshaping paradigm. *European Journal of Neuroscience* **23**: 1341-1351.

Der-Avakian A, D'Souza MS, Pizzagalli DA, Markou A. 2013. Assessment of reward responsiveness in the response bias probabilistic reward task in rats: implications for cross-species translational research. *Translational Psychiatry* **3**: e297.

Deroche-Gamonet V, Belin D, Piazza PV. 2004. Evidence for addiction-like behavior in the rat. *Science* **305**: 1014-1017.

Di Ciano P, Everitt BJ. 2004. Conditioned reinforcing properties of stimuli paired with self-adminstered cocaine, heroin or sucrose: implications for the persistence of addictive behavior. *Neuropharmacology* **47**: 202-213.

Dias R, Robbins TW, Roberts AC. 1996a. Dissociation in prefrontal cortex of affective and attentional shifts. *Nature* **380**: 69-72.

Dias R, Robbins TW, Roberts AC. 1996b. Primate analogue of the Wisconsin card sorting test: effects of excitotoxic lesions of the prefrontal cortex in the marmoset. *Behavioral Neuroscience* **110**: 872-886.

Dias R, Robbins TW, Roberts AC. 1997. Dissociable forms of inhibitory control within prefrontal cortex with an analog of the Wisconsin card sort test: restriction to novel situations and independence from "on-line" processing. *The Journal of Neuroscience* **17**: 9285-9297.

Dickinson A. 1994. Instrumental conditioning. In *Animal Learning and Cognition*, (ed. NJ Mackintosh), pp. 45-79. Academic Press, Inc., London.

Ducci F, Goldman D. 2012. The genetic basis of addictive disorders. *Psychiatric Clinics of North America* **35**: 495-519.

Dwyer D. 2017. Crossing the worm-brain barrier by using *Caenorhabditis elegans* to explore fundamentals of human psychiatric illness. *Molecular Neuropsychiatry* **3**: 170-179.

Eagle DM, Noschang C, d'Angelo L-SC, Noble C, Day JO, Dongelmans ML, Theobald DE, Mar AC, Urcelay GP, Morein-Zamir S et al. 2014. The dopamine D2/D3 receptor agonist quinpirole increases checking-like behaviour in an operant observing response task with uncertain reinforcement: a novel possible model of OCD. *Behavioural Brain Research* **264**: 207-229.

Epstein DH. 2020. Let's agree to agree: a comment on Hogarth (2020), with a plea for not-so-competing theories of addiction. *Neuropsychopharmacology* **45**: 715-716.

Everitt BJ, Dickinson A, Robbins TW. 2001. The neuropsychological basis of addictive behaviour. *Brain Research Reviews* **36**: 129-138.

Everitt BJ, Parkinson JA, Olmstead MC, Arroyo M, Robledo P, Robbins TW. 1999. Associative processes in addiction and reward: the role of amygdala-ventral striatal subsystems. *Annals of the New York Academy of Sciences* **877**: 412-438.

Everitt BJ, Robbins TW. 2005. Neural systems of reinforcement for drug addiction: from actions to habits to compulsion. *Nature Neuroscience* **8**: 1481-1489.

Everitt BJ, Robbins TW. 2013. From the ventral to the dorsal striatum: devolving views of their roles in drug addiction. *Neuroscience and Biobehavioral Reviews* **37**: 1946-1954.

Everitt BJ, Robbins TW. 2016. Drug addiction: updating actions to habits to compulsions ten years on. *Annual Review of Psychology* **67**: 8.1-8.28.

Exton-McGuinness MTJ, Lee JLC. 2015. Reduction in responding for sucrose and cocaine reinforcement by disruption of memory reconsolidation. *eNeuro* **2**: e0009-0015.2015.

Fanelli RR, Klein JT, Reese RM, Robinson DL. 2013. Dorsomedial and dorsolateral striatum exhibit distinct phasic neuronal activity during alcohol self-administration in rats. *European Journal of Neuroscience* **38**: 2637-2648.

Figee M, Pattij T, Willuhn I, Luigjes J, van den Brink W, Goudriaan A, Potenza MN, Robbins TW, Denys D. 2016. Compulsivity in obsessive-compulsive disorder and addictions. *European Neuropsychopharmacology* **26**: 856-868.

Fischbach S, Janak PH. 2019. Decreases in cued reward seeking after reward-paired inhibition of mesolimbic dopamine. *Neuroscience* **412**: 259-269.

Fischbach-Weiss S, Reese RM, Janak PH. 2018. Inhibiting mesolimbic dopamine neurons reduces the initiation and maintenance of instrumental responding. *Neuroscience* **372**: 306-315.

Fisher SD, Ferguson LA, Bertran-Gonzalez J, Balleine BW. 2020. Amygdala-cortical control of striatal plasticity drives the acquisition of goal-directed action. *Current Biology* **30**: 4541-4546.

Flagel SB, Watson SJ, Robinson TE, Akil H. 2007. Individual differences in the propensity to approach signals vs goals promote different adaptations in the dopamine system of rats. *Psychopharmacology* **191**: 599-607.

Flores Á, Fullana MÀ, Soriano-Mas C, Andero R. 2018. Lost in translation: how to upgrade fear memory research. *Molecular Psychiatry*: doi: 10.1038/s34180-34017-30006-34180.

Furlong TM, Corbit LH, Brown RA, Balleine BW. 2018. Methamphetamine promotes habitual action and alters the density of striatal glutamate receptor and vesicular proteins in dorsal striatum. *Addiction Biology* **23**: 857-867.

Gillan CM, Fineberg NA, Robbins TW. 2017. A trans-diagnostic perspective on obsessive-compulsive disorder. *Psychological Medicine* **47**: 1528-1548.

Gilpin NW, Richardson HN, Cole M, Koob GF. 2008. Vapor inhalation of alcohol in rats. *Current Protocols in Neuroscience* **9.29.1**.

Giuliano C, Peña-Oliver Y, Goodlett CR, Cardinal RN, Robbins TW, Bullmore ET, Belin D, Everitt BJ. 2018. Evidence for a long-lasting compulsive alcohol seeking phenotype in rats. *Neuropsychopharmacology* **43**: 728-738.

Goldstein RZ, Volkow ND. 2011. Dysfunction of the prefrontal cortex in addiction: neuroimaging findings and clinical implications. *Nature Reviews Neuroscience* **12**: 652-669.

Grant DA, Berg E. 1948. A behavioral analysis of degree of reinforcement and ease of shifting to new responses in a Weigl-type card-sorting problem. *Journal of Experimental Psychology* **38**: 404-411.

Haber SN, Fudge JL, McFarland NR. 2000. Striatonigrostriatal pathways in primates form an ascending spiral from the shell to the dorsolateral striatum. *The Journal of Neuroscience* **20**: 2369-2382.

Hales CA, Robinson ESJ, Houghton CJ. 2016. Diffusion modelling reveals the decision making processes underlying negative judgement bias in rats. *PLoS One* **11**: e0152592.

Heilig M, Epstein DH, Nader MA, Shaham Y. 2016. Time to connect: bringing social context into addiction neuroscience. *Nature Reviews Neuroscience* **17**: 592-599.

Heinz A, Schlagenhauf F, Beck A, Wackerhagen C. 2016. Dimensional psychiatry: mental disorders as dysfunctions of basic learning mechanisms. *Journal of Neural Transmission* **123**: 809-821.

Hogarth L. 2020. Addiction is driven by excessive goal-directed drug choice under negative affect: translational critique of habit and compulsion theory. *Neuropsychopharmacology* **45**: 720-735.

Hogarth L, Chase HW. 2011. Parallel goal-directed and habitual control of human drug-seeking: implications for dependence vulnerability. *Journal of Experimental Psychology: Animal Behavior Processes* **37**: 261-276.

Holmes EA, Craske MG, Graybiel AM. 2014. A call for mental-health science. *Nature* **511**: 287-289.

Hvoslef-Eide M, Mar AC, Nilsson SRO, Alsiö J, Heath CJ, Saksida LM, Robbins TW, Bussey TJ. 2015. The NEWMEDS rodent touchscreen test battery for cognition relevant to schizophrenia. *Psychopharmacology* **232**: 3853-3872.

Hyman SE, Malenka RC, Nestler EJ. 2006. Neural mechanisms of addiction: the role of reward-related learning and memory. *Annual Review of Neuroscience* **29**: 565-598.

Ikemoto S. 2007. Dopamine reward circuitry: two projection systems from the ventral midbrain to the nucleus accumbens-olfactory tubercle complex. *Brain Research Reviews* **56**: 27-78.

Insel T, Cuthbert B, Garvey M, Heinssen R, Pine DS, Quinn K, Sanislow C, Wang P. 2010. Research domain criteria (RDoC): toward a new classification framework for research on mental disorders. *American Journal of Psychiatry* **167**: 748-751.

Ito R, Dalley JW, Robbins TW, Everitt BJ. 2002. Dopamine release in the dorsal striatum during cocaine-seeking behavior under the control of a drug-associated cue. *The Journal of Neuroscience* **22**: 6247-6253.

Jean-Richard-Dit-Bressel P, Killcross S, McNally GP. 2018. Behavioral and neurobiological mechanisms of punishment: implications for psychiatric disorders. *Neuropsychopharmacology* **43**: 1639-1650.

Jentsch JD, Taylor JR. 1999. Impulsivity resulting from frontostriatal dysfunction in drug abuse: implications for the control of behavior by reward-related stimuli. *Psychopharmacology* **146**: 373-390.

Kenny PJ, Chen SA, Kitamura O, Markou A, Koob GF. 2006. Conditioned withdrawal drives heroin consumption and decreases reward sensitivity. *The Journal of Neuroscience* **26**: 5894-5900.

Koob GF. 2021. Drug addiction: hyperkatifeia/negative reinforcement as a framework for medications development. *Pharmacological Reviews* **73**: 163-201.

Koob GF, Le Moal M. 2008. Addiction and the brain antireward system. *Annual Review of Psychology* **59**: 29-53.

LeBlanc KH, Maidment NT, Ostlund SB. 2013. Repeated cocaine exposure facilitates the expression of incentive motivation and induces habitual control in rats. *PLoS One* **8**: e61355.

Lee JLC, Everitt BJ. 2008. Reactivation-dependent amnesia in pavlovian approach and instrumental transfer. *Learning & Memory* **15**: 597-602.

Lichtenberg NT, Pennington ZT, Holley SM, Greenfield VY, Cepeda C, Levine MS, Wassum KM. 2017. Basolateral amygdala to orbitofrontal cortex projections enable cue-triggered reward expectations. *The Journal of Neuroscience* **37**: 8374-8384.

Logrip ML, Zorrilla EP, Koob GF. 2012. Stress modulation of drug self-administration: implications for addiction comorbidity with post-traumatic stress disorder. *Neuropharmacology* **62**: 552-564.

Luijten M, Gillan CM, de Wit S, Franken IHA, Robbins TW, Ersche KD. 2020. Goal-directed and habitual control in smokers. *Nicotine & Tobacco Research* **22**: 188-195.

Mackintosh NJ. 1974. *The Psychology of Animal Learning.* Academic Press, London.

Malvaez M, Shieh C, Murphy MD, Greenfield VY, Wassum KM. 2019. Distinct cortical-amygdala projections drive reward value encoding and retrieval. *Nature Neuroscience* **22**: 762-769.

Marie N, Canestrelli C, Noble F. 2019. Role of pharmacokinetic and pharmacodynamic parameters in neuroadaptations induced by drugs of abuse, with a focus on opioids and psychostimulants. *Neuroscience and Biobehavioral Reviews* **106**: 217-226.

McAuliffe WE. 1982. A test of Wikler's theory of relapse: the frequency of relapse due to conditioned withdrawal sickness. *The International Journal of the Addictions* **17**: 19-33.

Mellentin AI, Skøt L, Nielsen B, Schippers GM, Nielsen AS, Stenager E, Juhl C. 2017. Cue exposure therapy for the treatment of alcohol use disorders: a meta-analytic review. *Clinical Psychology Review* **57**: 195-207.

Miller KJ, Shenhav A, Ludvig EA. 2019. Habits without values. *Psychological Review* **126**: 292-311.

Milton AL, Everitt BJ. 2009. NMDA receptors and beta-adrenergic receptors as molecular targets for the prevention of relapse to drug-seeking. *European Neuropsychopharmacology* **19**: S86-S87.

Milton AL, Everitt BJ. 2010. The psychological and neurochemical mechanisms of drug memory reconsolidation: implications for the treatment of addiction. *European Journal of Neuroscience* **31**: 2308-2319.

Milton AL, Everitt BJ. 2012. The persistence of maladaptive memory: addiction, drug memories and anti-relapse treatments. *Neuroscience & Biobehavioral Reviews* **36**: 1119-1139.

Milton AL, Holmes EA. 2018. Of mice and mental health: facilitating dialogue and seeing further. *Philosophical Transactions of the Royal Society B* **373**: pii:; 20170022.

Milton AL, Lee JLC, Butler VJ, Gardner RJ, Everitt BJ. 2008a. Intra-amygdala and systemic antagonism of NMDA receptors prevents the reconsolidation of drug-associated memory and impairs subsequently both novel and previously acquired drug-seeking behaviors. *The Journal of Neuroscience* **28**: 8230-8237.

Milton AL, Lee JLC, Everitt BJ. 2008b. Reconsolidation of appetitive memories for both natural and drug reinforcement is dependent on β-adrenergic receptors. *Learning & Memory* **15**: 88-92.

Milton AL, Schramm MJW, Wawrzynski J, Gore F, Oikonomou-Mpegeti F, Wang NQ, Samuel D, Economidou D, Everitt BJ. 2012. Antagonism at NMDA receptors, but not β-adrenergic receptors, disrupts the reconsolidation of pavlovian conditioned approach and instrumental transfer for ethanol-associated conditioned stimuli. *Psychopharmacology* **219**: 751-761.

Mobbs D, Kim JJ. 2015. Neuroethological studies of fear, anxiety, and risky decision-making in rodents and humans. *Current Opinion in Behavioral Sciences* **5**: 8-15.

Mollick JA, Kober H. 2020. Computational models of drug use and addiction: a review. *Journal of Abnormal Psychology* **129**: 544-555.

Morein-Zamir S, Shahper S, Fineberg NA, Eisele V, Eagle DM, Urcelay G, Robbins TW. 2018. Free operant observing in humans: a translational approach to compulsive certainty seeking. *Quarterly Journal of Experimental Psychology* **71**: 2052-2069.

Moscarello JM, Hartley CA. 2017. Agency and the calibration of motivated behavior. *Trends in Cognitive Sciences*: doi: 10.1016/j.tics.2017.1006.1008.

Munafò MR, Davey Smith G. 2018. Robust research needs many lines of evidence. *Nature* **553**: 399-401.

Murray EA, Wise SP, Drevets WC. 2011. Localization of dysfunction in major depressive disorder: prefrontal cortex and amygdala. *Biological Psychiatry* **69**: e43-54.

Murray JE, Belin D, Everitt BJ. 2012. Double dissociation of the dorsomedial and dorsolateral striatal control over the acquisition and performance of cocaine seeking. *Neuropsychopharmacology* **37**: 2456-2466.

Murray JE, Belin-Rauscent A, Simon M, Giuliano C, Benoit-Marand M, Everitt BJ, Belin D. 2015. Basolateral and central amygdala differentially recruit and maintain dorsolateral striatum-dependent cocaine-seeking habits. *Nature Communications* **6**: 10088.

Narayanan AS, Rothenfluh A. 2016. I believe I can fly!: Use of *Drosophila* as a model organism in neuropsychopharmacology research. *Neuropsychopharmacology* **41**: 1439-1446.

Nelson A, Killcross S. 2006. Amphetamine exposure enhances habit formation. *The Journal of Neuroscience* **26**: 3805-3812.

O'Donnell KJ, Meaney MJ. 2017. Fetal origins of mental health: the developmental origins of health and disease hypothesis. *American Journal of Psychiatry* **174**: 319-328.

Oikonomidis L, Santangelo AM, Shiba Y, Clarke HF, Robbins TW, Roberts AC. 2017. A dimensional approach to modeling symptoms of neuropsychiatric disorders in the marmoset monkey. *Developmental Neurobiology* **77**: 328-353.

Olesen J, Gustavsson A, Svensson M, Wittchen HU, Jönsson B, group Cs, Council EB. 2012. The economic cost of brain disorders in Europe. *European Journal of Neurology* **19**: 155-162.

Open Science Collaboration. 2015. Estimating the reproducibility of psychological science. *Science* **349**: aac4716.

Parkinson JA, Olmstead MC, Burns LH, Robbins TW, Everitt BJ. 1999. Dissociation in effects of lesions of the nucleus accumbens core and shell on appetitive pavlovian approach behavior and the potentiation of conditioned reinforcement and locomotor activity by D-amphetamine. *The Journal of Neuroscience* **19**: 2401-2411.

Pascoli V, Hiver A, Van Zessen R, Loureiro M, Achargui R, Harada M, Flakowski J, Lüscher C. 2018. Stochastic synaptic plasticity underlying compulsion in a model of addiction. *Nature* **564**: 366-371.

Pearce JM, Hall G. 1980. A model for Pavlovian learning: variations in the effectiveness of conditioned but not of unconditioned stimuli. *Psychological Review* **87**: 532-552.

Pelloux Y, Everitt BJ, Dickinson A. 2007. Compulsive drug seeking by rats under punishment: effects of drug taking history. *Psychopharmacology* **194**: 127-137.

Pelloux Y, Murray JE, Everitt BJ. 2013. Differential roles of the prefrontal cortical subregions and basolateral amygdala in compulsive cocaine seeking and relapse after voluntary abstinence in rats. *European Journal of Neuroscience* **38**: 3018-3026.

Peters SM, Pothuizen HHJ, Spruijt BM. 2015. Ethological concepts enhance the translational value of animal models. *European Journal of Pharmacology* **759**: 42-50.

Pizzagalli DA, Iosifescu D, Hallett LA, Ratner KG, Fava M. 2008. Reduced hedonic capacity in major depressive disorder: evidence from a probabilistic reward task. *Journal of Psychiatric Research* **43**: 76-87.

Puaud M, Higuera-Matas A, Brunault P, Everitt BJ, Belin D. 2021. The basolateral amygdala to nucleus accumbens core circuit mediates the conditioned reinforcing effects of cocaine-paired cues on cocaine seeking. *Biological Psychiatry* **89**: 356-365.

Rescorla RA, Wagner AR. 1972. A theory of Pavlovian conditioning: variations in the effectiveness of reinforcement and nonreinforcement. In *Classical Conditioning II: Current Research and Theory*, (ed. AH Black, WF Prokasy), pp. 64-99. Appleton-Century-Crofts, New York.

Richter-Levin G, Stork O, Schmidt MV. 2019. Animal models of PTSD: a challenge to be met. *Molecular Psychiatry* **24**.

Ritchie H, Roser M. 2018. Mental health., https://ourworldindata.org/mental-health.

Robbins TW, Cardinal RN. 2019. Computational psychopharmacology: a translational and pragmatic approach. *Psychopharmacology* **236**: 2295-2305.

Robbins TW, Gillan CM, Smith DG, de Wit S, Ersche KD. 2012. Neurocognitive endophenotypes of impulsivity and compulsivity: towards dimensional psychiatry. *Trends in Cognitive Sciences* **16**: 81-91.

Roberts AC, Robbins TW, Everitt BJ. 1988. The effects of intradimensional and extradimensional shifts on visual discrimination learning in humans and non-human primates. *The Quarterly Journal of Experimental Psychology* **40**: 321-341.

Robinson ESJ. 2018. Translational new approaches for investigating mood disorders in rodents and what they may reveal about the underlying neurobiology of major depressive disorder. *Philosophical Transactions of the Royal Society B*: doi: 10.1098/rstb.2017.0036.

Robinson TE, Berridge KC. 2000. The psychology and neurobiology of addiction: an incentive-sensitization view. *Addiction* **95**: 91-117.

Saddoris MP, Stamatakis A, Carelli RM. 2011. Neural correlates of Pavlovian-to-instrumental transfer in the nucleus accumbens shell are selectively potentiated following cocaine self-administration. *European Journal of Neuroscience* **33**: 2274-2287.

Sahakian BJ, Morris RG, Evenden JL, Heald A, Levy R, Philpot M, Robbins TW. 1988. A comparative study of visuospatial memory and learning in Alzheimer-type dementia and Parkinson's Disease. *Brain* **111**: 695-718.

Saunders BT, Richard JM, Margolis EB, Janak PH. 2018. Dopamine neurons create pavlovian conditioned stimuli with circuit-defined motivational properties. *Nature Neuroscience* **21**: 1072-1083.

Schlotz W, Phillips DIW. 2009. Fetal origins of mental health: evidence and mechanisms. *Brain, Behavior, and Immunity* **23**: 905-916.

Schramm MJW, Everitt BJ, Milton AL. 2016. Bidirectional modulation of alcohol-associated memory reconsolidation through manipulation of adrenergic signaling. *Neuropsychopharmacology* **41**: 1103-1111.

Schultz W, Apicella P, Ljungberg T. 1993. Responses of monkey dopamine neurons to reward and conditioned stimuli during successive steps of learning a delayed response task. *The Journal of Neuroscience* **13**: 900-913.

Shrout PE, Rodgers JL. 2018. Psychology, science, and knowledge construction: broadening perspectives from the replication crisis. *Annual Review of Psychology* **69**: 487-510.

Siegel S. 1977. Morphine tolerance acquisition as an associative process. *Journal of Experimental Psychology: Animal Behavior Processes* **3**: 1-13.

Sinha R, Shaham Y, Heilig M. 2011. Translational and reverse translational research on the role of stress in drug craving and relapse. *Psychopharmacology* **218**: 69-82.

Spanagel R, Noori HR, Heilig M. 2014. Stress and alcohol interactions: animal studies and clinical significance. *Trends in Neurosciences* **37**: 219-227.

Stam R. 2007. PTSD and stress sensitisation: a tale of brain and body. Part 2: animal models. *Neuroscience and Biobehavioral Reviews* **31**: 558-584.

Stuart SA, Butler P, Munafò MR, Nutt DJ, Robinson ESJ. 2013. A translational rodent assay of affective biases in depression and antidepressant therapy. *Neuropsychopharmacology* **38**: 1625-1635.

Stuart SA, Butler P, Munafò MR, Nutt DJ, Robinson ESJ. 2015. Distinct neuropsychological mechanisms may explain delayed- versus rapid-onset antidepressant efficacy. *Neuropsychopharmacology* **40**: 2165-2174.

Sutton RS, Barto AG. 1981. Toward a modern theory of adaptive networks: expectation and prediction. *Psychological Review* **88**: 135-170.

Szechtman H, Ahmari SE, Beninger RJ, Eilam D, Harvey BH, Edemann-Callesen H, Winter C. 2017. Obsessive-compulsive disorder: insights from animal models. *Neuroscience & Biobehavioral Reviews* **76**: 254-279.

Talmi D, Seymour B, Dayan P, Dolan RJ. 2008. Human pavlovian-instrumental transfer. *The Journal of Neuroscience* **28**: 360-368.

Tomie A, Brooks W, Zito B, Klein SB, Mowrer RR. 1989. Sign-tracking: the search for reward. In *Contemporary Learning Theories: Pavlovian Conditioning and the Status of Traditional Learning Theory*, pp. 191 -223. Lawrence Erlbaum Associates, Hillsdale, NJ.

Torregrossa MM, Corlett PR, Taylor JR. 2011. Aberrant learning and memory in addiction. *Neurobiology of Learning and Memory* **96**: 609-623.

Tronson NC, Taylor JR. 2013. Addiction: a drug-induced disorder of memory reconsolidation. *Current Opinion in Neurobiology* **23**: 573-580.

Vandaele Y, Ahmed SH. 2020. Habit, choice, and addiction. *Neuropsychopharmacology* **46**: 689-698.

Vandaele Y, Mahajan NR, Ottenheimer DJ, Richard JM, Mysore SP, Janak PH. 2019. Distinct recruitment of dorsomedial and dorsolateral striatum erodes with extended training. *eLife* **8**: e49536.

Vanderschuren LJMJ, Di Ciano P, Everitt BJ. 2005. Involvement of the dorsal striatum in cue-controlled cocaine seeking. *The Journal of Neuroscience* **25**: 8665-8670.

Vanderschuren LJMJ, Everitt BJ. 2004. Drug seeking becomes compulsive after prolonged cocaine self-adminstration. *Science* **305**: 1017-1019.

Venniro M, Banks ML, Heilig M, Epstein DH, Shaham Y. 2020. Improving translation of animal models of addiction and relapse by reverse translation. *Nature Reviews Neuroscience* **21**: 625-643.

Voon V, Dalley JW. 2016. Translatable and back-translatable measurement of impulsivity and compulsivity: convergent and divergent processes. *Current Topics in Behavioral Neurosciences* **28**: 53-91.

Wardle MC, Lopez-Gamundi P, Flagel SB. 2018. Measuring appetitive conditioned responses in humans. *Physiology & Behavior* **188**: 140-150.

Wassum KM, Izquierdo A. 2015. The basolateral amygdala in reward learning and addiction. *Neuroscience and Biobehavioral Reviews* **57**: 271-283.

Wikler A. 1948. Recent progress in research on the neurophysiologic basis of morphine addiction. *American Journal of Psychiatry* **105**: 329-338.

Wise RA, Bozarth MA. 1987. A psychomotor stimulant theory of addiction. *Psychological Review* **94**: 469-492.

Wittchen HU, Jacobi F, Rehm J, Gustavsson A, Svensson M, Jönsson B, Oleson J, Allgulander C, Alonso J, Faravelli C et al. 2011. The size and burden of mental disorders and other disorders of the brain in Europe 2010. *European Neuropsychopharmacology* **21**: 655-679.

Yehuda R, LeDoux JE. 2007. Response variation following trauma: a translational neuroscience approach to understanding PTSD. In *Neuron*, Vol 56, pp. 19-32.

Yin HH, Knowlton BJ, Balleine BW. 2004. Lesions of dorsolateral striatum preserve outcome expectancy but disrupt habit formation in instrumental learning. *European Journal of Neuroscience* **19**: 181-189.

Yin HH, Knowlton BJ, Balleine BW. 2005. Blockade of NMDA receptors in the dorsomedial striatum prevents action-outcome learning in instrumental conditioning. In *European Journal of Neuroscience*, Vol 22, pp. 505-512.

Zapata A, Minney VL, Shippenberg TS. 2010. Shift from goal-directed to habitual cocaine seeking after prolonged experience in rats. *The Journal of Neuroscience* **30**: 15457-15463.

| Term	Definition
Anhedonia	The inability to feel pleasure.
Categorical approach (to mental health disorders)	An approach to the study of mental health disorders in which the emphasis is put on the specific groups of symptoms required to warrant a particular diagnosis (e.g. as in the DSM-5).
Compulsive behaviour	Behaviour that is repeated persistently, despite no benefit (or even the production of adverse outcomes) for the individual.
Consummatory response	The final act of a chain of behaviours aimed at procuring a goal (e.g. eating, drinking, copulation).

Cue-induced reinstatement procedure	Behavioural procedure in which animals are trained to self-administer reward (e.g. drugs of abuse) in the presence of cues associated with the reward in a pavlovian manner. Responding in the absence of primary reward but in the presence of the cue is assessed at test after a manipulation (e.g. after extinction training, or enforced abstinence) as a measure of relapse-like behaviour.
Dimensional approach (to mental health disorders)	An approach to the study of mental health disorders in which the emphasis is put on underlying psychological processes that are dysfunctional, that may be dysfunctional across different diagnoses, and may co-occur with other dysfunctional psychological processes to give rise to different mental health disorders.
Drug-seeking behaviour	The instrumental act of working to procure drugs, which is dissociable from the consummatory act of consuming the drug.
Drug-taking behaviour	The consummatory act of self-administering the drug. For situations in which there is no clear consummatory behaviour (e.g. self-administration via the intravenous route) this may be represented as the final act in an instrumental chain leading to drug delivery, as in the 'seeking-taking' task.
Dysphoria	The state of intense generalised unhappiness.
Goal-directed behaviour	Behaviour that is elicited for the procurement of a particular outcome, supported by 'action-outcome' associations. Goal-directed behaviour is sensitive to changes in the value of the outcome or the contingency between the action and the outcome.
Habitual behaviour	Behaviour that is elicited automatically by stimuli in the environment, supported by 'stimulus-response' associations. Habitual behaviour is insensitive to changes in the value of the outcome and degradation of the contingency between the action and the outcome.
Impulsive behaviour	Acting without foresight or thinking.
Instrumental (association)	Learned behaviour in which an action is performed, at least initially due to perceived contingencies between the action and a particular outcome.
Interval schedule of reinforcement	A reinforcement schedule in which the first response after a specific time interval is reinforced. Interval schedules can bias towards habitual behaviour, as individuals often elicit greater responding than necessary for reward, thereby degrading the correlation between the action and reinforcement.
Negative (instrumental) reinforcement	An increase in the frequency or occurrence of instrumental behaviour in order to avoid an unwanted outcome (e.g. an increase in drug self-administration to avoid drug withdrawal).
Pavlovian (association)	Learned behaviour in which cues are associated with an outcome in a manner that is independent of the behaviour of the individual.
Pavlovian conditioned approach	The capacity of a pavlovian cue to support approach towards itself (also known as 'sign-tracking') or towards the location of reward delivery (also known as 'goal-tracking') when the cue is presented.
Pavlovian conditioned motivation	The capacity of a pavlovian cue to support increased instrumental responding when the cue is associated with the same outcome ('specific pavlovian-instrumental transfer') or an outcome of the same valence ('general pavlovian-instrumental transfer').
Pavlovian conditioned reinforcement	The capacity of a pavlovian cue to support instrumental behaviour in its own right and in the absence of primary reinforcement. Conditioned reinforcers can support instrumental behaviour over delays to primary reinforcement, and are persistent and resistant to extinction.

Pavlovian conditioned stimulus	An initially neutral cue which has been associated with an outcome through learning.
Pavlovian conditioned tolerance	A learned change in the degree of tolerance to a drug in the presence of pavlovian conditioned stimuli.
Pavlovian conditioned withdrawal	A state of withdrawal induced by the presentation of learned pavlovian conditioned stimuli that have previously been present during withdrawal.
Perseveration	Persistent, repetitive behaviour elicited beyond an appropriate point.
Punishment	A reduction in the frequency of occurrence instrumental responding produced by a negative outcome (e.g. a reduction in drug-seeking behaviour due to seeking responses being shocked).
Reconsolidation	The process by which memories can be updated under certain conditions, and require plasticity-related processes in order to persist in the brain in an updated form.
Self-administration procedure	Procedure in which an individual can administer drugs without the intervention of an experimenter. This can be oral (e.g. for alcohol) or may require the implantation of an intravenous line (e.g. for drugs such as cocaine and heroin).

**Table 1. Glossary of psychological terms.**

FIGURE LEGEND

Figure 1. The psychological approach to mental health disorders. Mental health disorders can be deconstructed into component psychological processes that have become dysfunctional or maladaptive, with some maladaptive processes being common across different disorders, and specific disorders comprising different maladaptive processes. The development or use of existing behavioural tasks that isolate the relevant psychological processes – particularly when designed to be language-independent and objective – facilitates comparison between humans, animals and computational models. The mechanistic insights provided by the data generated can then be used alongside research into other maladaptive psychological processes to 'reconstruct' the more complex mental health disorder.
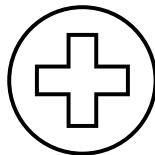
Mental health disorder

Reconstruction of disorder

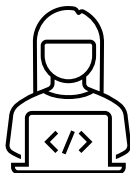Deconstruction into psychological processes
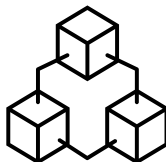
Mechanistic insight

Development/use of tasks to measure/quantify psychological processes

Human data

Computational data

Animal data

Translation / backtranslation