



REWARD PROCESSING IN MOTIVATIONAL AND AFFECTIVE DISORDERS

EDITED BY : Frank Ryan and Nikolina Skandali
PUBLISHED IN : Frontiers in Psychology



frontiers

Frontiers Copyright Statement

© Copyright 2007-2016 Frontiers Media SA. All rights reserved.

All content included on this site, such as text, graphics, logos, button icons, images, video/audio clips, downloads, data compilations and software, is the property of or is licensed to Frontiers Media SA ("Frontiers") or its licensees and/or subcontractors. The copyright in the text of individual articles is the property of their respective authors, subject to a license granted to Frontiers.

The compilation of articles constituting this e-book, wherever published, as well as the compilation of all other content on this site, is the exclusive property of Frontiers. For the conditions for downloading and copying of e-books from Frontiers' website, please see the Terms for Website Use. If purchasing Frontiers e-books from other websites or sources, the conditions of the website concerned apply.

Images and graphics not forming part of user-contributed materials may not be downloaded or copied without permission.

Individual articles may be downloaded and reproduced in accordance with the principles of the CC-BY licence subject to any copyright or other notices. They may not be re-sold as an e-book.

As author or other contributor you grant a CC-BY licence to others to reproduce your articles, including any graphics and third-party materials supplied by you, in accordance with the Conditions for Website Use and subject to any copyright notices which you include in connection with your articles and materials.

All copyright, and all rights therein, are protected by national and international copyright laws.

The above represents a summary only. For the full conditions see the Conditions for Authors and the Conditions for Website Use.

ISSN 1664-8714

ISBN 978-2-88919-986-0

DOI 10.3389/978-2-88919-986-0

About Frontiers

Frontiers is more than just an open-access publisher of scholarly articles: it is a pioneering approach to the world of academia, radically improving the way scholarly research is managed. The grand vision of Frontiers is a world where all people have an equal opportunity to seek, share and generate knowledge. Frontiers provides immediate and permanent online open access to all its publications, but this alone is not enough to realize our grand goals.

Frontiers Journal Series

The Frontiers Journal Series is a multi-tier and interdisciplinary set of open-access, online journals, promising a paradigm shift from the current review, selection and dissemination processes in academic publishing. All Frontiers journals are driven by researchers for researchers; therefore, they constitute a service to the scholarly community. At the same time, the Frontiers Journal Series operates on a revolutionary invention, the tiered publishing system, initially addressing specific communities of scholars, and gradually climbing up to broader public understanding, thus serving the interests of the lay society, too.

Dedication to Quality

Each Frontiers article is a landmark of the highest quality, thanks to genuinely collaborative interactions between authors and review editors, who include some of the world's best academicians. Research must be certified by peers before entering a stream of knowledge that may eventually reach the public - and shape society; therefore, Frontiers only applies the most rigorous and unbiased reviews.

Frontiers revolutionizes research publishing by freely delivering the most outstanding research, evaluated with no bias from both the academic and social point of view.

By applying the most advanced information technologies, Frontiers is catapulting scholarly publishing into a new generation.

What are Frontiers Research Topics?

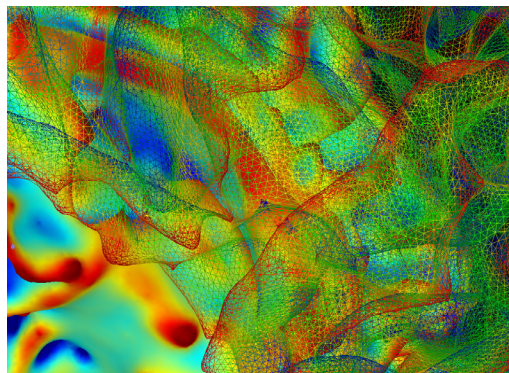
Frontiers Research Topics are very popular trademarks of the Frontiers Journals Series: they are collections of at least ten articles, all centered on a particular subject. With their unique mix of varied contributions from Original Research to Review Articles, Frontiers Research Topics unify the most influential researchers, the latest key findings and historical advances in a hot research area! Find out more on how to host your own Frontiers Research Topic or contribute to one as an author by contacting the Frontiers Editorial Office: researchtopics@frontiersin.org

REWARD PROCESSING IN MOTIVATIONAL AND AFFECTIVE DISORDERS

Topic Editors:

Frank Ryan, Imperial College, UK

Nikolina Skandali, University of Cambridge, UK



SUMA surface mapping of brain image

Image: National Institute of Mental Health, National Institutes of Health, Department of Health and Human Services

Preferential reward processing is the hallmark of addiction, where salient cues become overvalued and trigger compulsion. In depression, rewards appear to lose their incentive properties or become devalued. In the context of schizophrenia, aberrations in neural reward signalling are thought to contribute to the overvaluation of irrelevant stimuli on the one hand and the onset of negative symptoms on the other. Accordingly, reward processing has emerged as a key variable in contemporary, evidence based, diagnostic frameworks, such as the Research Domain Criteria launched by the United States National Institute of Mental Health. Delineation of the underlying mechanisms of aberrant or

blunted reward processing can be of trans-diagnostic importance across several neuropsychiatric disorders. Reward processing can become automatic thus raising the question of cognitive control, a core theme of this Topic, which aims at justifying the necessity of reward processing as a potential therapeutic target in clinical settings. Empirical and theoretical contributions on the following themes were expected to:

- * Explore new avenues of research by investigating the processing of rewards at the cognitive, behavioural, motivational, neural systems and individual difference levels. A developmental focus is promising in this regard, probing the core processes that shape reward processing and thus subsequent liability to motivational and affective disorders.
- * Develop and refine conceptual models of reward processing from computational neuroscience.
- * Promote greater understanding and development of emergent therapeutic approaches such as cognitive bias modification and behavioural approach or avoidance training. A key question is the feasibility of reversing or modifying maladaptive patterns of reward processing to therapeutic ends.

- * Refine and augment the evidential database for tried and tested therapies such as Contingency Management and Behavioural Activation by focusing on core cognitive processes mediating rewards.
- * Provide a potential dimensional approach for reward processing deficits that can be of trans-diagnostic importance in clinically relevant disorders, including depression and addiction
- * Investigate the subjective experience of pleasure- the hedonic aspect of reward seeking and consumption – and how this can be distinguished from the motivational, sometimes compulsive, component of reward pursuit. This promises more nuanced and effective interventions. Depression, for instance, could be seen as the restricted pursuit of pleasure rather than blunted pleasure experience; addiction can be viewed as accentuated drug seeking despite diminished consummatory pleasure. This aims to place motivation centre stage in both scenarios, emphasising the transdiagnostic theme of the Topic.
- * Temporal discounting of future rewards, whereby smaller, more immediate rewards are chosen even when significantly more valuable deferred rewards are available, is another trans-diagnostic phenomenon of interest in the in the present context. Factors that influence this, such as discounting of future reward are thought to reflect compulsion in the addictive context and hopelessness on the part of people experiencing depression. The executive cognitive processes that regulate this decision making are of both scientific and clinical significance. Empirical findings, theoretical contributions or commentaries bearing on cognitive or executive control were therefore welcome.

Citation: Ryan, F., Skandali, N., eds. (2016). Reward Processing in Motivational and Affective Disorders. Lausanne: Frontiers Media. doi: 10.3389/978-2-88919-986-0

Table of Contents

- 05 Editorial: Reward Processing in Motivational and Affective Disorders**
Frank Ryan and Nikolina Skandali
- 07 Measuring anhedonia: impaired ability to pursue, experience, and learn about reward**
Kristine Rømer Thomsen
- 18 Unifying treatments for depression: an application of the Free Energy Principle**
Adam M. Chekroud
- 26 The neuroscience of positive memory deficits in depression**
Daniel G. Dillon
- 38 Reduction in ventral striatal activity when anticipating a reward in depression and schizophrenia: a replicated cross-diagnostic finding**
Gonzalo Arrondo, Nuria Segarra, Antonio Metastasio, Hisham Ziauddeen, Jennifer Spencer, Niels R. Reinders, Robert B. Dudas, Trevor W. Robbins, Paul C. Fletcher and Graham K. Murray
- 50 The impact of threat of shock on the framing effect and temporal discounting: executive functions unperturbed by acute stress?**
Oliver J. Robinson, Rebecca L. Bond and Jonathan P. Roiser
- 59 Assessment of Tobacco-Related Approach and Attentional Biases in Smokers, Cravers, Ex-Smokers, and Non-Smokers**
Marcella L. Woud, Joyce Maas, Reinout W. Wiers, Eni S. Becker and Mike Rinck
- 70 Priming of conflicting motivational orientations in heavy drinkers: robust effects on self-report but not implicit measures**
Lisa C. G. Di Lemma, Joanne M. Dickson, Pawel Jedras, Anne Roefs and Matt Field
- 86 Embracing comorbidity: a way toward understanding the role of motivational and control processes in cannabis use disorders**
Janna Cousijn
- 91 The computational psychiatry of reward: broken brains or misguided minds?**
M. Moutoussis, G. W. Story and R. J. Dolan
- 100 A Computational Analysis of Aberrant Delay Discounting in Psychiatric Disorders**
Giles W. Story, Michael Moutoussis and Raymond J. Dolan



Editorial: Reward Processing in Motivational and Affective Disorders

Frank Ryan^{1*} and Nikolina Skandali²

¹ Centre for Mental Health, Division of Brain Sciences, Imperial College London, London, UK, ² Department of Psychiatry, Behavioural and Clinical Neuroscience Institute, University of Cambridge, Cambridge, UK

Keywords: reward, anhedonia, depression, schizophrenia, addiction, computational neuroscience, Bayesian models

The Editorial on the Research Topic

Reward Processing in Motivational and Affective Disorders

Reward prediction and valuation are central to decision making (Schultz et al., 1997; O'Doherty, 2004), and thus motivate and guide human action. Faulty reward processing can be pragmatically viewed as compromised decision-making, reflected in making suboptimal choices. A primary aim of this Research Topic is to provide a dimensional approach for reward processing. In accordance with the Research Domain Criteria, a newly proposed research classification of mental health disorders based on behavioral dimensions and neurobiological findings (Insel et al., 2010), reward processing deficits can be of transdiagnostic importance. This prominence has been largely influenced by the evolution of novel neuroimaging techniques, experimental cognitive psychology findings, and the application of computational modeling in simulating human behavior. This body of work has increased knowledge of the neural mechanisms underlying aberrant reward processing.

We maintain that construing patterns or expressions of reward processing as potential biomarkers, or indices of psychological vulnerability, can facilitate early detection and intervention in the clinical arena. Additionally, therapeutic approaches originating in the psychology laboratory aimed at modifying or reversing cognitive biases or behavioral approach biases linked to aberrant reward processes are showing promise in preventing relapse in the context of addiction (see Gladwin et al., 2016). It is this twin promise of enhanced prediction of vulnerability or risk and ultimately improved clinical outcomes, combined with a deeper understanding of brain functions, that motivated us to gather together this unique series of articles linked by the common thread of reward processing.

The *Topic* includes four original articles exploring reward processing in schizophrenia, depression, addiction and in the context of stress or anxiety. These empirical contributions are complemented by three review articles, two theoretical contributions and an opinion piece. In tandem with laboratory findings, these more conceptual articles put emphasis on the role of the integrity of neuromodulatory systems implicated in reward processing as well as the remarkable insights that can be derived from the implementation of computational modeling.

Rømer Thomsen, set the scene by outlining the subcomponents of reward processing: wanting, liking and learning. This parsing of reward processing enables a critical analysis of the concept of anhedonia, suggesting that deficits in reward processing are not restricted to “liking” (the subjective, pleasurable, experience of being rewarded). Other components, “wanting” and mechanisms underlying learning about rewards, can also be disrupted and contribute to the development and maintenance of disorders such as addiction and depression.

OPEN ACCESS

Edited and reviewed by:

Gianluca Castelnuovo,
Catholic University of the Sacred
Heart, Italy

*Correspondence:

Frank Ryan
f.ryan@imperial.ac.uk

Specialty section:

This article was submitted to
Psychology for Clinical Settings,
a section of the journal
Frontiers in Psychology

Received: 18 March 2016

Accepted: 12 August 2016

Published: 30 August 2016

Citation:

Ryan F and Skandali N (2016)
Editorial: Reward Processing in
Motivational and Affective Disorders.
Front. Psychol. 7:1288.
doi: 10.3389/fpsyg.2016.01288

Arrondo et al. demonstrated blunted reward anticipation in people with a diagnosis of schizophrenia and depression. This attenuated striatal response to the prospect of monetary reward correlated with depressive symptoms in the schizophrenia group, but did not cohere with clinical symptoms of depression in the depressed cohort. Di Lemma et al. and colleagues found that, contrary to predictions, approach and avoidance tendencies following positive and negatively themed videos *did* change in parallel challenging the expectation that these are independent processes. Woud et al. investigated broadly similar processes investigating cohorts of current or former tobacco smokers. The researchers reported no attentional or behavioral approach biases in either current smokers, nicotine deprived smokers and ex-smokers, thus challenging incentive theories of addiction. These three innovative studies raise important questions for further research, and refine experimental methods in the process.

Robinson et al. applied a stress manipulation paradigm in order to study the effect of acute stress on two well-established biases in decision-making, temporal discounting and the framing effect. The researchers observed mood alterations in response to experimentally induced stress, but no effects on decision-making processes. Acute stress impacted on low level “bottom-up” perceptual biases, but higher level executive processes were unperturbed. The findings support the application of psychotherapeutic approaches aiming to enhance cognitive control as an apparently resilient component of therapeutic intervention for affective disorders.

Chekroud discussed the distortions in reward sensitivity and/or reward learning that contribute to the development of depression. He described the implementation of the free-energy principle, which views the brain as a “predictive machine,” aimed at reducing surprise (i.e., free energy) by constructing congruent cognitive models and optimizing actions. One potential clinical application is that changing cognitive representations using pharmacological agents or psychotherapy will be a necessarily gradual rather than immediate process. Also on the topic of depression, Dillon assigned a pivotal role to reward processing in the formation of long term memories. He concluded that impaired reward processing reflected distorted mesolimbic dopaminergic transmission, thus impeding the transfer of short-term memories into long-term episodic memory storage. Consequentially, in order to recover from depression, not only

will somebody who is depressed need to overcome a memory bias for the recall of negative events, they will also struggle to recall positive events.

Cousijn highlighted the role of fronto-parietal and limbic brain networks that are implicated in vulnerability to cannabis use disorders, other substance misuse disorders and increased risk of anxiety and depression. When these cognitive control systems are compromised individuals are more likely to reach out for immediately available rewards whether they are linked to substance use or the powerful negative reinforcement that occurs when emotional distress is alleviated by avoidance, thus increasing the possibility of anhedonia and depression.

Moutoussis et al. differentiated between optimal decisions delivering the best possible rewards and a conceptualization of psychiatric disorder based on suboptimal reward processing. In this context, rewards are milestones or surrogates to strategic goals such as health, wellbeing and social affiliation. The researchers emphasized the importance of considering the patient's autonomy by pointing out the need for the clinician to engage in a dialog to elicit the patient's values and goals. Story et al. identified two factors involved in delayed reward discounting in psychiatric disorders: the opportunity cost that waiting for a delayed reward entails and the associated uncertainty of reward delivery.

The theoretical insights and experimental findings presented in this topic justify further exploration of the mechanisms underlying the anticipation, valuation and pursuit of rewards. In tandem with this, adapting and applying these findings in clinical settings could, we believe, provide additional therapeutic benefit.

AUTHOR CONTRIBUTIONS

All authors listed, have made substantial, direct and intellectual contribution to the work, and approved it for publication.

FUNDING

Dr. NS is supported by a UK Medical Research Council Doctoral grant studentship. Dr. FR wishes to acknowledge the support and encouragement he received from his employer Camden & Islington NHS Foundation Mental Health Trust and the Centre for Mental Health, Imperial College.

REFERENCES

- Gladwin, T. E., Wiers, C. E., and Wiers, R. W. (2016). Cognitive neuroscience of cognitive retraining for addiction medicine: from mediating mechanisms to questions of efficacy. *Prog. Brain Res.* 224, 323–344. doi: 10.1016/bs.pbr.2015.07.021
- Insel, T., Cuthbert, B., Garvey, M., Heinssen, R., Pine, D. S., Quinn, K., et al. (2010). Research domain criteria (RDoC): toward a new classification framework for research on mental disorders. *Am. J. Psychiatry* 167, 748–751. doi: 10.1176/appi.ajp.2010.09091379
- O'Doherty, J. P. (2004). Reward representations and reward-related learning in the human brain: insights from neuroimaging. *Curr. Opin. Neurobiol.* 14, 769–776. doi: 10.1016/j.conb.2004.10.016
- Schultz, W., Dayan, P., and Montague, P. R. (1997). A neural substrate of prediction and reward. *Science* 275, 1593–1599.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2016 Ryan and Skandali. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Measuring anhedonia: impaired ability to pursue, experience, and learn about reward

Kristine Rømer Thomsen*

Centre for Alcohol and Drug Research, Department of Psychology and Behavioural Sciences, Aarhus University, Aarhus C, Denmark

OPEN ACCESS

Edited by:

Frank Ryan,
Imperial College, UK

Reviewed by:

Richard J. Tunney,
University of Nottingham, UK

Mike J. F. Robinson,
Wesleyan University, USA

*Correspondence:

Kristine Rømer Thomsen,
Centre for Alcohol and Drug
Research, Department of Psychology
and Behavioural Sciences, Aarhus
University, Bartholins Allé 10,
8000 Aarhus C, Denmark
krt.crf@psy.au.dk

Specialty section:

This article was submitted to
Psychology for Clinical Settings,
a section of the journal
Frontiers in Psychology

Received: 29 May 2015

Accepted: 03 September 2015

Published: 17 September 2015

Citation:

Rømer Thomsen K (2015) Measuring
anhedonia: impaired ability to pursue,
experience, and learn about reward.
Front. Psychol. 6:1409.
doi: 10.3389/fpsyg.2015.01409

Ribot's (1896) long standing definition of anhedonia as "the inability to experience pleasure" has been challenged recently following progress in affective neuroscience. In particular, accumulating evidence suggests that reward consists of multiple subcomponents of wanting, liking and learning, as initially outlined by Berridge and Robinson (2003), and these processes have been proposed to relate to appetitive, consummatory and satiety phases of a pleasure cycle. Building on this work, we recently proposed to reconceptualize anhedonia as "impairments in the ability to pursue, experience, and/or learn about pleasure, which is often, but not always accessible to conscious awareness." (Rømer Thomsen et al., 2015). This framework is in line with Treadway and Zald's (2011) proposal to differentiate between motivational and consummatory types of anhedonia, and stresses the need to combine traditional self-report measures with behavioral measures or procedures. In time, this approach may lead to improved clinical assessment and treatment. In line with our reconceptualization, increasing evidence suggests that reward processing deficits are not restricted to impaired hedonic impact in major psychiatric disorders. Successful translations of animal models have led to strong evidence of impairments in the ability to pursue and learn about reward in psychiatric disorders such as major depressive disorder, schizophrenia, and addiction. It is of high importance that we continue to systematically target impairments in all phases of reward processing across disorders using behavioral testing in combination with neuroimaging techniques. This in turn has implications for diagnosis and treatment, and is essential for the purposes of identifying the underlying neurobiological mechanisms. Here I review recent progress in the development and application of behavioral procedures that measure subcomponents of anhedonia across relevant patient groups, and discuss methodological caveats as well as implications for assessment and treatment.

Keywords: anhedonia, reward, pleasure, motivation, learning, depression, schizophrenia, addiction

Introduction

The generally accepted understanding of the term anhedonia has remained almost unaltered since Ribot (1896) first defined it as the "inability to experience pleasure" over a century ago. However, during the last 5 years the term has been subject to debate and some progress has been made in terms of elucidating the underlying neurobiological mechanisms. A number of recent reviews (Treadway and Zald, 2011; Der-Avakian and Markou, 2012; Whitton et al., 2015), summarize

this progress and offer improved understanding of the underlying neurobiology. However, their conceptual understanding of anhedonia diverge. Treadway and Zald (2011) made a convincing case to differentiate between motivational and consummatory types of anhedonia and introduced the term decisional anhedonia to emphasize the influence of anhedonic symptoms on decision-making. In contrast, Der-Avakian and Markou (2012) recently argued that deficits in motivational and decision-making processes (albeit disturbed, e.g., in depressed patients) should not be labeled under the umbrella of anhedonia.

Overall, findings from affective neuroscience have challenged Ribot's (1896) definition, which is restricted to subjectively experienced pleasure. Accumulating evidence suggests that reward consists of multiple subcomponents and processes of wanting, liking and learning (Robinson and Berridge, 2003; Berridge and Kringelbach, 2008) and these processes have been proposed to relate to appetitive, consummatory and satiety phases of a pleasure cycle (Kringelbach et al., 2012). Building on this work, we recently proposed to reconceptualize anhedonia as "impaired ability to pursue, experience and/or learn about pleasure, which is often, but not always accessible to conscious awareness" (Rømer Thomsen et al., 2015, p. 2).

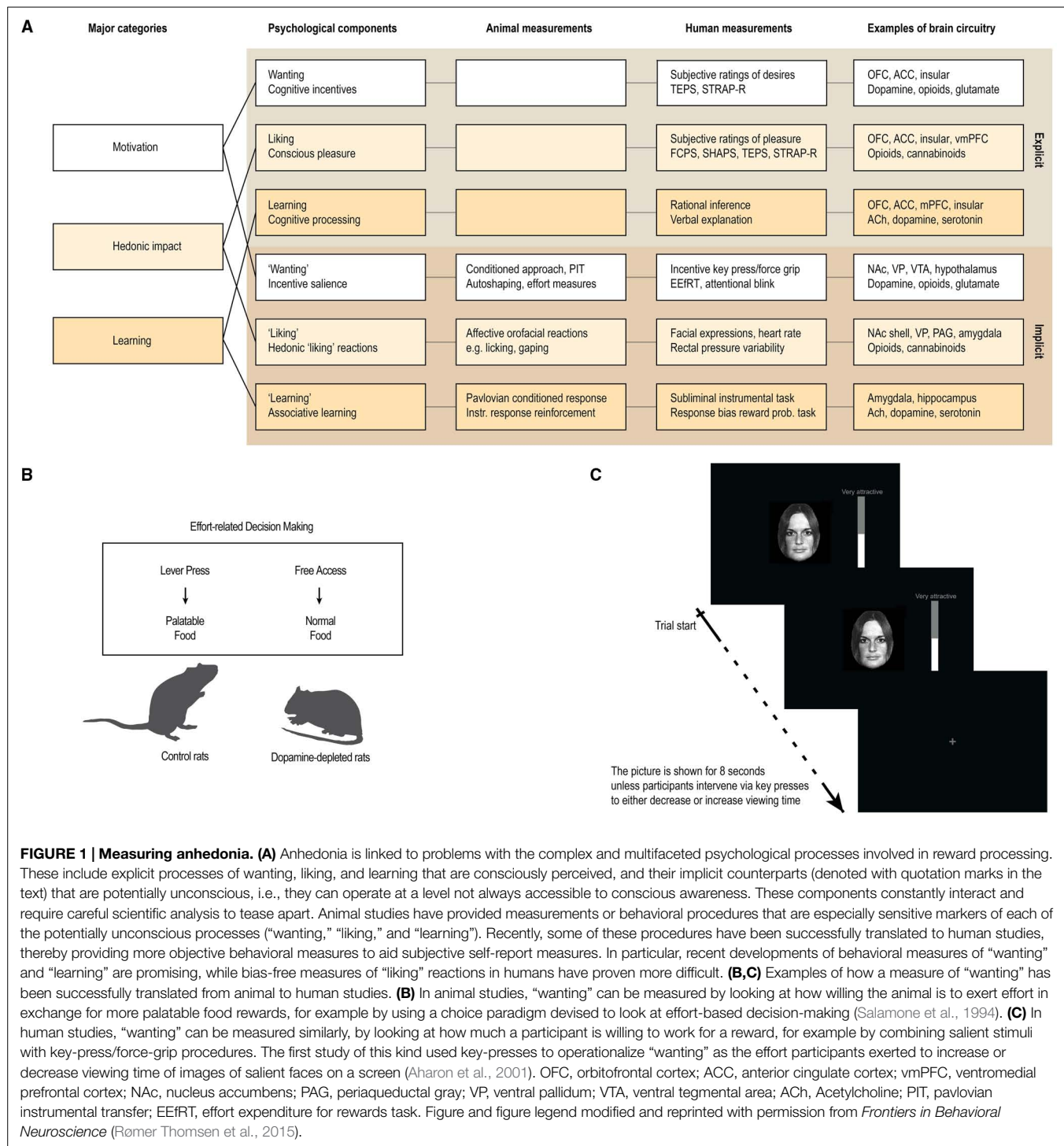
The parsing of reward into wanting, liking and learning components was originally introduced by Robinson and Berridge (1993) in their influential incentive sensitization theory of drug addiction. The theory has received support in animal and human studies of drug addiction (Vezina and Leyton, 2009; Leyton and Vezina, 2013) and recently also in terms of behavioral addiction like Gambling Disorder (Leyton and Vezina, 2012; Rømer Thomsen et al., 2014). In Robinson and Berridge's taxonomy they differentiate between core reactions that are not necessarily conscious ("wanting," "liking," and "learning") and their conscious counterparts (wanting, liking, and learning, i.e., denoted without quotation marks; Berridge and Robinson, 2003; Berridge and Kringelbach, 2008). In other words, reward can be parsed into three main components—motivation, hedonic impact and learning—and each of these components consist of both conscious and unconscious subcomponents (see **Figure 1A**). For example, motivation consists of "(1) core incentive salience 'wanting' processes that are not necessarily conscious (e.g., cue-triggered 'wanting' for food or drugs) and (2) conscious desires for incentives or cognitive goals" (Berridge and Kringelbach, 2008, p. 2). Hedonic impact consists of "(1) core 'liking' reactions that need not necessarily be conscious and (2) conscious experiences of pleasure, in the ordinary sense of the word, which may be elaborated out of core 'liking' reactions by brain mechanisms of awareness" (Berridge and Kringelbach, 2008, p. 2). Similarly, learning (or learned predictions) include "(1) implicit knowledge as well as associative conditioning, such as basic pavlovian and instrumental associations and (2) explicit and cognitive predictions" (Berridge and Kringelbach, 2008, p. 2).

The subcomponents of reward constantly interact through the appetitive, consummatory and satiety phases of a pleasure cycle, but can be teased apart using systematic scientific analysis. Self-report measures can help identify the conscious components (wanting, liking, and learning) and provide valuable information on this level of processing. However, self-report measures are of

course limited in their ability to capture unconscious processes, as well as in their ability to parse out contributions that may have been made by any of the unconscious processes, considering that these processes interact strongly together. In contrast, behavioral procedures from animal studies provide useful markers of the core "wanting," "liking," and "learning" reactions (**Figure 1B**). For example, "liking" reactions have been studied in rodents by measuring the affective orofacial expressions that are elicited in response to sweet tastes (Pfaffmann et al., 1977; Grill and Norgren, 1978a,b), and a number of procedures have been developed to study "wanting" in rodents, e.g., by measuring the effort exerted to obtain rewards (Salamone et al., 2007) or the ability of reward-related cues to act as motivational magnets (Wyvell and Berridge, 2000). In recent years, some of these animal models have been successfully translated to human studies and provide valuable behavioral measures of subcomponents of reward, which can complement traditional self-report measures (**Figure 1C**).

Overall, findings from animal and human studies applying these types of measures support the view that reward is a complex process consisting of several psychological components that correspond to partly dissociable neurobiological mechanisms (Berridge and Robinson, 2003; Berridge and Kringelbach, 2008, 2015). For example, there is strong evidence that dopamine plays an important role in "wanting," but not in "liking" reactions. In animal and human studies where "wanting" and "liking" reactions have been systematically teased apart, specific manipulation of dopamine signaling has failed to shift "liking" reactions to rewards (Berridge and Valenstein, 1991; Peciña et al., 2003; Ward et al., 2012). In contrast, there is accumulating evidence that dopamine plays an important role in "wanting" processes. For example, elevation of dopamine has been shown to increase willingness to work for a food reward in rodents (Bardgett et al., 2009), while dopamine attenuation or blockade has the opposite effect (Cousins and Salamone, 1994; Salamone et al., 2007). Similarly, evidence from human studies suggests that amphetamine/L-Dopa-induced elevated dopamine increases subjective ratings of drug wanting, but not subjective ratings of drug liking during consumption (Leyton et al., 2002, 2007; Higgins et al., 2012). Recently, Wardle et al. (2011) provided evidence that elevated levels of dopamine increase willingness to work for reward in humans using a behavioral measure.

Building on the framework set forward by Berridge and Robinson (2003) suggesting that reward consists of multiple subcomponents of wanting, liking, and learning and recent proposals relating these processes to the appetitive, consummatory and satiety phases of a pleasure cycle (Kringelbach et al., 2012), we recently proposed to reconceptualize anhedonia as "impairments in the ability to pursue, experience and/or learn about pleasure" (Rømer Thomsen et al., 2015, p. 2). In this conceptualization of anhedonia, impairments in each of the subcomponents can lead to a malfunctioning pleasure system. Normally, wanting, liking, and learning processes are balanced over time, however this balance can be compromised by impairments in each of the components. Depending on which of the subcomponents are most affected, and how the components are affected, this can lead to distinct subtypes of anhedonia, that are associated with distinct imbalances of the pleasure system



(Rømer Thomsen et al., 2015). For example, patients suffering from major depressive disorder often describe a diminished ability to pursue and experience pleasure, i.e., a progressive decrease in some (or all) of the reward components. In contrast, drug addiction can be characterized by excessive wanting for the drug of choice, which grows over time independently of drug liking. While anhedonia has traditionally been conceived as diminished responses (typically, diminished subjectively

experienced pleasure), "our proposed framework acknowledges that both too much and too little activity in specific parts of the pleasure system can lead to pathological changes. This is for example illustrated in the excessive wanting for drugs in drug addiction or in disorders with hypersexuality" (Rømer Thomsen et al., 2015, p. 15).

It is important to note, that in this terminology (Rømer Thomsen et al., 2015) *pleasure* and *pleasure system* is not restricted

to hedonic impact, but is instead used to encompass all of the phases of reward processing. This is in contrast to the dominating terminology, where *pleasure* is restricted to the hedonic impact of a reward, while *reward* is used to encompass all of the reward-related processes (see, e.g., Berridge and Kringelbach, 2008). In an attempt to avoid misunderstandings, I have changed the wording of our definition in the present paper to reflect the dominating terminology. Hence, anhedonia is defined here as “impairments in the ability to pursue, experience, and/or learn about *reward*.”

In line with our proposed reconceptualization of anhedonia, there has been a growing bulk of evidence suggesting that reward processing deficits are not restricted to impaired hedonic impact in psychiatric disorders typically associated with anhedonia. Findings from the past 5 years suggest that motivational and learning processes are impaired, e.g., in patients suffering from major depressive disorder (subsequently referred to as depression) and schizophrenia (Treadway and Zald, 2011; Fervaha et al., 2013b; Rømer Thomsen et al., 2015; Whitton et al., 2015). Part of this work is based on successful translations of animal models, thereby paving the way for validated behavioral paradigms that can supplement traditional self-report measures. These efforts are exciting and hold promise in terms of elucidating the role of subcomponents of anhedonia and the underlying neurobiological mechanisms across major psychiatric disorders. Here I review recent progress in the development and application of behavioral procedures that measure subcomponents of anhedonia in relevant patient groups (including patients suffering from depression, schizophrenia and addiction) and discuss implications for clinical assessment and treatment.

Measuring Subcomponents of Anhedonia

In line with the generally accepted understanding of anhedonia as “decreased subjective experience of pleasure” [as per Ribot’s (1896) definition], the most popular way of measuring anhedonia has been self-report scales or questionnaires like The Fawcett–Clark Pleasure Scale (FCPS; Fawcett et al., 1983) or The Snaith–Hamilton Pleasure Scale (SHAPS; Snaith et al., 1995). The majority of these instruments are restricted to measuring subjective experiences of hedonic impact (i.e., liking), but some of the more recently developed questionnaires also include aspects of reward motivation (i.e., wanting). For example, The Temporal Experience of Pleasure Scale (TEPS; Gard et al., 2006) differentiates between anticipatory and consummatory experiences of pleasure, and The Sensitivity To Reinforcement of Addictive and other Primary Rewards (STRAP-R; Goldstein et al., 2010) measures liking and wanting of drug and non-drug rewards under various situations (e.g., current and hypothetical). Building on Robinson and Berridge’s incentive sensitization theory (Robinson and Berridge, 1993; Robinson et al., 2013), Lende (2005) developed a short Incentive Salience Scale that measures key aspects of drug wanting and has been used to predict addiction status.

While these instruments provide useful information about the conscious components of anhedonia, they are of course limited in their ability to capture unconscious components. Similar to research on reward (Berridge and Robinson, 2003; Berridge and

Kringelbach, 2008), it is crucial to differentiate between conscious and unconscious components of anhedonia (Rømer Thomsen et al., 2015). Accumulating evidence suggests that we do not always know what motivates our behavior or brings us pleasure (Aharon et al., 2001; Winkielman et al., 2005; Moeller et al., 2009; Parsons et al., 2011), and there is convincing evidence that reward (also) affects our behavior on an unconscious level (Winkielman et al., 2005; Pessiglione et al., 2007, 2008; Aarts et al., 2008).

During the last 5 years a number of validated and useful behavioral procedures have been developed that can be used to measure impairments in the described subcomponents of reward (Figure 1). Of particular relevance are recent developments in behavioral procedures that can be used to measure impairments in the ability to pursue and learn about reward.

Impaired Ability to Pursue Reward

A large number of animal models have been developed to study motivational processes by looking at behavior related to obtainment of rewards such as food. Of particular relevance here are models of the effort exerted to obtain rewards [e.g., by measuring how eagerly the animal runs for rewards in a runway (Berridge and Valenstein, 1991; Peciña et al., 2003) or its willingness to exert effort in exchange for more palatable food rewards (Salamone et al., 1994, 2007)] and of the ability of reward-related cues to act as motivational magnets [e.g., by measuring pavlovian instrumental transfer (Wyvell and Berridge, 2000, 2001)]. Recently, some of these models have been successfully translated to studies of humans and the reported findings offer intriguing information on the role of reward motivation across major psychiatric disorders.

The effort expenditure for rewards task (EEfRT), which was developed by Treadway et al. (2009), represent a good example of how a validated animal model of motivation (Salamone et al., 1994) can be successfully translated to human studies. The EEfRT is an effort-based decision-making task, where reduced reward motivation is operationalized as a decreased willingness to choose greater-effort/greater-reward over less-effort/less-reward options with varying probability (Treadway et al., 2009). Recently, the task has been applied to relevant clinical populations and provide evidence of reduced willingness to expend effort for rewards in patients with subsyndromal depression, first-episode depression and remitted depression, compared to controls (Treadway et al., 2012; Yang et al., 2014).

A recent longitudinal study of reward seeking behavior in individuals at risk of depression provides intriguing evidence of diminished reward motivation as a potential precursor of depression (Rawal et al., 2013). Adolescent offspring of depressed parents performed the Cambridge Gambling Task in order to measure betting behavior under different odds. Compared to healthy adolescents and adolescents with externalizing disorders, the adolescent offspring of depressed parents showed diminished reward seeking (i.e., betted less at favorable odds). Importantly, the magnitude of this diminished response predicted depressive symptoms, depression-onset and functional impairment 1 year later (Rawal et al., 2013).

Several recent studies have reported decreased willingness to work for rewards using the EEfRT or similar tasks in patients

suffering from schizophrenia (Fervaha et al., 2013c; Gold et al., 2013; Barch et al., 2014). For example, Barch et al. (2014) reported that patients with schizophrenia were less willing to work harder when the size of the rewards increased or when the rewards were more probable, compared to control participants. Furthermore, among patients with schizophrenia, there was an association between choosing fewer greater-effort/greater-reward choices in the task and having more severe negative symptoms (self-reported) and worse community and work function (reported by caretaker; Barch et al., 2014).

Overall these findings are exciting and promising by providing strong evidence of reduced reward motivation across major psychiatric disorders. However, as we have previously stressed (Rømer Thomsen et al., 2015) participants are working for abstract rewards in these tasks, and not fundamental rewards (as in the animal models). Whether abstract and fundamental rewards are treated in the same way remains an open question, but emerging evidence suggests that there are important differences in the underlying brain processing (Sescousse et al., 2013a,b).

Other groups have used a related measure of reward motivation in humans by combining a key-pressing procedure with salient face stimuli (Aharon et al., 2001; Parsons et al., 2011). In these tasks, “wanting” is operationalized as the amount of work participants perform (i.e., by pressing a key) in order to change the duration they view images of adult/infant faces on a screen. Findings from these studies provide support for a dissociation of conscious liking ratings of salient face stimuli and the behavioral measure of “wanting” (Aharon et al., 2001; Parsons et al., 2011). For example, heterosexual males used more effort to keep female compared to male faces on a screen, but in a self-report task they rated the faces as equally attractive (Aharon et al., 2001). The use of salient face stimuli in combination with a key-pressing procedure represents a promising way to study possible impairments in the ability (or willingness) to work for fundamental social rewards in humans.

Moeller et al. (2009) have used a similar key-pressing paradigm in combination with salient drug-related stimuli and provide evidence for increased “wanting” of drug-related stimuli in drug addicted: cocaine addicted used more effort to view cocaine-related stimuli in a behavioral choice task, compared to control participants. Furthermore, they reported dissociation between a self-report measure of hedonic impact and a behavioral measure of motivation in cocaine addicted individuals: in the self-report task they rated pictures of pleasant scenes as more pleasant than cocaine-related pictures, however, in the behavioral choice task they did not show this preference (Moeller et al., 2009). These findings are in line with the incentive sensitization theory (Robinson and Berridge, 1993; Robinson et al., 2013) which argues that cue-triggered “wanting” of drug-related stimuli is enhanced in drug addicted individuals, and that these “wanting” processes are partly dissociable from “liking” processes in the brain, and in behavior. The reported dissociation between self-reported hedonic impact and a behavioral measure of motivation also reflects the impaired insight that characterizes drug addicted individuals (Goldstein et al., 2009; Moeller and Goldstein, 2014).

A related and promising measure of effort is the use of force-grip procedures which allows us to quantify various aspects of

effort including; how much effort is exerted over time, how fast participants start to squeeze, or how fast the force is increased (Aarts et al., 2008). By combining force-grip procedures with subliminal priming paradigms it becomes possible to study motivational processes that we are not aware of (Pessiglione et al., 2007; Aarts et al., 2008). For example, it has been shown that subliminally priming of the concept of exertion (i.e., words such as “exert” or “vigorous”) can prepare people for forceful action, and when primes are accompanied with a rewarding stimulus (i.e., a consciously visible positive word) they are motivated to spend more effort (Aarts et al., 2008). In a similar set-up, Pessiglione et al. (2007) studied unconscious motivation with an Incentive Force Task, where the amount and reportability of monetary rewards participants could gain through physical effort varied. Pessiglione et al. (2007) reported that even when participants were unable to report how much money was at stake, they still used more effort for larger rewards. These paradigms have yet to be applied to samples of relevant patients, but they represent a promising way to investigate impairments in unconscious reward motivation.

Another important component of reward motivation is the ability of reward-related cues to capture our attention and act as motivational magnets. In human studies, one way of operationalizing the ability of certain stimuli to capture our attention is by using variants of the attentional blink paradigm. Studies using this type of measure provide evidence that drug addicted individuals display an attentional bias toward drug-related visual cues and that this bias is correlated with self-reported craving (Wiers and Stacy, 2006; Field et al., 2009; Tibboel et al., 2010). For example, heavy social drinkers have reduced attentional blink for alcohol-related stimuli, which is consistent with the hypothesis of enhanced attentional bias for salient drug-related cues (Tibboel et al., 2010). Recent evidence suggests that a similar mechanism is present in behavioral addiction like Gambling Disorder (Brevers et al., 2011a,b; Rømer Thomsen et al., 2014). For example, in an attentional blink paradigm problem gamblers exhibited enhanced processing of gambling-related cues compared to neutral cues (Brevers et al., 2011b), and in a change detection task problem gamblers were faster at detecting gambling-related stimulus changes compared to neutral (Brevers et al., 2011a). Taken together, these findings support the hypothesis of an attentional bias toward addiction-related stimuli in drug and behavioral addiction.

Impaired Ability to Learn About Reward

There is an extensive literature from animal and human studies investigating the ability to learn from experiences with reward and punishment. Recently, some of these paradigms have been applied to relevant patient groups and provide evidence of impairments in the ability to learn about reward in patients suffering from depression and schizophrenia.

In a series of studies, Pizzagalli and colleagues have investigated impairments in the ability, or propensity, to develop a response bias toward stimuli that are more frequently rewarded than others using a probabilistic reward task (Pizzagalli et al., 2005, 2008; Pechtel et al., 2013; Vrieze et al., 2013). The task has been applied to patients with varying degree of depressive and anhedonia symptoms and findings from these studies consistently

show evidence of impaired reward learning. In the first study, Pizzagalli et al. (2005) showed that in participants with low levels of depressive symptoms there was an increase in the response bias over time, which was not present in participants with high levels of depressive symptoms. Subsequent studies of clinical populations show evidence of diminished reward responsiveness in depressed patients compared with controls (Pizzagalli et al., 2008; Vrieze et al., 2013), in patients with remitted depression compared with controls (Pechtel et al., 2013), and in depressed patients with high vs. low levels of anhedonia symptoms (Vrieze et al., 2013).

Of relevance here are also studies using probabilistic learning tasks that differentiate between reward-guided and punishment-guided learning. So far, this type of paradigm has not been systematically applied to clinically depressed patients, but one study reported evidence of blunted reward- and punishment-guided learning in depressed patients compared with controls (Chase et al., 2010). More data is available from patients suffering from schizophrenia. Compared to controls, patients suffering from schizophrenia consistently show deficits in reward-guided learning, while findings regarding punishment-guided learning are conflicting (Waltz et al., 2007, 2011; Strauss et al., 2011; Gold et al., 2012; Yilmaz et al., 2012; Fervaha et al., 2013a).

Studies targeting impairments in unconscious reward learning are intriguing, considering recent evidence of reward learning occurring outside our awareness. Pessiglione et al. (2008) used a subliminal instrumental conditioning task, where cues predicting monetary reward or punishment are subliminally presented, and showed that participants develop a propensity to choose cues associated with reward, even though the cues are not consciously perceived. These findings support the notion that cues related to reward and punishment (also) affect behavior and decision-making processes on an unconscious level and underscores the need to study reward processing deficits on both conscious and unconscious levels. This type of paradigm has yet to be applied to relevant patients, but represents a promising method to study potential impairments in unconscious reinforcement learning.

In animal studies, the conditioned place preference (CPP) procedure has long been used to study the development of preferences for environments or stimuli which have previously been associated with rewarding drug intake through the process of classical conditioning (Tzschentke, 1998, 2007). Recently, Mayo et al. (2013) successfully translated the CPP procedure into a human drug conditioning task and showed that healthy participants develop a behavioral preference for cues that have been paired with drug intake (a dose of methamphetamine), compared with cues that have been paired with placebo. These findings were recently replicated and extended by including a broader range of measures of the conditioned drug response, including self-report, behavioral and psychophysiological measures. After the conditioning procedure, participants showed an increase in behavioral preference, positive emotional reactivity, and attentional bias toward the cue associated with drug intake, compared with the cue associated with placebo (Mayo and de Wit, 2015). This paradigm represents a promising method to study individual determinants of classical conditioning and is therefore highly relevant for the disorders discussed here. For example, this type of paradigm can shed light on individual risk factors in

the development of sensitized responses to drugs/drug-related cues and blunted responses to other types of rewards (e.g., social, sexual, and sensory) in drug addiction, and similarly in behavioral addiction such as Gambling Disorder. This procedure is also highly promising in terms of studying deficient associative learning in patients suffering from depression and schizophrenia, preferably by using different types of rewards.

In line with the strong evidence suggesting that “wanting,” “liking,” and “learning” processes are dissociable in the brain and in behavior, it is important to note that the impaired reward learning reviewed here is not necessarily related to impairments in the ability to learn about “liking” (i.e., the hedonic impact of a reward), but could as easily be due to a reduced or modified sensitivity to the rewarding properties of the stimulus in the absence of “liking.” Future studies are needed to tease these differences apart.

Impaired Ability to Experience Pleasure

While successful models have been developed to study aspects of reward motivation and reward learning in humans, it has proven more difficult to develop behavioral procedures that measure hedonic impact in humans. In animal studies, hedonic impact of pleasurable stimuli has been successfully studied by measuring affective orofacial expressions elicited by the hedonic impact of sweet tastes. Studies applying taste-reactivity paradigms have convincingly shown that sweet tastes elicit rhythmic licking of lips (i.e., facial “liking” reactions) and bitter tastes elicit gapes (i.e., facial “disliking” reactions) in rodents and human infants (Steiner, 1973, 1974; Pfaffmann et al., 1977; Grill and Norgren, 1978a,b; Steiner et al., 2001). However, these affective orofacial measures are not easily translated to (adult) human studies, because we learn to control and mimic orofacial reactions to food as we grow up.

The hedonic impact of other types of rewards (than food) appears to be easier to measure behaviorally, or physiologically. Although mostly taboo, there is an increasing interest in the mechanisms underlying sexual pleasures (Georgiadis and Kringelbach, 2012; Georgiadis et al., 2012), and a number of measures have been developed to quantify pleasure-elicited “liking” reactions to sexual pleasures, e.g., by measuring rectal pressure variability and self-reported level of sexual arousal (Georgiadis et al., 2006). Although impairments related to sexual activity and sexual pleasures are still taboo, they represent a promising area of research that can help shed light on impairments in hedonic impact in relevant patient groups, including patients suffering from depression, schizophrenia and addiction.

A number of studies have measured facial reactions to pictures of emotional facial expressions and there is some evidence of a blunted response to positive facial expressions in depressed patients (Bylsma et al., 2008). For example, Dimberg (1982, 1990) has used electromyographic (EMG) recordings to detect emotion-related facial movements and shown that we elicit distinct facial reactions in response to emotional facial expressions, which partly reflects a tendency to mimic the facial expression. Studies have shown that these reactions are elicited very rapidly (Dimberg and Thunberg, 1998) and even when participants are not aware that they are being exposed to facial stimuli (Dimberg et al., 2000). Although it is unlikely that all changes in facial musculature

are related to emotion, EMG recordings of facial reactions may provide a way to investigate deficits in “liking” reactions to social pleasure (e.g., happy facial expressions). These rapid facial reactions have already been related to empathy (Dimberg et al., 2011), however more work is needed to confirm that they are in fact indicators of pleasure “liking.”

So far, the most popular way of measuring hedonic impact in humans has been to measure self-reported hedonic reactivity (i.e., subjective ratings of pleasure) to pleasant solutions and odors in a here-and-now setting. Surprisingly, the majority of studies report similar, or higher, pleasantness ratings in depressed patients compared to controls in response to sweet solutions (Amsterdam et al., 1987; Berlin et al., 1998; Scinska et al., 2004; Swiecicki et al., 2009; Dichter et al., 2010) and various odors (Steiner et al., 1993; Pause et al., 2001; Lombion-Pouthier et al., 2006; Scinska et al., 2008; Clepce et al., 2010). Similarly, evidence from studies of patients suffering from schizophrenia does not suggest that this patient group experiences lower levels of hedonic reactivity to pleasurable stimuli compared with controls (Heerey and Gold, 2007; Barch and Dowd, 2010; Strauss and Gold, 2012).

Interestingly, the same patient groups (depressed and schizophrenic) report diminished enjoyment in studies where they are asked to rate prospective, retrospective, or hypothetical experiences (McFarland and Klein, 2009; Watson and Naragon-Gainey, 2010; Strauss and Gold, 2012). One way of interpreting this discrepancy is that anhedonic patients retain core “liking” reactions, but do not cognitively value them in the same way as they did before (Dichter et al., 2010; Berridge and Kringelbach, 2011). This interpretation should however, be seen in the light of standard clinical examinations where depressed patients often present with behavioral characteristics that do not only imply impairments in cognitive evaluations of their experiences. For example, clinicians often report less smiling and less reactivity to stimuli (in general) which might reflect diminished “liking.” The disagreement—between laboratory based studies using taste-reactivity paradigms and clinical observations of patients—underscores the need to consider methodological aspects. The laboratory based studies reviewed here (where they failed to show diminished “liking” reactions to pleasurable solutions and odors in depressed and schizophrenic patients) were all based on self-reported ratings of hedonic impact. It remains an open question whether behavioral or physiological measures of “liking” will inform us differently.

Implications for Assessment and Treatment

The generally accepted understanding of anhedonia as “diminished subjectively experienced pleasure” is reflected in current diagnostic classification systems. For example, in the DSM 5 anhedonia is one of two main symptoms needed for the diagnosis of depression and is defined as “decreased interest and pleasure in most activities most of the day” (American Psychiatric Association, 2013). In this definition of anhedonia, wanting and liking components are collapsed, which is in contrast to the accumulating evidence suggesting that these processes are in fact dissociable in the brain and in behavior. For example, findings

from animal and human studies suggest that dopamine plays a crucial role in reward motivation (“wanting” and wanting), but not in hedonic impact (“liking” and liking; Berridge and Robinson, 2003; Berridge and Kringelbach, 2008).

Further, findings from affective neuroscience suggest that reward processing deficits are *not* restricted to impaired hedonic impact. As reviewed here, increasing evidence suggests that the ability to pursue and learn about reward is compromised in patients suffering from depression, schizophrenia, and drug/behavioral addiction (Treadway and Zald, 2011; Rømer Thomsen et al., 2015; Whitton et al., 2015). In contrast, it is less clear whether core “liking” reactions are in fact compromised in, e.g., depression and schizophrenia.

The growing evidence that reward processing deficits are not restricted to diminished experience of pleasure across major psychiatric disorders stresses the need to consider impairments in reward wanting and reward learning in clinical assessments. As a start, self-report instruments could be elaborated to reflect all phases of reward processing. The motivational aspect has already been included in some of the more recent questionnaires (e.g., the TEPS and the STRAP-R), but so far the learning component has been absent. Considering the growing evidence that unconscious components of reward affect our behavior, and are not always accompanied by conscious awareness (Berridge and Winkielman, 2003; Pessiglione et al., 2007, 2008; Aarts et al., 2008), it is highly debatable whether self-report instruments are sufficient in clinical assessments. Or whether they should be complemented by behavioral procedures. For example, behavioral measures of “wanting” could compliment self-report questionnaires in clinical assessments with advantage and help guide subsequent treatment. Depending on which subcomponents of reward processing are mainly affected, different medical treatments may be afforded. For example, depressed patients characterized by impaired ability to pursue pleasurable activities may benefit from medical interventions that target neurotransmitter systems such as the mesolimbic dopamine system and the opioid system, which have been shown to play a crucial role in reward motivation (Treadway and Zald, 2011; Soskin et al., 2013; Rømer Thomsen et al., 2015).

These insights are also relevant in terms of psychological treatment options. For example, cognitive behavioral therapy (CBT) has so far shown more promising treatment effects than pharmacological treatments in patients suffering from drug and behavioral addiction (Gooding and Tarrier, 2009; Potenza et al., 2011; Bullock and Potenza, 2012). In the context of addiction, CBT is expected to improve the individual’s control over motivation by increasing awareness of cues that trigger craving and by learning skills that enable new patterns of thinking and acting (Potenza et al., 2011). These efforts are important and efficiently target conscious feelings of craving. However, this type of cognitive intervention has limited efficacy in terms of targeting unconscious mechanisms. In particular, cue-induced craving reactions that occur outside our awareness are not likely to be targeted in CBT, but play an important role in maintaining the addictive behavior as outlined, e.g., by the incentive sensitization theory of addiction (Robinson et al., 2013). Hence, although CBT reduces some of the cognitive layers of responsiveness to drug cues, it is very likely that unconscious layers persist (Robinson et al., 2013).

Other types of psychological interventions may provide a way to target unconscious “wanting” (or “craving”) mechanisms, such as mindfulness based interventions that aim to improve the individual’s awareness of bodily and emotional signals (Garland et al., 2014). There is some evidence to suggest that mindfulness based interventions can reduce consumption and craving of a number of substances in substance users, although more randomized controlled trials are warranted (Chiesa and Serretti, 2014). For example, in a recent randomized controlled trial Tang et al. (2013) reported that brief meditation training reduced smoking by 60% in smokers who wanted to quit smoking, which was accompanied by increased activity in brain regions related to self-control and self-awareness. These findings foster hope that mindfulness based interventions can improve self-control and awareness of otherwise unconscious “wanting” reactions, and stresses the need to consider these types of treatments in combination with CBT, although more randomized controlled studies are warranted.

Concluding Remarks

Ribot’s (1896) long standing definition of anhedonia as “the inability to experience pleasure” has been challenged following progress in affective neuroscience, and in particular following pioneering work suggesting that reward consists of multiple subcomponents that can be divided into the processes of wanting, liking and learning (Berridge and Kringelbach, 2008). Recent proposals to reconceptualize anhedonia as motivational or

consummatory subtypes of anhedonia (Treadway and Zald, 2011), or as impaired ability to pursue, experience, and/or learn about pleasure (Rømer Thomsen et al., 2015) have paved the way for objective behavioral measures to complement traditional self-report measures of anhedonia. As reviewed here, a number of behavioral procedures have been developed that can be used to measure impairments in reward motivation and reward learning, while behavioral measures of hedonic impact have proven more difficult. Findings from studies applying these methods support the new conceptualizations of anhedonia by providing robust evidence that reward processing deficits are not restricted to impaired hedonic impact in major psychiatric disorders. Instead, there is increasing evidence of impairments in the ability to pursue and learn about reward in, e.g., depression and schizophrenia. This progress is essential for the purposes of identifying the underlying neurobiological mechanisms of anhedonia, and has important clinical implications for assessment and treatment of anhedonia. For example, self-report measures of anhedonia could be elaborated to reflect all phases of reward processing and it is debatable whether self-report measures of anhedonia are sufficient, or whether they should be complemented by behavioral measures in clinical assessments.

Acknowledgments

I am grateful for the support of the Ministry of Children, Gender Equality, Integration and Social Affairs (Denmark).

References

- Aarts, H., Custers, R., and Marien, H. (2008). Preparing and motivating behavior outside of awareness. *Science* 319, 1639. doi: 10.1126/science.1150432
- Aharon, I., Etcoff, N., Arieli, D., Chabris, C. F., O'Connor, E., and Breiter, H. C. (2001). Beautiful faces have variable reward value: fMRI and behavioral evidence. *Neuron* 32, 537–551. doi: 10.1016/S0896-6273(01)00491-3
- American Psychiatric Association. (2013). *Diagnostic and Statistical Manual of Mental Disorders: DSM-V*. Washington DC: American Psychiatric Association.
- Amsterdam, J. D., Settle, R. G., Doty, R. L., Abelman, E., and Winokur, A. (1987). Taste and smell perception in depression. *Biol. Psychiatry* 22, 1481–1485. doi: 10.1016/0006-3223(87)90108-9
- Barch, D. M., and Dowd, E. C. (2010). Goal representations and motivational drive in schizophrenia: the role of prefrontal-striatal interactions. *Schizophr. Bull.* 36, 919–934. doi: 10.1093/schbul/sbq068
- Barch, D. M., Treadway, M. T., and Schoen, N. (2014). Effort, anhedonia, and function in schizophrenia: reduced effort allocation predicts amotivation and functional impairment. *J. Abnorm. Psychol.* 123, 387–397. doi: 10.1037/a0036299
- Bardgett, M. E., Depenbrock, M., Downs, N., Points, M., and Green, L. (2009). Dopamine modulates effort-based decision making in rats. *Behav. Neurosci.* 123, 242–251. doi: 10.1037/a0014625
- Berlin, I., Givry-Steiner, L., Lecrubier, Y., and Puech, A. J. (1998). Measures of anhedonia and hedonic responses to sucrose in depressive and schizophrenic patients in comparison with healthy subjects. *Eur. Psychiatry* 13, 303–309. doi: 10.1016/S0924-9338(98)80048-5
- Berridge, K. C., and Kringelbach, M. L. (2008). Affective neuroscience of pleasure: reward in humans and animals. *Psychopharmacology* 199, 457–480. doi: 10.1007/s00213-008-1099-6
- Berridge, K. C., and Kringelbach, M. L. (2011). Building a neuroscience of pleasure and well-being. *Psychol. Well Being* 1, 1–3. doi: 10.1186/2211-1522-1-3
- Berridge, K. C., and Kringelbach, M. L. (2015). Pleasure systems in the brain. *Neuron* 86, 646–664. doi: 10.1016/j.neuron.2015.02.018
- Berridge, K. C., and Robinson, T. E. (2003). Parsing reward. *Trends Neurosci.* 26, 507–513. doi: 10.1016/S0166-2236(03)00233-9
- Berridge, K. C., and Valenstein, E. S. (1991). What psychological process mediates feeding evoked by electrical stimulation of the lateral hypothalamus? *Behav. Neurosci.* 105, 3–14. doi: 10.1037/0735-7044.105.1.3
- Berridge, K. C., and Winkelman, P. (2003). What is an unconscious emotion? (The case for unconscious “liking”). *Cogn. Emot.* 17, 181–211. doi: 10.1080/02699930302289
- Brevers, D., Cleeremans, A., Bechara, A., Laloyaux, C., Kornreich, C., Verbanck, P., et al. (2011a). Time course of attentional bias for gambling information in problem gambling. *Psychol. Addict. Behav.* 25, 675–682. doi: 10.1037/a0024201
- Brevers, D., Cleeremans, A., Tibboel, H., Bechara, A., Kornreich, C., Verbanck, P., et al. (2011b). Reduced attentional blink for gambling-related stimuli in problem gamblers. *J. Behav. Ther. Exp. Psychiatry* 42, 265–269. doi: 10.1016/j.jbtep.2011.01.005
- Bullock, S. A., and Potenza, M. N. (2012). Pathological gambling: neuropsychopharmacology and treatment. *Curr. Psychopharmacol.* 1, 67–85. doi: 10.2174/2211556011201010067
- Bylsma, L. M., Morris, B. H., and Rottenberg, J. (2008). A meta-analysis of emotional reactivity in major depressive disorder. *Clin. Psychol. Rev.* 28, 676–691. doi: 10.1016/j.cpr.2007.10.001
- Chase, H. W., Frank, M. J., Michael, A., Bullmore, E. T., Sahakian, B. J., and Robbins, T. W. (2010). Approach and avoidance learning in patients with major depression and healthy controls: relation to anhedonia. *Psychol. Med.* 40, 433–440. doi: 10.1017/S0033291709990468
- Chiesa, A., and Serretti, A. (2014). Are mindfulness-based interventions effective for substance use disorders? A systematic review of the evidence. *Subst. Use Misuse* 49, 492–512. doi: 10.3109/10826084.2013.770027
- Clepece, M., Gossler, A., Reich, K., Kornhuber, J., and Thuerauf, N. (2010). The relation between depression, anhedonia and olfactory hedonic estimates—a pilot study in major depression. *Neurosci. Lett.* 471, 139–143. doi: 10.1016/j.neulet.2010.01.027
- Cousins, M. S., and Salamone, J. D. (1994). Nucleus accumbens dopamine depletions in rats affect relative response allocation in a novel cost/benefit procedure. *Pharmacol. Biochem. Behav.* 49, 85–91. doi: 10.1016/0091-3057(94)90460-X

- Der-Avakian, A., and Markou, A. (2012). The neurobiology of anhedonia and other reward-related deficits. *Trends Neurosci.* 35, 68–77. doi: 10.1016/j.tins.2011.11.005
- Dichter, G. S., Smoski, M. J., Kampov-Polevoy, A. B., Gallop, R., and Garbutt, J. C. (2010). Unipolar depression does not moderate responses to the Sweet Taste Test. *Depress. Anxiety* 27, 859–863. doi: 10.1002/da.20690
- Dimberg, U. (1982). Facial reactions to facial expressions. *Psychophysiology* 19, 643–647. doi: 10.1111/j.1469-8986.1982.tb02516.x
- Dimberg, U. (1990). Facial electromyography and emotional reactions. *Psychophysiology* 27, 481–494. doi: 10.1111/j.1469-8986.1990.tb01962.x
- Dimberg, U., Andreasson, P., and Thunberg, M. (2011). Emotional empathy and facial reactions to facial expressions. *J. Psychophysiol.* 25, 26–31. doi: 10.1027/0269-8803/a000029
- Dimberg, U., and Thunberg, M. (1998). Rapid facial reactions to emotional facial expressions. *Scand. J. Psychol.* 39, 39–45. doi: 10.1111/1467-9450.00054
- Dimberg, U., Thunberg, M., and Elmehed, K. (2000). Unconscious facial reactions to emotional facial expressions. *Psychol. Sci.* 11, 86–89. doi: 10.1111/1467-9280.00221
- Fawcett, J., Clark, D. C., Scheftner, W. A., and Gibbons, R. D. (1983). Assessing anhedonia in psychiatric patients. The pleasure scale. *Arch. Gen. Psychiatry* 40, 79–84. doi: 10.1001/archpsyc.1983.01790010081010
- Fervaha, G., Agid, O., Foussias, G., and Remington, G. (2013a). Impairments in both reward and punishment guided reinforcement learning in schizophrenia. *Schizophr. Res.* 150, 592–593. doi: 10.1016/j.schres.2013.08.012
- Fervaha, G., Foussias, G., Agid, O., and Remington, G. (2013b). Neural substrates underlying effort computation in schizophrenia. *Neurosci. Biobehav. Rev.* 37, 2649–2665. doi: 10.1016/j.neubiorev.2013.09.001
- Fervaha, G., Graff-Guerrero, A., Zakzanis, K. K., Foussias, G., Agid, O., and Remington, G. (2013c). Incentive motivation deficits in schizophrenia reflect effort computation impairments during cost-benefit decision-making. *J. Psychiatr. Res.* 47, 1590–1596. doi: 10.1016/j.jpsychires.2013.08.003
- Field, M., Munafo, M. R., and Franken, I. H. (2009). A meta-analytic investigation of the relationship between attentional bias and subjective craving in substance abuse. *Psychol. Bull.* 135, 589–607. doi: 10.1037/a0015843
- Gard, D. E., Germans Gard, M., Kring, A. M., and John, O. P. (2006). Anticipatory and consummatory components of the experience of pleasure: a scale development study. *J. Res. Pers.* 40, 1086–1102. doi: 10.1016/j.jrp.2005.11.001
- Garland, E. L., Froeliger, B., and Howard, M. O. (2014). Mindfulness training targets neurocognitive mechanisms of addiction at the attention-appraisal-emotion interface. *Front. Psychiatry* 4:173. doi: 10.3389/fpsy.2013.00173
- Georgiadis, J. R., Korteckaas, R., Kuipers, R., Nieuwenburg, A., Pruim, J., Reinders, A. A., et al. (2006). Regional cerebral blood flow changes associated with clitorally induced orgasm in healthy women. *Eur. J. Neurosci.* 24, 3305–3316. doi: 10.1111/j.1460-9568.2006.05206.x
- Georgiadis, J. R., and Kringelbach, M. L. (2012). The human sexual response cycle: brain imaging evidence linking sex to other pleasures. *Prog. Neurobiol.* 98, 49–81. doi: 10.1016/j.pneurobio.2012.05.004
- Georgiadis, J. R., Kringelbach, M. L., and Pfaus, J. G. (2012). Sex for fun: a synthesis of human and animal neurobiology. *Nat. Rev. Urol.* 9, 486–498. doi: 10.1038/nrurol.2012.151
- Gold, J. M., Strauss, G. P., Waltz, J. A., Robinson, B. M., Brown, J. K., and Frank, M. J. (2013). Negative symptoms of schizophrenia are associated with abnormal effort-cost computations. *Biol. Psychiatry* 74, 130–136. doi: 10.1016/j.biopsych.2012.12.022
- Gold, J. M., Waltz, J. A., Matveeva, T. M., Kasanova, Z., Strauss, G. P., Herbener, E. S., et al. (2012). Negative symptoms and the failure to represent the expected reward value of actions: behavioral and computational modeling evidence. *Arch. Gen. Psychiatry* 69, 129–138. doi: 10.1001/archgenpsychiatry.2011.1269
- Goldstein, R. Z., Craig, A. D., Bechara, A., Garavan, H., Childress, A. R., Paulus, M. P., et al. (2009). The Neurocircuitry of impaired insight in drug addiction. *Trends Cogn. Sci.* 13, 372–380. doi: 10.1016/j.tics.2009.06.004
- Goldstein, R. Z., Woicik, P. A., Moeller, S. J., Telang, F., Jayne, M., Wong, C., et al. (2010). Liking and wanting of drug and non-drug rewards in active cocaine users: the STRAP-R questionnaire. *J. Psychopharmacol.* 24, 257–266. doi: 10.1177/0269881108096982
- Gooding, P., and Tarrier, N. (2009). A systematic review and meta-analysis of cognitive-behavioural interventions to reduce problem gambling: hedging our bets? *Behav. Res. Ther.* 47, 592–607. doi: 10.1016/j.brat.2009.04.002
- Grill, H. J., and Norgren, R. (1978a). The taste reactivity test. I. Mimetic responses to gustatory stimuli in neurologically normal rats. *Brain Res.* 143, 263–279. doi: 10.1016/0006-8993(78)90568-1
- Grill, H. J., and Norgren, R. (1978b). The taste reactivity test. II. Mimetic responses to gustatory stimuli in chronic thalamic and chronic decerebrate rats. *Brain Res.* 143, 281–297. doi: 10.1016/0006-8993(78)90569-3
- Heerey, E. A., and Gold, J. M. (2007). Patients with schizophrenia demonstrate dissociation between affective experience and motivated behavior. *J. Abnorm. Psychol.* 116, 268–278. doi: 10.1037/0021-843X.116.2.268
- Kringelbach, M. L., Stein, A., and Van Hartevelt, T. J. (2012). The functional human neuroanatomy of food pleasure cycles. *Physiol. Behav.* 106, 307–316. doi: 10.1016/j.physbeh.2012.03.023
- Lende, D. H. (2005). Wanting and drug use: a biocultural approach to the analysis of addiction. *Ethos* 33, 100–124. doi: 10.1525/eth.2005.33.1.100
- Leyton, M., Aan Het Rot, M., Booij, L., Baker, G. B., Young, S. N., and Benkelfat, C. (2007). Mood-elevating effects of d-amphetamine and incentive salience: the effect of acute dopamine precursor depletion. *J. Psychiatry Neurosci.* 32, 129–136.
- Leyton, M., Boileau, I., Benkelfat, C., Diksic, M., Baker, G., and Dagher, A. (2002). Amphetamine-induced increases in extracellular dopamine, drug wanting, and novelty seeking: a PET/[¹¹C]raclopride study in healthy men. *Neuropsychopharmacology* 27, 1027–1035. doi: 10.1016/S0893-133X(02)00366-4
- Leyton, M., and Vezina, P. (2012). On cue: striatal ups and downs in addictions. *Biol. Psychiatry* 72, e21–e22. doi: 10.1016/j.biopsych.2012.04.036
- Leyton, M., and Vezina, P. (2013). Striatal ups and downs: their roles in vulnerability to addictions in humans. *Neurosci. Biobehav. Rev.* 37, 1999–2014. doi: 10.1016/j.neubiorev.2013.01.018
- Liggins, J., Pihl, R. O., Benkelfat, C., and Leyton, M. (2012). The dopamine augmenting L-DOPA does not affect positive mood in healthy human volunteers. *PLoS ONE* 7:e28370. doi: 10.1371/journal.pone.0028370
- Lombion-Pouthier, S., Vandel, P., Nezelof, S., Haffen, E., and Millot, J. L. (2006). Odor perception in patients with mood disorders. *J. Affect. Disord.* 90, 187–191. doi: 10.1016/j.jad.2005.11.012
- Mayo, L. M., and de Wit, H. (2015). Acquisition of responses to a methamphetamine-associated cue in healthy humans: self-report, behavioral, and psychophysiological measures. *Neuropsychopharmacology* 40, 1734–1741. doi: 10.1038/npp.2015.21
- Mayo, L. M., Fraser, D., Childs, E., Momenan, R., Hommer, D. W., De Wit, H., et al. (2013). Conditioned preference to a methamphetamine-associated contextual cue in humans. *Neuropsychopharmacology* 38, 921–929. doi: 10.1038/npp.2013.3
- McFarland, B. R., and Klein, D. N. (2009). Emotional reactivity in depression: diminished responsiveness to anticipated reward but not to anticipated punishment or to nonreward or avoidance. *Depress. Anxiety* 26, 117–122. doi: 10.1002/da.20513
- Moeller, S. J., and Goldstein, R. Z. (2014). Impaired self-awareness in human addiction: deficient attribution of personal relevance. *Trends Cogn. Sci.* 18, 635–641. doi: 10.1016/j.tics.2014.09.003
- Moeller, S. J., Maloney, T., Parvaz, M. A., Dunning, J. P., Alia-Klein, N., Woicik, P. A., et al. (2009). Enhanced choice for viewing cocaine pictures in cocaine addiction. *Biol. Psychiatry* 66, 169–176. doi: 10.1016/j.biopsych.2009.02.015
- Parsons, C. E., Young, K. S., Kumari, N., Stein, A., and Kringelbach, M. L. (2011). The motivational salience of infant faces is similar for men and women. *PLoS ONE* 6:e20632. doi: 10.1371/journal.pone.0020632
- Pause, B. M., Miranda, A., Goder, R., Aldenhoff, J. B., and Ferstl, R. (2001). Reduced olfactory performance in patients with major depression. *J. Psychiatr. Res.* 35, 271–277. doi: 10.1016/S0022-3956(01)00029-2
- Pechtel, P., Dutra, S. J., Goetz, E. L., and Pizzagalli, D. A. (2013). Blunted reward responsiveness in remitted depression. *J. Psychiatr. Res.* 47, 1864–1869. doi: 10.1016/j.jpsychires.2013.08.011
- Peciña, S., Cagniard, B., Berridge, K. C., Aldridge, J. W., and Zhuang, X. (2003). Hyperdopaminergic mutant mice have higher “wanting” but not “liking” for sweet rewards. *J. Neurosci.* 23, 9395–9402.
- Pessiglione, M., Petrovic, P., Daunizeau, J., Palminteri, S., Dolan, R. J., and Frith, C. D. (2008). Subliminal instrumental conditioning demonstrated in the human brain. *Neuron* 59, 561–567. doi: 10.1016/j.neuron.2008.07.005
- Pessiglione, M., Schmidt, L., Draganski, B., Kalisch, R., Lau, H., Dolan, R., et al. (2007). How the brain translates money into force: a neuroimaging study of subliminal motivation. *Science* 316, 904–906. doi: 10.1126/science.1140459

- Pfaffmann, C., Norgren, R., and Grill, H. J. (1977). Sensory affect and motivation. *Ann. N. Y. Acad. Sci.* 290, 18–34. doi: 10.1111/j.1749-6632.1977.tb39713.x
- Pizzagalli, D. A., Iosifescu, D., Hallett, L. A., Ratner, K. G., and Fava, M. (2008). Reduced hedonic capacity in major depressive disorder: evidence from a probabilistic reward task. *J. Psychiatr. Res.* 43, 76–87. doi: 10.1016/j.jpsychires.2008.03.001
- Pizzagalli, D. A., Jahn, A. L., and O'Shea, J. P. (2005). Toward an objective characterization of an anhedonic phenotype: a signal-detection approach. *Biol. Psychiatry* 57, 319–327. doi: 10.1016/j.biopsych.2004.11.026
- Potenza, M. N., Sofuoglu, M., Carroll, K. M., and Rounsaville, B. J. (2011). Neuroscience of behavioral and pharmacological treatments for addictions. *Neuron* 69, 695–712. doi: 10.1016/j.neuron.2011.02.009
- Rawal, A., Collishaw, S., Thapar, A., and Rice, F. (2013). 'The risks of playing it safe': a prospective longitudinal study of response to reward in the adolescent offspring of depressed parents. *Psychol. Med.* 43, 27–38. doi: 10.1017/S0033291712001158
- Ribot, T. (1896). *La Psychologie des Sentiment* [The Psychology of Feelings]. Paris: Felix Alcan.
- Robinson, M. J. F., Robinson, T. E., and Berridge, K. C. (2013). "Incentive salience and the transition to addiction," in *Biological Research on Addiction*, ed. P. Miller (San Diego: Academic Press), 391–399.
- Robinson, T. E., and Berridge, K. C. (1993). The neural basis of drug craving: an incentive-sensitization theory of addiction. *Brain Res. Brain Res. Rev.* 18, 247–291. doi: 10.1016/0165-0173(93)90013-P
- Robinson, T. E., and Berridge, K. C. (2003). Addiction. *Annu. Rev. Psychol.* 54, 25–53. doi: 10.1146/annurev.psych.54.101601.145237
- Rømer Thomsen, K., Fjorback, L. O., Møller, A., and Lou, H. C. (2014). Applying incentive sensitization models to behavioral addiction. *Neurosci. Biobehav. Rev.* 45C, 343–349. doi: 10.1016/j.neubiorev.2014.07.009
- Rømer Thomsen, K., Whybrow, P. C., and Kringelbach, M. L. (2015). Reconceptualizing anhedonia: novel perspectives on balancing the pleasure networks in the human brain. *Front. Behav. Neurosci.* 9:49. doi: 10.3389/fnbeh.2015.00049
- Salamone, J. D., Correa, M., Farrar, A., and Mingote, S. M. (2007). Effort-related functions of nucleus accumbens dopamine and associated forebrain circuits. *Psychopharmacology (Berl.)* 191, 461–482. doi: 10.1007/s00213-006-0668-9
- Salamone, J. D., Cousins, M. S., McCullough, L. D., Carriero, D. L., and Berkowitz, R. J. (1994). Nucleus accumbens dopamine release increases during instrumental lever pressing for food but not free food consumption. *Pharmacol. Biochem. Behav.* 49, 25–31. doi: 10.1016/0091-3057(94)90452-9
- Scinska, A., Sienkiewicz-Jarosz, H., Kuran, W., Ryglewicz, D., Rogowski, A., Wrobel, E., et al. (2004). Depressive symptoms and taste reactivity in humans. *Physiol. Behav.* 82, 899–904. doi: 10.1016/j.physbeh.2004.07.012
- Scinska, A., Wrobel, E., Korkosz, A., Zatorski, P., Sienkiewicz-Jarosz, H., Lojkowska, W., et al. (2008). Depressive symptoms and olfactory function in older adults. *Psychiatry Clin. Neurosci.* 62, 450–456. doi: 10.1111/j.1440-1819.2008.01824.x
- Sescousse, G., Barbalat, G., Domenech, P., and Dreher, J. C. (2013a). Imbalance in the sensitivity to different types of rewards in pathological gambling. *Brain* 136, 2527–2538. doi: 10.1093/brain/awt126
- Sescousse, G., Caldu, X., Segura, B., and Dreher, J. C. (2013b). Processing of primary and secondary rewards: a quantitative meta-analysis and review of human functional neuroimaging studies. *Neurosci. Biobehav. Rev.* 37, 681–696. doi: 10.1016/j.neubiorev.2013.02.002
- Snaith, R. P., Hamilton, M., Morley, S., Humayan, A., Hargreaves, D., and Trigwell, P. (1995). A scale for the assessment of hedonic tone. The Snaith-Hamilton Pleasure Scale. *Br. J. Psychiatry* 167, 99–103. doi: 10.1192/bjp.167.1.99
- Soskin, D. P., Holt, D. J., Sacco, G. R., and Fava, M. (2013). Incentive salience: novel treatment strategies for major depression. *CNS Spectr.* 18, 307–314. doi: 10.1017/S1092852913000345
- Steiner, J. E. (1973). The gustofacial response: observation on normal and anencephalic newborn infants. *Symp. Oral Sens. Percept.* 4, 254–278.
- Steiner, J. E. (1974). Discussion paper: innate, discriminative human facial expressions to taste and smell stimulation. *Ann. N. Y. Acad. Sci.* 237, 229–233. doi: 10.1111/j.1749-6632.1974.tb49858.x
- Steiner, J. E., Glaser, D., Hawilo, M. E., and Berridge, K. C. (2001). Comparative expression of hedonic impact: affective reactions to taste by human infants and other primates. *Neurosci. Biobehav. Rev.* 25, 53–74. doi: 10.1016/S0149-7634(00)00051-8
- Steiner, J. E., Lidar-Lifschitz, D., and Perl, E. (1993). Taste and odor: reactivity in depressive disorders, a multidisciplinary approach. *Percept. Mot. Skills* 77, 1331–1346. doi: 10.2466/pms.1993.77.3f.1331
- Strauss, G. P., Frank, M. J., Waltz, J. A., Kasantova, Z., Herbener, E. S., and Gold, J. M. (2011). Deficits in positive reinforcement learning and uncertainty-driven exploration are associated with distinct aspects of negative symptoms in schizophrenia. *Biol. Psychiatry* 69, 424–431. doi: 10.1016/j.biopsych.2010.10.015
- Strauss, G. P., and Gold, J. M. (2012). A new perspective on anhedonia in schizophrenia. *Am. J. Psychiatry* 169, 364–373. doi: 10.1176/appi.ajp.2011.11030447
- Swiecki, L., Zatorski, P., Bzinkowska, D., Sienkiewicz-Jarosz, H., Szyndler, J., and Scinska, A. (2009). Gustatory and olfactory function in patients with unipolar and bipolar depression. *Prog. Neuropsychopharmacol. Biol. Psychiatry* 33, 827–834. doi: 10.1016/j.pnpbp.2009.03.030
- Tang, Y. Y., Tang, R., and Posner, M. I. (2013). Brief meditation training induces smoking reduction. *Proc. Natl. Acad. Sci. U.S.A.* 110, 13971–13975. doi: 10.1073/pnas.1311887110
- Tibboel, H., De Houwer, J., and Field, M. (2010). Reduced attentional blink for alcohol-related stimuli in heavy social drinkers. *J. Psychopharmacol.* 24, 1349–1356. doi: 10.1177/0269881109106977
- Treadway, M. T., Bossaller, N. A., Shelton, R. C., and Zald, D. H. (2012). Effort-based decision-making in major depressive disorder: a translational model of motivational anhedonia. *J. Abnorm. Psychol.* 121, 553–558. doi: 10.1037/a0028813
- Treadway, M. T., Buckholtz, J. W., Schwartzman, A. N., Lambert, W. E., and Zald, D. H. (2009). Worth the 'Effort': The effort expenditure for rewards task as an objective measure of motivation and anhedonia. *PLoS ONE* 4:e6598. doi: 10.1371/journal.pone.0006598
- Treadway, M. T., and Zald, D. H. (2011). Reconsidering anhedonia in depression: lessons from translational neuroscience. *Neurosci. Biobehav. Rev.* 35, 537–555. doi: 10.1016/j.neubiorev.2010.06.006
- Tzschentke, T. M. (1998). Measuring reward with the conditioned place preference paradigm: a comprehensive review of drug effects, recent progress and new issues. *Prog. Neurobiol.* 56, 613–672. doi: 10.1016/S0304-0082(98)00060-4
- Tzschentke, T. M. (2007). Measuring reward with the conditioned place preference (CPP) paradigm: update of the last decade. *Addict. Biol.* 12, 227–462. doi: 10.1111/j.1369-1600.2007.00070.x
- Vezina, P., and Leyton, M. (2009). Conditioned cues and the expression of stimulant sensitization in animals and humans. *Neuropharmacology* 56(Suppl. 1), 160–168. doi: 10.1016/j.neuropharm.2008.06.070
- Vrieze, E., Pizzagalli, D. A., Demyttenaere, K., Hompes, T., Sienaert, P., De Boer, P., et al. (2013). Reduced reward learning predicts outcome in major depressive disorder. *Biol. Psychiatry* 73, 639–645. doi: 10.1016/j.biopsych.2012.10.014
- Waltz, J. A., Frank, M. J., Robinson, B. M., and Gold, J. M. (2007). Selective reinforcement learning deficits in schizophrenia support predictions from computational models of striatal-cortical dysfunction. *Biol. Psychiatry* 62, 756–764. doi: 10.1016/j.biopsych.2006.09.042
- Waltz, J. A., Frank, M. J., Wiecki, T. V., and Gold, J. M. (2011). Altered probabilistic learning and response biases in schizophrenia: behavioral evidence and neurocomputational modeling. *Neuropsychology* 25, 86–97. doi: 10.1037/a0020882
- Ward, R. D., Simpson, E. H., Richards, V. L., Deo, G., Taylor, K., Glendinning, J. I., et al. (2012). Dissociation of hedonic reaction to reward and incentive motivation in an animal model of the negative symptoms of schizophrenia. *Neuropsychopharmacology* 37, 1699–1707. doi: 10.1038/npp.2012.15
- Wardle, M. C., Treadway, M. T., Mayo, L. M., Zald, D. H., and De Wit, H. (2011). Amping up effort: effects of d-amphetamine on human effort-based decision-making. *J. Neurosci.* 31, 16597–16602. doi: 10.1523/JNEUROSCI.4387-11.2011
- Watson, D., and Naragon-Gainey, K. (2010). On the specificity of positive emotional dysfunction in psychopathology: evidence from the mood and anxiety disorders and schizophrenia/schizotypy. *Clin. Psychol. Rev.* 30, 839–848. doi: 10.1016/j.cpr.2009.11.002
- Whitton, A. E., Treadway, M. T., and Pizzagalli, D. A. (2015). Reward processing dysfunction in major depression, bipolar disorder and schizophrenia. *Curr. Opin. Psychiatry* 28, 7–12. doi: 10.1097/YCO.0000000000000122
- Wiers, R. W., and Stacy, A. W. (eds). (2006). *Handbook of Implicit Cognition and Addiction*. Thousand Oaks, CA: Sage Publications, Inc.
- Winkielman, P., Berridge, K. C., and Wilbarger, J. L. (2005). Unconscious affective reactions to masked happy versus angry faces influence consumption

- behavior and judgements of value. *Pers. Soc. Psychol. Bull.* 31, 121–135. doi: 10.1177/0146167204271309
- Wyvell, C. L., and Berridge, K. C. (2000). Intra-accumbens amphetamine increases the conditioned incentive salience of sucrose reward: enhancement of reward “wanting” without enhanced “liking” or response reinforcement. *J. Neurosci.* 20, 8122–8130.
- Wyvell, C. L., and Berridge, K. C. (2001). Incentive sensitization by previous amphetamine exposure: increased cue-triggered “wanting” for sucrose reward. *J. Neurosci.* 21, 7831–7840.
- Yang, X. H., Huang, J., Zhu, C. Y., Wang, Y. F., Cheung, E. F., Chan, R. C., et al. (2014). Motivational deficits in effort-based decision making in individuals with subsyndromal depression, first-episode and remitted depression patients. *Psychiatry Res.* 220, 874–882. doi: 10.1016/j.psychres.2014.08.056
- Yilmaz, A., Simsek, F., and Gonul, A. S. (2012). Reduced reward-related probability learning in schizophrenia patients. *Neuropsychiatr. Dis. Treat.* 8, 27–34. doi: 10.2147/NDT.S26243
- Conflict of Interest Statement:** The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.
- Copyright © 2015 Rømer Thomsen. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Unifying treatments for depression: an application of the Free Energy Principle

Adam M. Chekroud^{1,2*}

¹ Department of Psychology, Yale University, New Haven, CT, USA

² Department of Neuroscience, Oxford University, Oxford, UK

Edited by:

Nikolina Skandali, University of Cambridge, UK

Reviewed by:

Philip R. Corlett, Yale University

School of Medicine, USA

Christoph Daniel Mathys, University

College London, UK

*Correspondence:

Adam M. Chekroud, Department of Psychology, Yale University, 2

Hillhouse Avenue, New Haven, CT

06511, USA

e-mail: adam.chekroud@yale.edu

Major Depressive Disorder is a debilitating and increasingly prevalent psychiatric condition (Compton et al., 2006; Andersen et al., 2011). At present, its primary treatments are antidepressant medications and psychotherapy. Curiously, although the pharmacological effects of antidepressants manifest within hours, remission of clinical symptoms takes a number of weeks—if at all. Independently, support has grown for an idea—proposed as early as Helmholtz (von Helmholtz, 1924)—that the brain is a prediction machine, holding generative models¹ for the purpose of inferring causes of sensory information (Dayan et al., 1995; Rao and Ballard, 1999; Kniill and Pouget, 2004; Friston et al., 2006; Friston, 2010). If the brain does indeed represent a collection of beliefs about the causal structure of the world, then the depressed phenotype may emerge from a collection of depressive beliefs. These beliefs are modified gradually through successive combinations of expectations with observations. As a result, phenotypic remission ought to take some time as the brain's relevant statistical structures become less pessimistic.

Keywords: major depressive disorder, predictive coding, free-energy principle, antidepressants, computational psychiatry, generative models, antidepressants efficacy

THE FREE-ENERGY PRINCIPLE

The free-energy principle has been proposed as a unifying framework that simultaneously links perception and action, and formalizes the roles of brain theories including attention, motor control, and perceptual learning (Friston et al., 2006; Friston, 2010; Clark, 2013). It is a mathematical description whereby the brain is a predictive device that builds statistical models of the world and then seeks to minimize “free energy,” an approximation of surprise. Free energy depends on a number of quantities, including the internal states of the brain, the external environment, and exchanges between the two (through action and perception). Mathematically, free energy is a statistical quantity that approximates the surprise in a sensory input, and it rests on two probability densities: a recognition density and a generative model (Friston et al., 2006; Friston, 2010; Clark, 2013).

The first component—the recognition density—is an approximate probability distribution of the causes of sensory data (Friston, 2010). This is quite a simple concept to apply: when a visual neuron fires in response to a horizontal bar in an area of the visual field, we could think that the stimulus caused the neural response. Equally, from the point of the neuron, we can consider its firing to reflect a probability that our sensation of a stimulus was caused by a horizontal bar in a particular area. The second component—the generative model—is a joint probability density between data and their causes from which samples can

be drawn (Friston, 2010). In the present context, the generative model seeks to capture the statistical structure of its sensory environment by tracking the web of causes of that statistical structure. Crucially, the inferences we make about causes are not restricted to immediate sensory signals (e.g., “the switch caused light”), nor even time-varying/transitive inferences (e.g., “that bird is flying”) (Friston, 2010). The brain's model of the world also includes time-invariant regularities that afford structure to our world e.g., “gravity makes things fall,” but could equally be “I am in control of my own actions.” Having said this, it is important to note that Bayes rule and Bayesian brain theory do not guarantee veridical associations.

MODELS IN THE BRAIN?

A “belief” in the context of the free-energy principle is formalized as a probability distribution of an external state as internally represented by its sufficient statistics (Huys and Dayan, 2009; Friston, 2010; Mathys et al., 2011). For the purposes of this article, a “belief” is simplified as a prediction about the cause of an observation, given a particular circumstance and some previous experience. A “depressive belief” can then be considered as any consistent (negative) bias in these predictions, or vice versa². If our beliefs are to be useful, and reflect genuine associations rather than random co-occurrences we must consider the prior observations of all elements concerned (Fletcher and Frith, 2009). A simple thought experiment illustrates the concept of belief well:

¹Technically there is only one generative model, but for the purpose of this essay I refer to multiple internal models since the hierarchical structure supports many processing levels (as in Clark, 2013)

²Interested readers are encouraged to read an excellent formal treatment of model inversion by FitzGerald et al. (2014), which explores the implications of approximate Bayesian inference on behavior.

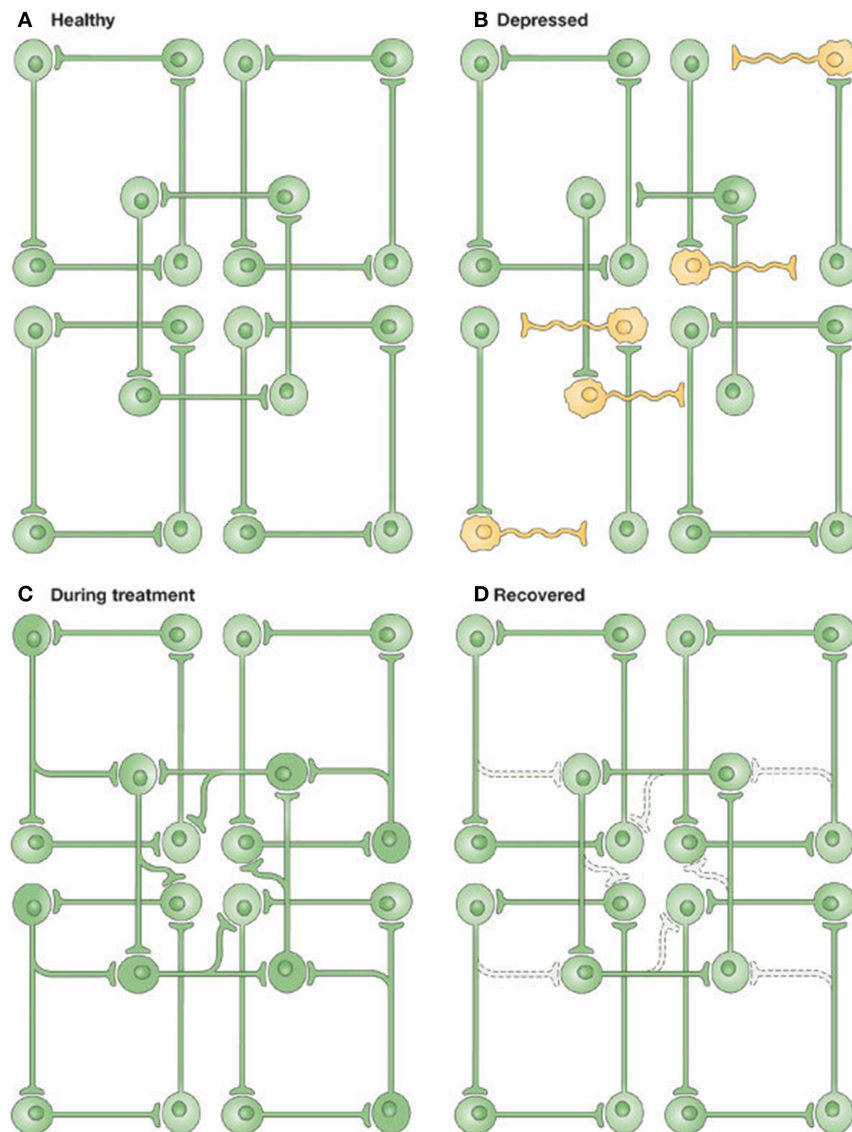


FIGURE 1 | Illustrating the network hypothesis of Depression. (A) In the healthy brain, information is distributed amongst partially overlapping brain networks. **(B)** In Depression, some information processing is altered. **(C)**

Antidepressant treatments enhance connectivity in neural networks. **(D)** Activity-dependent synaptic pruning stabilizes the network. Figure and caption reproduced with permission from Castrén (2005).

upon meeting a three-legged dog, one needs to recall all the previous times one encountered four-legged dogs to avoid the false prediction that dogs only have three legs.

Bayes' rule, a mathematical theorem, offers a mechanism for how beliefs should develop over time: updating as a function of past experiences (the prior), and the current experience (the likelihood) to produce a posterior belief or expectation. This interplay between likelihoods and priors may sound abstract, but it has the very practical implication that all our experiences depend on our knowledge of their predictability. The connection between the free-energy principle, predictive coding³ and the

Bayesian brain rests on the fact that minimizing free energy corresponds to variational Bayesian inference. This may sound technical; however, it brings an important insight to the table: namely, all quantities involved in making predictions must jointly minimize surprise or free-energy. Notably, proposed quantities include synaptic activity (encoding beliefs about the current state of the world), synaptic efficacy (encoding regularities and causal structure) and synaptic gain (encoding the precision of beliefs) (Corlett et al., 2009, 2011; Adams et al., 2013). This three way split provides a natural framework to understand perceptual inference, learning, and the encoding of uncertainty, respectively. Crucially,

³Predictive coding refers to a class of theories in which the brain is held to continually generate models of the world based on context and information

from memory to predict sensory input (Rao and Ballard, 1999; Friston and Kiebel, 2009; Clark, 2013).

to optimize any one set of these quantities one needs the optimal values of the others. The implicit circular dependency means that disruptions to inference, learning or the encoding of uncertainty will necessarily cause abnormalities in the other domains. Of particular importance here is the notion of precision. In predictive coding, precision corresponds to the (synaptic) gain applied to prediction errors and plays the role of a learning rate. We will return to this later when considering the link between neuromodulators, synaptic gain and their effects on perceptual inference and learning.

PERCEPTION AND BELIEF: WE SEE WHAT WE WANT TO SEE

Exchanges between our brain's internal states and our external environment are bidirectional. That is, the brain draws its input through perception as it forms a model of the world, and then engages the external environment through action. It is this sampling of the environment that dictates our sensations, thus completing an action-perception cycle. Consider an intuitive example that occurs as we wander through our bedroom in complete darkness. We anticipate what we might touch in the world around us (expectations), and then feel around accordingly as we attempt to confirm these expectations (selective sampling). This process—whereby an agent selectively samples the sensory inputs that it expects—is known as active inference (Friston et al., 2009; Friston and Kiebel, 2009). In most real-life cases, there is already considerable contextual (i.e., prior) information in place when we encounter new information (Friston et al., 2006; Clark, 2013). There is, therefore, the potential for many prior expectations to be primed, alter the processing of incoming sensory information, and influence future environment sampling through action (Friston, 2010; Clark, 2013).

WHEN THINGS GO WRONG

Predictions are only as good as the model that generates them. Disadvantages begin to creep into the system when beliefs become either inaccurate or inflexible (Fletcher and Frith, 2009; Ma, 2012). Recall that all of our experiences are influenced by our beliefs. Experiences that are in line with beliefs become predictable, strengthen the original belief and eliminate the need for the energy consuming processing of predictable sensations; because they have already been predicted and provide no “news-worthy” information. When an incorrect belief gains strength, it can result in one ignoring potentially informative experiences, or a range of other misattributions (Fletcher and Frith, 2009). The bidirectional belief-action relationship means that any inaccuracies in our model of the world might result in abnormal perception or action, and vice versa (Fletcher and Frith, 2009; Friston, 2010). Additionally, since the model must have a neural basis, correct predictions (in the form of some distributed neural network) could plausibly be disrupted by neurobiological changes.

CHANGING THE MODEL: MINIMIZING FREE-ENERGY

Friston's original proposition offered two mechanisms for minimizing free-energy: through optimizing actions, and optimizing representations (Friston, 2010). In other words, we must either change the inputs to the model, or change the internal states.

Returning to the example of walking through a dark room, there are two ways in which we might minimize surprise. One could sample differently (through action), e.g., turn on a light. Alternatively, one could change expectations (perceptual inference), e.g., entertain the alternative belief that you have woken up in a hotel room as opposed to your bedroom. It is critical to note here that both action and perception constitute an iterative cycle and depend upon each other. This contextualizes the three way dependencies between perceptual inference, learning and precision noted above: in other words, any changes in action rest upon changes in perception that—at some level—depend upon perceptual learning. Because perceptual learning proceeds at a much slower timescale than inference, our beliefs (which underlie action) do not change immediately; rather we successively combine past and current experience to optimize our generative model of the world. As such, rectifying a depressive model of the world (and thus the depressive phenotype) will be a gradual process. More specifically, this gradual process corresponds to the acquisition of generative models and involves the suppression of free-energy or prediction errors (over time) by changing connection strengths in the generative model. It is this process that one might consider to be the target of therapeutic interventions (e.g., by increasing learning rates—as is discussed later).

ANTIDEPRESSANTS: REPAIRING REPRESENTATIONS?

Interesting parallels arise when considering depression from a free energy viewpoint. Anhedonia—a decreased interest in rewarding stimuli—is a cardinal symptom in the diagnosis of depression. Computational theories of reward-guided learning hold that future reward expectations depend heavily on the difference between actual and expected reward outcomes, i.e., prediction errors (Rescorla and Wagner, 1972; Sutton and Barto, 1998). Neurally, a close link has long been noted between prediction error signals and the firing of dopaminergic neurons during associative learning (Schultz et al., 1997). For instance, one recent study used optogenetic techniques to demonstrate a causal relationship between dopamine and anhedonia (Tye et al., 2013). Here, optogenetic silencing of midbrain (VTA) dopaminergic neurons was shown to induce a lack of sucrose preference (a homolog of anhedonia) in mice, while optogenetic stimulation of the same neurons relieved anhedonia (Hayes, 2013; Tye et al., 2013). However, while animal models suggest that phasic prediction error signaling is impaired in anhedonia, a recent behavioral meta-analysis of human data suggests otherwise (Huys et al., 2013).

Computationally, anhedonia can arise through either a primary insensitivity to reward, or disrupted ability to learning about rewards (Huys et al., 2013). Huys et al. (2013) directly contrasted these alternatives, conducting a model-based Bayesian meta-analysis of six datasets where depressed patients completed a probabilistic reward learning experiment. They found that reward sensitivity⁴—but not learning—was impaired in MDD patients, but a dopamine agonist pramipexole showed the

⁴It may be important to clarify that while the Huys paper demonstrated impaired reward sensitivity for *wanting* in MDD patients, it seems that *liking* rewards remains intact in MDD (Dichter et al., 2010).

opposite pattern (Huys et al., 2013). Serotonin, on the other hand, is understood to selectively modulate behavioral and neural representations of reward value (Seymour et al., 2012). Specifically, it has been shown through acute depletion of a serotonin precursor (tryptophan) that serotonin depletion leads to impaired reward sensitivity in humans (Seymour et al., 2012). Other model-based differences in reward processing between depressed patients and controls have also been shown: depressed patients have blunted prediction error signals compared to healthy controls (Kumar et al., 2008; Gradin et al., 2011), and fail to adjust reaction times (e.g., post-error slowing) in the same way as control participants (Steele et al., 2007). As a brief but important note, either deficit would lead to inaccuracies in our recognition model: no longer faithfully reflecting reward causalities in our interactions with the world. Here we see important examples of how neuromodulation (serotonin and dopamine) can adversely affect beliefs about precision (sensitivity or gain and learning rates) to produce suboptimal inference and learning, respectively.

Seemingly abstract differences between computational quantities may carry important implications for the treatment of depression. Disruptions to neural representations of either reward sensitivity or reward learning would introduce inaccuracies to our generative model of the world. Indeed, a distorted mapping between actions and rewards could conceivably explain a number of depressive symptoms, particularly the feelings of hopelessness, distorted appetite, and anhedonia/decreased interest in pleasurable stimuli. However, the precise mechanism by which this occurs is critical to treatment strategies: while either impairment might lead to similar behavioral deficits, exactly which impairment a patient has carries implications for their treatment. At present, serotonin-targeting treatments are the first-line antidepressant method, but do not seem to alleviate depressive symptoms in many patients (Harmer and Cowen, 2013). After failed SSRI treatment, dopamine-targeting treatments can be attempted (Rush et al., 2006; Trivedi et al., 2006). It is entirely plausible that depressive symptoms, which broadly result from impaired reward processing, might stem from either, or both, impairments to serotonergic reward sensitivity and dopaminergic reward learning.

Perhaps, therefore, a reinforcement learning experiment might have predictive value over which treatment will be effective. If model-based analyses show a patient has a learning rate impairment, they may be more suited to dopaminergic treatment. If a patient has impaired reward sensitivity, perhaps serotonergic interventions ought to work. Considering behaviorally-dissociable recognition density distortions offers an interesting re-appraisal of inconsistent antidepressant success, with potential therapeutic implications. In fact, there is no reason to limit our investigation to these parameters. In one recent example (Diaconescu et al., 2014), sophisticated computational modeling was applied to a social learning task (based on Behrens et al., 2008) to investigate the mechanisms by which we infer the intentions of others. Such analyses characterize both social and non-social aspects of learning behavior extensively, and would enable researchers to consider potential abnormalities in MDD in a rich fashion. This kind of computational psychiatric approach is becoming increasingly popular, and has enjoyed

recent success across a range of disorders (Montague et al., 2012; Corlett and Fletcher, 2014; Stephan and Mathys, 2014), including psychosis (Corlett et al., 2009, 2011), borderline personality disorder (Fineberg et al., 2014), schizophrenia (Fletcher and Frith, 2009), and delusions (Moutoussis et al., 2011).

REWIRING BELIEFS

Since beliefs must be stored in the brain, using antidepressants to correct aberrant models of the world ought also to require some neurophysiological restructuring. This is in line with extant explanations for the delay in antidepressant efficacy. One early hypothesis for the delay in clinical effect of SSRIs argued that the desensitization of serotonin autoreceptors on serotonergic bodies and terminals is required before SSRIs can fully increase serotonin nervous transmission (Blier and de Montigny, 1994). In line with this suggestion, clinical trials combining SSRIs with serotonin autoreceptor antagonists have shown a faster and enhanced antidepressant effect (Whale et al., 2010). More recent alternative suggestions concerned the effects of SSRIs on neurotrophins and cellular processes generating new neurons and synapses. Animal models of depression highlight decreased BDNF production, neurogenesis, and synaptic plasticity: effects that are reversed by repeated administration of SSRIs (Santarelli et al., 2003; Castrén, 2005; Castrén and Rantamäki, 2010). Although interesting attempts have been made to apply the free-energy principle to monoaminergic (dopamine) transmission and inference (Friston et al., 2012), the present essay makes no prescriptions as to what specific neurobiological changes would happen as models become “less depressed.”

PSYCHOTHERAPY: BREAKING THE ACTION-PERCEPTION CYCLE?

Correcting representations (the perceptual side to free-energy) might be one way of treating a depressive model of the world, but it is not the only way. Earlier I described the notion of active inference, whereby an agent selectively samples the environment in line with its model of the world, using the intuitive example of wandering in the dark (Friston, 2010). Another way in which we can influence the model is by changing its inputs, that is, optimizing actions to sample the environment differently. For instance, it may be that interactions with certain people or objects are further enhancing depressive symptoms, and/or (conversely) that a lack of positive actions is having a similar effect, reminiscent of learned helplessness models of depression. There is evidence in line with this; several studies have noted depressed patients spend significantly longer looking at negative stimuli (Matthews and Antes, 1992; Eizenman et al., 2003; Caseras et al., 2007; Seth, 2013)—perhaps this excessive negative sampling is skewing the inputs to our models. But this notion also extends beyond an individual's physical actions; excessively sampling negative causal relationships might also distort an agent's model of the world. Indeed, this idea of active sampling concurs with recent theoretical work linking interoceptive inference and emotion, where emotion is held to emerge from cognitive appraisals of physiological states (Seth, 2013). One recent computational paper attempted to model emotional valence as the second time-derivative of free-energy, where emotional valence regulates the learning rate of the

causes of sensory inputs (Joffily and Coricelli, 2013). In more plain terms: “when sensations increasingly violate the agent’s expectations, [emotional] valence is negative and increases the learning rate. Conversely, when sensations increasingly fulfill the agent’s expectations, [emotional] valence is positive and decreases the learning rate” (Joffily and Coricelli, 2013). Put simply, active inference requires us to sample the world in accordance with our expectations. If expectations imply the world is a rather hostile place, then it will be sampled as such. There are clear analogies with learned helplessness models of depression here. Interestingly, learned helplessness can be (Bayes) optimal, if the world is indeed persistently hostile and has a low volatility.

Note again the crucial role of the learning rate in facilitating the (re)learning of a generative model. Under predictive coding, an implementation of the free-energy principle, the learning rate increases with the expected volatility of environmental contingencies, but volatility is only one factor influencing it (Mathys et al., 2011, 2014). Behaviorally, it has been shown that healthy human subjects assess volatility in an optimal manner—that is, increase their learning rate when the environment is more volatile (Behrens et al., 2007). In this study, the authors demonstrated that the optimal estimate of environmental volatility was reflected in the fMRI signal in the anterior cingulate cortex (ACC), and variations in this signal predicted between-subject variations in learning rate. Although no study has specifically investigated the ability of depressed patients to optimally update learning rate according to their environment, one study showed that controls—but not patients—significantly activated the ACC when given negative feedback during a gambling task (Gradin et al., 2011). Of course, this itself does not mean that ACC activity significantly differed between controls and patients (Gelman and Stern, 2006).

Nonetheless, it seems optimizing actions in order to change a model of the world is reflected in psychotherapy approaches. APA (2010)⁵ guidelines for depression psychotherapy include helping people “gradually incorporate enjoyable, fulfilling activities back into their lives,” and “improve patterns of interacting with other people that contribute to their depression,” both of which would constitute optimization of actions under the free-energy framework. Essentially, breaking any actions or sampling mechanisms that further a depressive model of the world appears to be a recommendation that the free-energy framework makes, that psychotherapy treatments have already taken.

FREE-ENERGY: A HOLISTIC APPROACH?

It is worth briefly setting this approach in the context of other accounts of depression. It is true that we already have elegant emotional/cognitive accounts of depression (Harmer and Cowen, 2013), and there are many putative biological explanations of disruption and restoration at cellular (Castrén and Rantamäki, 2010) and molecular (Berman et al.) levels. However, pharmacological level explanations often lose sight of the multidimensional nature of the depressive phenotype; and emotional or high-level explanations are difficult to relate directly back to the neurobiology. In fact, a predictive coding approach resembles a previous

information-processing level approach, illustrated in **Figure 1**, known as the “network hypothesis of depression” (Nestler et al., 2002; Castrén, 2005). At a minimum, this is encouraging: it suggests that the free-energy framework is largely consistent with theories of depression at multiple levels, and offers a plausible alternative that also unifies global brain theories in biological and physical sciences. In future, it may offer an opportunity for researchers to directly transition between depression’s many levels of research in a principled, model-based fashion (Montague et al., 2012; Friston et al., 2014; Stephan and Mathys, 2014). For example, there may be different underlying problems in MDD, with different behavioral ways of testing, with specific therapeutic implications.

It is also interesting to relate the free-energy approach to the literature on Depressive Realism, a claim that depressed people are sometimes better at evaluating instrumentality than non-depressed people (Alloy and Abramson, 1979; Alloy et al.). The claim appears robust, if small: a recent meta-analysis of 75 studies indicated a small overall depressive realism effect, although both depressed and non-depressed individuals showed a substantial “optimism bias” (Moore and Fresco, 2012). Some compelling model-driven research suggests the effect may be driven by contextual processing differences, rather than depressed individuals having consistently low expectation of control (Msetfi et al., 2005). This is partially supported by one recent pharmacological study showing that amongst a group of 15 non-depressed participants, acute tryptophan depletion improved contingency judgments for participants with particularly low scores on the Beck Depression Inventory (BDI < 6; Chase et al., 2011). In a free-energy view, “control” in the clinical psychological context corresponds to outcome entropy, and it directly influences an individual’s belief about what kinds of outcome distributions are likely. “Maladaptive” priors or generalization tendencies could equally result in differences in perceived control, although “maladaptive” here requires some clarification. Since both depressed and non-depressed individuals typically show an optimism bias, “maladaptive” is simply with reference to non-depressed individuals, rather than a comment on optimality. Although a detailed analysis of entropy and perceived control is beyond the scope of the current article, Huys and Dayan (2009) offer an excellent mathematical treatment of behavioral control from a Bayesian perspective.

The free-energy approach detailed in this review is not, however, an exhaustive account of depression. Symptoms of low mood and anhedonia may be cardinal symptoms in MDD but they are not the only ones: the accompanying loss of appetite, sleep disturbance, diurnal fluctuation, low energy and somatic symptoms are a key part of the illness. Furthermore, these additional symptoms can sometimes be the ones that are slowest to resolve. It is possible that wider symptoms may emerge as a behavioral consequence of a distorted generative model: for instance, if food rewards are no longer subjectively rewarding then loss of appetite or motivation to eat is understandable, if not predictable. In addition, although this review focused on the most common treatments for depression—monoaminergic antidepressants and psychotherapy—there is now preliminary evidence that intravenous administration of ketamine and other

⁵ Available online at: http://psychiatryonline.org/pb/assets/raw/sitewide/practice_guidelines/guidelines/mdd.pdf

glutamatergic drugs can have remarkably quick—but transient—antidepressant effects in unipolar and bipolar depression (aan het Rot et al., 2010; Aan Het Rot et al., 2012; McGirr et al., 2014). The speed of ketamine's antidepressant efficacy here may appear problematic for a free-energy interpretation at first glance. However, few treatments in psychiatry or medicine are effective after a single dose, and ketamine is no exception: patients often return to the depressed state without a course of treatment over a number of weeks (aan het Rot et al., 2010; McGirr et al., 2014). From a free-energy perspective, ketamine can be considered a faster vehicle for repairing representations, but one that nonetheless takes some time to repair the generative model. In addition, from a neurobiological perspective, ketamine's acute and sustained antidepressant effects have been hypothesized to depend on synaptogenesis (Li et al., 2010), in reminiscent fashion to monoaminergic antidepressants. Further insight comes from Bayesian treatments of psychosis using ketamine as a model (Corlett et al., 2009, 2011). Here, distinct influences have been proposed for ketamine in the short and long term. In the short term, it is thought that ketamine briefly disturbs cortical inference by blocking NMDA receptors, and impairing the specification of top-down prior expectancies (Corlett et al., 2011). With chronic ketamine use, however, there is a compensatory increase in the number and function of NMDA receptors; longer-lasting changes that can give way to a delusional phenotype and depressed mood rather than remission from depression (Morgan et al., 2010; Corlett et al., 2011).

CONCLUSION

Under the free-energy principle the brain is an active prediction engine that seeks to establish a model of the causal structure of our environment, and minimize long-term surprise. The brain makes inferences about causal relationships at many levels of abstraction, and there is growing neural evidence in line with this theory. If the brain does indeed represent a collection of beliefs about the causal structure of the world, then the depressed phenotype emerges from a collection of depressive beliefs. The two mechanisms by which free-energy is minimized (and perhaps, how agents survive) are by optimizing actions, and optimizing representations. The two are markedly reminiscent of depression's two main therapies: psychotherapy and antidepressants, respectively. Distorted representations of the world might stem from distortions in reward representation, and correcting these through monoaminergic interventions might be a solution to anhedonia symptoms in particular. Similarly, a distorted sampling mechanism may exacerbate depressed mood, and require psychotherapies in an attempt to break the spiral of self-defeating actions. Either way, solutions ought not to be immediate: beliefs are changed gradually through successive combinations of past experiences and current observations. Irrespective of the formal insights into putative pathophysiology in depression, it may be the case that the holistic (theoretical) framework on offer here may be useful in cognitive behavior therapy. In other words, it may provide a rationale for the conjoint use of psychotherapeutic and pharmacological approaches that could be useful for both the therapist and patient alike. One thing is clear: depression is a multi-faceted illness in which disruptions to beliefs,

emotions, perception and action are intertwined. Perhaps, therefore, our approach must intertwine beliefs, emotions, perceptions and actions accordingly.

ACKNOWLEDGMENTS

I would like to thank Prof. Karl Friston, as well as the two reviewers, for their detailed comments on the manuscript. In addition, I would like to thank Profs. Catherine Harmer, Phil Burnet, Jutta Joorman, and Gregory McCarthy for their thoughtful discussions and advice. A. M. Chekroud was supported by the Philip Wright Scholarship, offered by Wadham College (Oxford).

REFERENCES

- aan het Rot, M., Collins, K. A., Murrough, J. W., Perez, A. M., Reich, D. L., Charney, D. S., et al. (2010). Safety and efficacy of repeated-dose intravenous ketamine for treatment-resistant depression. *Biol. Psychiatry* 67, 139–145. doi: 10.1016/j.biopsych.2009.08.038
- Aan Het Rot, M., Zarate, C. A., Charney, D. S., and Mathew, S. J. (2012). Ketamine for depression: where do we go from here? *Biol. Psychiatry* 72, 537–547. doi: 10.1016/j.biopsych.2012.05.003
- Adams, R. A., Stephan, K. E., Brown, H. R., Frith, C. D., and Friston, K. J. (2013). The computational anatomy of psychosis. *Front. Psychiatry* 4:47. doi: 10.3389/fpsyt.2013.00047
- Alloy, L. B., and Abramson, L. Y. (1979). Judgment of contingency in depressed and nondepressed students: sadder but wiser? *J. Exp. Psychol. Gen.* 108, 441–485. doi: 10.1037/0096-3445.108.4.441
- Alloy, L. B., Kelly, K. A., Mineka, S., and Clements, C. M. (1990). "Comorbidity of anxiety and depressive disorders: a helplessness-hopelessness perspective," in *Comorbidity of Mood and Anxiety Disorders*, eds J. D. Maser and C. R. Cloninger (Washington, DC: American Psychiatric Press), 499–453.
- Andersen, I., Thielen, K., Bech, P., Nygaard, E., and Diderichsen, F. (2011). Increasing prevalence of depression from 2000 to 2006. *Scand. J. Public Health* 39, 857–863. doi: 10.1177/1403494811424611
- Behrens, T. E. J., Hunt, L. T., Woolrich, M. W., and Rushworth, M. F. S. (2008). Associative learning of social value. *Nature* 456, 245–249. doi: 10.1038/nature07538
- Behrens, T. E. J., Woolrich, M. W., Walton, M. E., and Rushworth, M. F. S. (2007). Learning the value of information in an uncertain world. *Nat. Neurosci.* 10, 1214–1221. doi: 10.1038/nn1954
- Berman, R. M., Sporn, J., Charney, D. S., and Mathew, S. J. (2009). "Principles of the pharmacotherapy of depression," in *Neurobiology of Mental Illness*, eds D. S. Charney and E. J. Nestler (New York, NY: Guilford), 491–515.
- Blier, P., and de Montigny, C. (1994). Current advances and trends in the treatment of depression. *Trends Pharmacol. Sci.* 15, 220–226. doi: 10.1016/0165-6147(94)90315-8
- Caseras, X., Garner, M., Bradley, B. P., and Mogg, K. (2007). Biases in visual orienting to negative and positive scenes in dysphoria: an eye movement study. *J. Abnorm. Psychol.* 116, 491–497. doi: 10.1037/0021-843X.116.3.491
- Castrén, E. (2005). Is mood chemistry? *Nat. Rev. Neurosci.* 6, 241–246. doi: 10.1038/nrn1629
- Castrén, E., and Rantamäki, T. (2010). The role of BDNF and its receptors in depression and antidepressant drug action: reactivation of developmental plasticity. *Dev. Neurobiol.* 70, 289–297. doi: 10.1002/dneu.20758
- Chase, H. W., Crockett, M. J., Msetfi, R. M., Murphy, R. A., Clark, L., Sahakian, B. J., et al. (2011). 5-HT modulation by acute tryptophan depletion of human instrumental contingency judgements. *Psychopharmacology (Berl.)* 213, 615–623. doi: 10.1007/s00213-010-1934-4
- Clark, A. (2013). Whatever next? Predictive brains, situated agents, and the future of cognitive science. *Behav. Brain Sci.* 36, 181–204. doi: 10.1017/S0140525X12000477
- Compton, W. M., Conway, K. P., Stinson, F. S., and Grant, B. F. (2006). Changes in the prevalence of major depression and comorbid substance use disorders in the United States between 1991–1992 and 2001–2002. *Am. J. Psychiatry* 163, 2141–2147. doi: 10.1176/ajp.2006.163.12.2141

- Corlett, P. R., and Fletcher, P. C. (2014). Computational psychiatry: a Rosetta Stone linking the brain to mental illness. *Lancet Psychiatry* 1, 399–402. doi: 10.1016/S2215-0366(14)70298-6
- Corlett, P. R., Frith, C. D., and Fletcher, P. C. (2009). From drugs to deprivation: a Bayesian framework for understanding models of psychosis. *Psychopharmacology* 206, 515–530. doi: 10.1007/s00213-009-1561-0
- Corlett, P. R., Honey, G. D., Krystal, J. H., and Fletcher, P. C. (2011). Glutamatergic model psychoses: prediction error, learning, and inference. *Neuropsychopharmacology* 36, 294–315. doi: 10.1038/npp.2010.163
- Dayan, P., Hinton, G. E., Neal, R. M., and Zemel, R. S. (1995). The Helmholtz machine. *Neural Comput.* 7, 889–904. doi: 10.1162/neco.1995.7.5.889
- Diaconescu, A. O., Mathys, C., Weber, L. A. E., Daunizeau, J., Kasper, L., Lomakina, E. I., et al. (2014). Inferring on the intentions of others by hierarchical Bayesian learning. *PLoS Comput. Biol.* 10:e1003810. doi: 10.1371/journal.pcbi.1003810
- Dichter, G. S., Smoski, M. J., Kampov-Polevoy, A. B., Gallop, R., and Garbutt, J. C. (2010). Unipolar depression does not moderate responses to the Sweet Taste Test. *Depress. Anxiety* 27, 859–863. doi: 10.1002/da.20690
- Eizenman, M., Yu, L. H., Grupp, L., Eizenman, E., Ellenbogen, M., Gemar, M., et al. (2003). A naturalistic visual scanning approach to assess selective attention in major depressive disorder. *Psychiatry Res.* 118, 117–128. doi: 10.1016/S0165-1781(03)00068-4
- Fineberg, S. K., Steinfeld, M., Brewer, J., and Corlett, P. R. (2014). A computational account of Borderline Personality Disorder: impaired predictive learning about self and others through bodily simulation. *Front. Psychiatry* 5:111. doi: 10.3389/fpsy.2014.00111
- FitzGerald, T. H. B., Dolan, R. J., and Friston, K. J. (2014). Model averaging, optimal inference, and habit formation. *Front. Hum. Neurosci.* 8:457. doi: 10.3389/fnhum.2014.00457
- Fletcher, P. C., and Frith, C. D. (2009). Perceiving is believing: a Bayesian approach to explaining the positive symptoms of schizophrenia. *Nat. Rev. Neurosci.* 10, 48–58. doi: 10.1038/nrn2536
- Friston, K. (2010). The free-energy principle: a unified brain theory? *Nat. Rev. Neurosci.* 11, 127–138. doi: 10.1038/nrn2787
- Friston, K., and Kiebel, S. (2009). Predictive coding under the free-energy principle. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 364, 1211–1221. doi: 10.1098/rstb.2008.0300
- Friston, K., Kilner, J., and Harrison, L. (2006). A free energy principle for the brain. *J. Physiol. Paris* 100, 70–87. doi: 10.1016/j.jphysparis.2006.10.001
- Friston, K. J., Daunizeau, J., and Kiebel, S. J. (2009). Reinforcement learning or active inference? *PLoS ONE* 4:e6421. doi: 10.1371/journal.pone.0006421
- Friston, K. J., Shiner, T., FitzGerald, T., Galea, J. M., Adams, R., Brown, H., et al. (2012). Dopamine, affordance and active inference. *PLoS Comput. Biol.* 8:e1002327. doi: 10.1371/journal.pcbi.1002327
- Friston, K. J., Stephan, K. E., Montague, R., and Dolan, R. J. (2014). Computational psychiatry: the brain as a phantastic organ. *Lancet Psychiatry* 1, 148–158. doi: 10.1016/S2215-0366(14)70275-5
- Gelman, A., and Stern, H. (2006). The difference between “significant” and “not significant” is not itself statistically significant. *Am. Stat.* 60, 328–331. doi: 10.1198/000313006X152649
- Gradin, V. B., Kumar, P., Waiter, G., Ahearn, T., Stickler, C., Milders, M., et al. (2011). Expected value and prediction error abnormalities in depression and schizophrenia. *Brain* 134, 1751–1764. doi: 10.1093/brain/awr059
- Harmer, C. J., and Cowen, P. J. (2013). “It’s the way that you look at it”—a cognitive neuropsychological account of SSRI action in depression. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 368:20120407. doi: 10.1098/rstb.2012.0407
- Hayes, G. (2013). Dynamic and diverse roles of dopamine. *Cell* 152, 663–665. doi: 10.1016/j.cell.2013.01.045
- Huys, Q. J. M., and Dayan, P. (2009). A Bayesian formulation of behavioral control. *Cognition* 113, 314–328. doi: 10.1016/j.cognition.2009.01.008
- Huys, Q. J., Pizzagalli, D. A., Bogdan, R., and Dayan, P. (2013). Mapping anhedonia onto reinforcement learning: a behavioural meta-analysis. *Biol. Mood Anxiety Disord.* 3:12. doi: 10.1186/2045-5380-3-12
- Joffily, M., and Coricelli, G. (2013). Emotional valence and the free-energy principle. *PLoS Comput. Biol.* 9:e1003094. doi: 10.1371/journal.pcbi.1003094
- Knill, D. C., and Pouget, A. (2004). The Bayesian brain: the role of uncertainty in neural coding and computation. *Trends Neurosci.* 27, 712–719. doi: 10.1016/j.tins.2004.10.007
- Kumar, P., Waiter, G., Ahearn, T., Milders, M., Reid, I., and Steele, J. D. (2008). Abnormal temporal difference reward-learning signals in major depression. *Brain* 131, 2084–2093. doi: 10.1093/brain/awn136
- Li, N., Lee, B., Liu, R.-J., Banasr, M., Dwyer, J. M., Iwata, M., et al. (2010). mTOR-dependent synapse formation underlies the rapid antidepressant effects of NMDA antagonists. *Science* 329, 959–964. doi: 10.1126/science.1190287
- Ma, W. J. (2012). Organizing probabilistic models of perception. *Trends Cogn. Sci.* 16, 511–518. doi: 10.1016/j.tics.2012.08.010
- Mathys, C., Daunizeau, J., Friston, K. J., and Stephan, K. E. (2011). A bayesian foundation for individual learning under uncertainty. *Front. Hum. Neurosci.* 5:39. doi: 10.3389/fnhum.2011.00039
- Mathys, C. D., Lomakina, E. I., Daunizeau, J., Iglesias, S., Brodersen, K. H., Friston, K. J., et al. (2014). Uncertainty in perception and the Hierarchical Gaussian Filter. *Front. Hum. Neurosci.* 8:825. doi: 10.3389/fnhum.2014.00825
- Matthews, G. R., and Antes, J. R. (1992). Visual attention and depression: cognitive biases in the eye fixations of the dysphoric and the nondepressed. *Cogn. Ther. Res.* 16, 359–371. doi: 10.1007/BF01183287
- McGirr, A., Berlim, M. T., Bond, D. J., Fleck, M. P., Yatham, L. N., and Lam, R. W. (2014). A systematic review and meta-analysis of randomized, double-blind, placebo-controlled trials of ketamine in the rapid treatment of major depressive episodes. *Psychol. Med.* 45, 693–704. doi: 10.1017/S0033291714001603
- Montague, P. R., Dolan, R. J., Friston, K. J., and Dayan, P. (2012). Computational psychiatry. *Trends Cogn. Sci.* 16, 72–80. doi: 10.1016/j.tics.2011.11.018
- Moore, M. T., and Fresco, D. M. (2012). Depressive realism: a meta-analytic review. *Clin. Psychol. Rev.* 32, 496–509. doi: 10.1016/j.cpr.2012.05.004
- Morgan, C. J. A., Muetzelfeldt, L., and Curran, H. V. (2010). Consequences of chronic ketamine self-administration upon neurocognitive function and psychological wellbeing: a 1-year longitudinal study. *Addiction* 105, 121–133. doi: 10.1111/j.1360-0443.2009.02761.x
- Moutoussis, M., Bentall, R. P., El-Derey, W., and Dayan, P. (2011). Bayesian modelling of Jumping-to-Conclusions bias in delusional patients. *Cogn. Neuropsychiatry* 16, 422–447. doi: 10.1080/13546805.2010.548678
- Msetfi, R. M., Murphy, R. A., Simpson, J., and Kornbrot, D. E. (2005). Depressive realism and outcome density bias in contingency judgments: the effect of the context and intertrial interval. *J. Exp. Psychol. Gen.* 134, 10–22. doi: 10.1037/0096-3445.134.1.10
- Nestler, E. J., Barrot, M., DiLeone, R. J., Eisch, A. J., Gold, S. J., and Monteggia, L. M. (2002). Neurobiology of depression. *Neuron* 34, 13–25. doi: 10.1016/S0896-6273(02)00653-0
- Rao, R. P., and Ballard, D. H. (1999). Predictive coding in the visual cortex: a functional interpretation of some extra-classical receptive-field effects. *Nat. Neurosci.* 2, 79–87. doi: 10.1038/4580
- Rescorla, R. A., and Wagner, A. R. (1972). “A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement,” in *Classical Conditioning II: Current Research and Theory*, Vol. 2, eds A. H. Black and W. F. Prokasy (New York, NY: Appleton-Century-Crofts), 64–99.
- Rush, A. J., Trivedi, M. H., Wisniewski, S. R., Stewart, J. W., Nierenberg, A. A., Thase, M. E., et al. (2006). Bupropion-SR, sertraline, or venlafaxine-XR after failure of SSRIs for depression. *N. Engl. J. Med.* 354, 1231–1242. doi: 10.1056/NEJMoa052963
- Santarelli, L., Saxe, M., Gross, C., Surget, A., Battaglia, F., Dulawa, S., et al. (2003). Requirement of hippocampal neurogenesis for the behavioral effects of antidepressants. *Science* 301, 805–809. doi: 10.1126/science.1083328
- Schultz, W., Dayan, P., and Montague, P. R. (1997). A neural substrate of prediction and reward. *Science* 275, 1593–1599. doi: 10.1126/science.275.5306.1593
- Seth, A. K. (2013). Interoceptive inference, emotion, and the embodied self. *Trends Cogn. Sci.* 17, 565–573. doi: 10.1016/j.tics.2013.09.007
- Seymour, B., Daw, N. D., Roiser, J. P., Dayan, P., and Dolan, R. (2012). Serotonin selectively modulates reward value in human decision-making. *J. Neurosci.* 32, 5833–5842. doi: 10.1523/JNEUROSCI.0053-12.2012
- Steele, J. D., Kumar, P., and Ebmeier, K. P. (2007). Blunted response to feedback information in depressive illness. *Brain* 130, 2367–2374. doi: 10.1093/brain/awn150
- Stephan, K. E., and Mathys, C. (2014). Computational approaches to psychiatry. *Curr. Opin. Neurobiol.* 25, 85–92. doi: 10.1016/j.conb.2013.12.007

- Sutton, R. S., and Barto, A. G. (1998). *Reinforcement Learning*. Cambridge, MA: MIT Press.
- Trivedi, M. H., Fava, M., Wisniewski, S. R., Thase, M. E., Quitkin, F., Warden, D., et al. (2006). Medication augmentation after the failure of SSRIs for depression. *N. Engl. J. Med.* 354, 1243–1252. doi: 10.1056/NEJMoa052964
- Tye, K. M., Mirzabekov, J. J., Warden, M. R., Ferenczi, E. A., Tsai, H.-C., Finkelstein, J., et al. (2013). Dopamine neurons modulate neural encoding and expression of depression-related behaviour. *Nature* 493, 537–541. doi: 10.1038/nature11740
- von Helmholtz, H. (1924). *Helmholtz's Treatise on Physiological Optics*. Menasha, WI: George Banta Publishing Company.
- Whale, R., Terao, T., Cowen, P., Freemantle, N., and Geddes, J. (2010). Pindolol augmentation of serotonin reuptake inhibitors for the treatment of depressive disorder: a systematic review. *J. Psychopharmacol.* 24, 513–520. doi: 10.1177/0269881108097714
- Conflict of Interest Statement:** The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.
- Received: 14 November 2014; accepted: 30 January 2015; published online: 20 February 2015.
- Citation: Chekroud AM (2015) Unifying treatments for depression: an application of the Free Energy Principle. *Front. Psychol.* 6:153. doi: 10.3389/fpsyg.2015.00153
- This article was submitted to *Psychology for Clinical Settings*, a section of the journal *Frontiers in Psychology*.
- Copyright © 2015 Chekroud. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



The neuroscience of positive memory deficits in depression

Daniel G. Dillon*

Motivated Learning and Memory Laboratory, Center for Depression, Anxiety and Stress Research, McLean Hospital, Harvard Medical School, Belmont, MA, USA

OPEN ACCESS

Edited by:

Frank Ryan,
Imperial College London, UK

Reviewed by:

Carlos E. Norte,
Federal University of Rio de Janeiro,
Brazil

Roger C. McIntosh,
University of Miami, USA

*Correspondence:

Daniel G. Dillon,
Motivated Learning and Memory
Laboratory, Center for Depression,
Anxiety and Stress Research, McLean
Hospital, Harvard Medical School,
115 Mill Street, Belmont,
MA 02478, USA
ddillon@mclean.harvard.edu

Specialty section:

This article was submitted to
Psychology for Clinical Settings,
a section of the journal
Frontiers in Psychology

Received: 29 June 2015

Accepted: 13 August 2015

Published: 07 September 2015

Citation:

Dillon DG (2015) The neuroscience
of positive memory deficits
in depression.
Front. Psychol. 6:1295.
doi: 10.3389/fpsyg.2015.01295

Adults with unipolar depression typically show poor episodic memory for positive material, but the neuroscientific mechanisms responsible for this deficit have not been characterized. I suggest a simple hypothesis: weak memory for positive material in depression reflects disrupted communication between the mesolimbic dopamine pathway and medial temporal lobe (MTL) memory systems during encoding. This proposal draws on basic research showing that dopamine release in the hippocampus is critical for the transition from early- to late-phase long-term potentiation (LTP) that marks the conversion of labile, short-term memories into stable, long-term memories. Neuroimaging and pharmacological data from healthy humans paint a similar picture: activation of the mesolimbic reward circuit enhances encoding and boosts retention. Unipolar depression is characterized by anhedonia—loss of pleasure—and reward circuit dysfunction, which is believed to reflect negative effects of stress on the mesolimbic dopamine pathway. Thus, I propose that the MTL is deprived of strengthening reward signals in depressed adults and memory for positive events suffers accordingly. Although other mechanisms are important, this hypothesis holds promise as an explanation for positive memory deficits in depression.

Keywords: depression, reward, memory, anhedonia, hippocampus

Introduction

Unipolar depression impairs episodic memory (Burt et al., 1995; Zakzanis et al., 1998), and the dominant hypothesis is that stress-induced changes in the hippocampus are responsible (Sapolsky, 1996, 2000; Sheline et al., 1999; MacQueen et al., 2003; MacQueen and Frodl, 2011; Huang et al., 2013; Travis et al., 2014). The idea is straightforward: stress is a potent risk factor for depression (Kendler et al., 1999; Monroe and Harkness, 2005), and the dense concentration of glucocorticoid receptors in the human hippocampus (Wang et al., 2013) makes it a prime candidate for stress-induced neurotoxicity¹. Indeed, the hippocampus is smaller in adults with recurrent depression (MacQueen and Frodl, 2011), and post-mortem exams reveal shrunken hippocampal neurons and glial cells in depressed individuals (Stockmeier et al., 2004). Because the hippocampus is the seat of episodic memory (Squire, 1992), a causal chain running from stress to depression to hippocampal volume reductions provides an appealing account of memory deficits in depression.

However, there are reasons to think the hippocampal stress hypothesis would benefit from supplementation. First, although adults with recurrent depression typically show memory deficits and hippocampal volume reductions, strong evidence for a direct relationship between these two phenomena is lacking (MacQueen et al., 2003; Travis et al., 2014). Second, research on

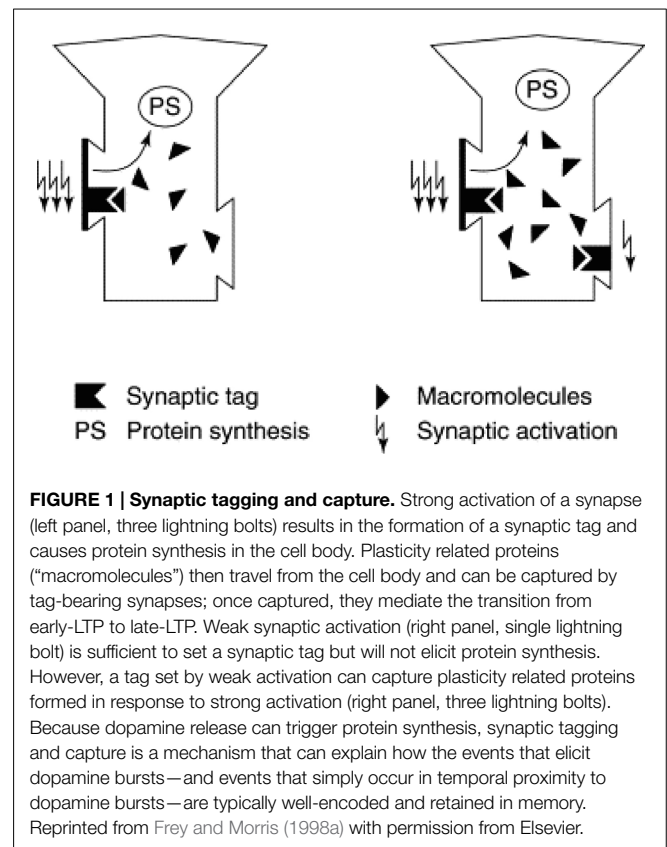
¹For brevity, I will use the terms “depression” and “memory” to refer to unipolar depression and episodic memory, respectively, throughout the remainder of the manuscript unless otherwise noted.

emotional memory suggests an important role for neural mechanisms implicated in positive emotional responses. Excessive sadness is one of two cardinal symptoms of major depressive disorder (MDD; American Psychiatric Association, 2013), and since there is a tendency for depressed adults to preferentially attend to and ruminate on negative information (Beck et al., 1979; Nolen-Hoeksema, 1991), one might expect better memory for negative material in depressed adults, as information consistent with one's mood is preferentially encoded (Bower, 1981, 1987). Indeed, this effect is often found (e.g., Gotlib et al., 2004; Hamilton and Gotlib, 2008; Matt et al., 1992)—but positive memory deficits are also robust. Healthy adults often show better memory for positive versus negative or neutral material, but this advantage is reduced in depression (e.g., Gotlib et al., 2004; Hamilton and Gotlib, 2008), and a seminal meta-analysis found that this group difference is more reliable than enhanced memory for negative material in depression (Burt et al., 1995). Why is memory for positive material impaired in depressed adults? I propose that it reflects anhedonia—the second cardinal symptom of MDD—and its association with dysfunction in mesolimbic dopamine circuits that respond to reward (Schultz, 1998).

In this article I review key studies from cellular, behavioral, and human neuroscience that underscore the critical role of dopamine transmission in the persistence of episodic memories (for more extensive reviews, see Frey and Morris, 1998a; Lisman and Grace, 2005; Shohamy and Adcock, 2010; Lisman et al., 2011). When considered alongside growing evidence of reward system dysfunction in depression (Eshel and Roiser, 2010; Treadway and Zald, 2011; Dillon et al., 2014b), these data invite the following inference: memory for positive material is impaired in depressed adults because, on average, mesolimbic dopamine circuits do not mount an adequate response to reward in such individuals, compromising interactions between reward and memory systems that ensure memory retention. As stress is the most likely cause of weak dopaminergic reward responses in depression (Dillon et al., 2014b; Pizzagalli, 2014), this hypothesis extends the existing literature: stress cannot only induce hippocampal volume reductions, it can also perturb reward circuits and preferentially disrupt the formation of positive memories.

Synaptic Tagging and Capture

The synaptic tagging and capture (STC) hypothesis forms the foundation of this proposal (Frey and Morris, 1997, 1998a,b; Frey and Frey, 2008). STC solves a routing problem confronted by the brain's cellular learning and memory mechanism, long-term potentiation (LTP; Bliss and Lømo, 1973). To appreciate the nature of the problem, consider a population of hippocampal neurons communicating across a synapse. If an experimenter applies weak electrical stimulation to the pre-synaptic neurons and measures the post-synaptic response, she will observe an increase in activation (specifically, excitatory field potentials) that will decay back to baseline within about 3 to 6 h; this transient response is called early-LTP (Frey and Morris, 1997). But if she applies strong stimulation, she will record increased post-synaptic activation that can be maintained for days, weeks,



or months (Abraham, 2003); this sustained response is called late-LTP. Late-LTP reflects the operation of molecular processes that structurally remodel the connection between pre- and post-synaptic neurons, transforming a country path into an eight-lane highway and making information trafficking easier (Baudry et al., 2011). The routing problem arises because most of the plasticity-related proteins (PRPs) needed for remodeling are synthesized in the body of the neuron, but LTP is synapse-specific. There are thousands of synapses per neuron (Pakkenberg et al., 2003), all of them far downstream from the cell body, so getting PRPs to the right synapses is a significant challenge. In other words, if late-LTP amounts to building a bridge between neurons, then the concrete is mixed at a rural plant and must be delivered to a construction site in a busy neighborhood downtown—how does the delivery driver find his way?

Synaptic tagging and capture provides an appealing answer (Figure 1). It turns out that even weak activation is sufficient to set molecular tags that mark a synapse as a candidate for strengthening. If only weak activation occurs, PRPs will not be synthesized and the tags will fade away, consistent with the transient nature of early-LTP (Frey and Morris, 1998a,b; Rogerson et al., 2014). However, if PRPs are synthesized they can begin their trek down the axon without a determined destination, simply stopping at any synapse that displays a tag. When they do so, structural remodeling occurs and a stronger synaptic connection is in place: late-LTP.

This raises another issue. Neurons use an on-demand inventory system and synthesize PRPs as they are needed. In the laboratory,

strong electrical stimulation drives PRP synthesis, but what triggers their production in nature?

Dopamine is critical (Frey and Morris, 1998a; Smith et al., 2005; Lisman et al., 2011). Direct evidence for this claim comes from an *in vitro* study that used fluorescence imaging to track the production of new proteins in hippocampal neurons (Smith et al., 2005). This work showed that applying D1/D5 dopamine receptor *agonists* to the hippocampus results in increased protein synthesis, an effect that can be blocked by applying D1/D5 receptor *antagonists*. Furthermore, the newly synthesized proteins included a subunit of the AMPA receptor, which is a key contributor to late-LTP: increased post-synaptic AMPA receptor density is a major part of the “bridge” between neurons (Malinow, 2003). Finally, if dopamine-driven protein synthesis contributes to LTP, then one would expect application of D1/D5 receptor agonists and antagonists to influence the post-synaptic response to pre-synaptic stimulation in opposite directions. Indeed, this was observed: for a given level of stimulation, application of the agonists doubled the response frequency of post-synaptic neurons, but application of the antagonists blocked this effect. This is compelling evidence that hippocampal dopamine release is crucial for the synthesis of proteins that mediate late-LTP.

There is also a wealth of evidence regarding the role of dopamine in the synthesis of PRPs from studies that inferred their presence (or absence) by examining late-LTP. For instance: (1) the concentration of hippocampal dopamine increases following late-LTP induction (Frey et al., 1990); (2) the application of D1/D5 receptor agonists can directly induce late-LTP, skipping early-LTP entirely (Huang and Kandel, 1995); and (3) the oral administration of L-DOPA—a dopamine precursor used to treat Parkinson’s Disease—lowers the stimulation necessary to transition from early to late-LTP in the rodent hippocampus, as does application of a D1/D5 receptor agonist (Kusuki et al., 1997). All of these effects can be blocked by administration of D1/D5 receptor antagonists or protein synthesis inhibitors.

In summary, dopamine release drives PRP synthesis, which enables the transition from early to late-LTP (Frey and Morris, 1998a). Critically, any event capable of driving dopamine release is also strong enough to place tags on the synapses it activates. Since PRPs are sequestered by tag-bearing synapses, this is a mechanism for memory formation: dopamine release stabilizes LTP for the events that caused its release. If dopamine release is disrupted, memory for those events will suffer accordingly. Because many of the events that cause dopamine release also elicit positive emotional responses, disrupting the mesolimbic dopamine circuit should preferentially impair long-term memory for emotionally positive events.

Dopamine Supports the Retention of Episodic Memory in Non-human Animals

One could accept the findings reviewed above and yet question whether dopamine is relevant to episodic memory. After all, most of the work just described was conducted *in vitro* rather than in behaving animals. Even if dopamine proves important for hippocampal function in rodents, classic accounts of episodic memory hinge on conscious experience (Tulving, 1993). Given the

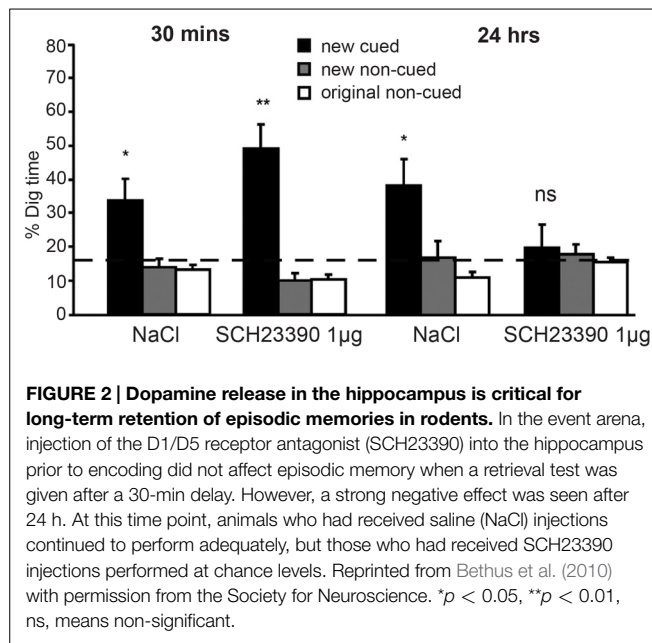
challenges associated with assessing consciousness in lab rats, one could fairly ask: Can we study episodic memory in animals?

The answer is yes. An alternative approach to episodic memory conceptualizes it not in terms of consciousness, but rather as memory for events-in-context: not only knowing that an event happened, but also knowing the spatial and temporal circumstances in which it happened (Clayton and Dickinson, 1998; Allen and Fortin, 2013). In other words, memory for an episode, defined as events taking place in a certain space or time, or in a particular sequence (MacDonald et al., 2013).

An elegant study used a method that met these criteria to show that memory persistence depends on the activation of hippocampal dopamine receptors (Bethus et al., 2010). The study used a sand-filled rectangular “event arena” with six wells in which rats could dig. On each training trial, rats were placed in one of four start boxes and given a food pellet with one of six flavors. The rat was then allowed to enter the arena and search for food. Critically, the “cue” pellet given in the start box determined which well contained additional pellets. The pairing of cues to wells was stable, and after 16 days the rats were running to the correct wells on about 80% of trials. This degree of accuracy is especially impressive because the use of four start boxes located in different places meant that a strategy based on landmarks would necessarily fail. Instead, the animals must have developed a map into which the cue-well associations were embedded.

With training complete, the experimenters tested episodic memory by introducing a novel cue flavor and placing two new wells in the arena, along with the original six. Only one of the two new wells was loaded with pellets. The rats explored the arena until they found the pellets, thus encoding a new cue-well association. Next, an identical trial was used to test memory, following delays of either 30 min or 24 h. Rats showed excellent memory regardless of the delay, running to the new cued well and avoiding both the new uncued well and the original six wells. However, performance after 24 h was at chance following hippocampal lesions administered after training but prior to encoding (Tse et al., 2007), and a similar impairment was observed when NMDA receptor blockers were injected into the hippocampus (Bethus et al., 2010). Because activation of NMDA receptors is essential to LTP (Lynch, 2004), these results indicate that 24-h memory for material learned in the event arena requires hippocampal LTP. In other words, performance in this task displays the hallmarks of episodic memory: rapid encoding of events-in-context, with long-term retention dependent on LTP in the hippocampus.

Having established that (episodic) memory for novel cue-well pairs depends on hippocampal LTP, the experimenters next demonstrated a critical role for dopamine in memory persistence (Figure 2). When they injected D1/D5 receptor antagonists into the hippocampus prior to encoding and tested retrieval after a 30-min delay, performance was unaffected. However, delaying the retrieval test by 24 h revealed a profound impairment, with performance no better than chance: given the new cue, rats ran to the uncued well as often as to the cued well (importantly, memory for the original six cue-well pairings was unimpaired). Control experiments showed that this effect did not reflect state-dependent retrieval: if the antagonists were injected into the hippocampus at encoding then 24 h memory was impaired



whether or not the antagonists were also injected at retrieval, and injecting the antagonists only at retrieval had no effect. In other words, using a drug to block dopamine release in the hippocampus at encoding impaired long-term memory, and this effect could not be rescued by injecting the drug again just prior to the memory test. These data extend *in vitro* studies reviewed in the prior section to behavior *in vivo*: just as dopamine is more important for late than early LTP, so dopamine release in the hippocampus is more important for long-term versus short-term retention of episodic memories (for additional evidence, see Wang et al., 2010).

One need not inject dopaminergic agents to observe these effects, as well-chosen behavioral interventions yield similar results. For instance, novelty elicits mesolimbic dopamine release in non-human animals (Lisman and Grace, 2005), and exposure to novelty enhances memory persistence in the event arena. In a modified version of the task, rats encoded a new association between a cue flavor and the location of a single well—the only well available—before completing a cued retrieval test in which six wells were available (Wang et al., 2010). This procedure was repeated daily for 6 months, with a different cue-well association learned and tested each day (the experimenters drew an analogy to parking one's car in different spots within the same lot). When a single pellet served as the reward for finding the correct well during learning, retrieval for cue-well pairings was adequate after a 30-min delay but decayed substantially after 24 h. However, 24-h memory was rescued by a simple manipulation: exposing the rats to a novel environment 30 min after learning.

Synaptic tagging and capture can account for this finding. Encoding the cue-well pairing results in synaptic activation that is sufficient to set a tag but not strong enough to drive PRP synthesis because of the meager reward (1 pellet) delivered for finding the correct well. Consequently, if nothing else happens, late-LTP will not occur and the cue-well memory will fade away within 24 h. However, exposure to a novel context drives dopamine release in the mesolimbic circuit, and this is sufficient to trigger PRP

synthesis. The molecular tags set during cue-well learning will not fade in the 30 min post-encoding, so when novelty exposure triggers dopamine release and PRP synthesis, the tags will still be set and can sequester PRPs, leading to late-LTP and robust 24 h memory. This account was strengthened by the fact that the positive effect of novelty on memory retention was blocked if novelty exposure was accompanied by the injection of D1/D5 receptor antagonists or protein synthesis inhibitors; either of these agents can disrupt PRP synthesis and thus deny the synapse a chance at late-LTP.

Note that enhanced memory persistence following a post-encoding manipulation (here, novelty exposure) is consistent with STC: as long as synaptic tags are present when PRPs are made available, the transition from early to late-LTP will occur and a lasting memory will be formed, regardless of whether the PRPs are synthesized before or after the tags are set (Frey and Morris, 1998a). STC also predicts a time window governed by the decay of the tags, and this study found evidence for such a window: novelty exposure 6 h post-encoding did not rescue 24-h memory and injecting protein synthesis inhibitors 6 h post-encoding did not impair 24-h memory. Thus, tag setting and PRP capture are complete within 6 h after encoding (other experiments find evidence for considerably more narrow windows, e.g., 2 h in Moncada and Viola, 2007).

Finally, using larger reward (3 pellets) at encoding had the same effect as novelty exposure: memory accuracy was sustained at the 24-h test, and injection of D1/D5 receptor antagonists into the hippocampus blocked this effect. Remarkably, this blockade of “strong” reward memory could be prevented by novelty exposure. When rats explored a novel environment prior to completing encoding trials in which 3 pellet reward were given, 24-h memory was intact even if encoding occurred under dopamine receptor blockade. This confirms a central prediction of STC: once PRPs are made available—here, by novelty exposure—they can be captured by tags that are set shortly afterward, even if PRP production is blocked during tag setting (see also Moncada and Viola, 2007; Ballarini et al., 2009).

In summary, rodents can form episodic memories, the persistence of which depends on dopamine release into the hippocampus. Furthermore, memory strength can be bolstered by behavioral manipulations that trigger dopamine release, such as the opportunity to explore a new environment or the receipt of 3 pellets rather than one for successful encoding. The next section highlights the fact that these manipulations have similar effects in humans.

Anticipation of Reward and Novelty Support Episodic Memory in Healthy Humans

A role for dopamine in episodic memory formation is consistent with anatomical studies in non-human animals (Lisman and Grace, 2005). In the rodent, there are direct projections from the ventral tegmental area (VTA) to the hippocampus, along with indirect projections from the hippocampus to the VTA that go through the nucleus accumbens and pallidum. The VTA-to-hippocampus connection permits the dopaminergic modulation

of hippocampal LTP that has been discussed thus far. Meanwhile, the hippocampal-to-VTA connection is critical for triggering dopamine bursts in the first instance. The hippocampus holds a representation of the current context and detects deviations from that context. When a novel stimulus or an unexpected reward is detected, the hippocampus registers the deviation and transmits that information to the VTA, leading to burst firing of dopamine neurons.

Do humans have the same pathways? Definitive anatomical studies have not been done, but an investigation of spontaneous functional connectivity in fMRI data is suggestive (Kahn and Shohamy, 2013). In two large samples ($n = 100$ and $n = 894$), analysis of resting state fMRI signals revealed significant correlations among the body of the hippocampus, the nucleus accumbens, and the VTA. Although functional connectivity does not imply anatomical connectivity, the fact that correlated activation among these regions was detectable at rest is encouraging because it implies synchronous patterns of activation in the absence of external stimulation, which suggests stable communication between these regions in humans.

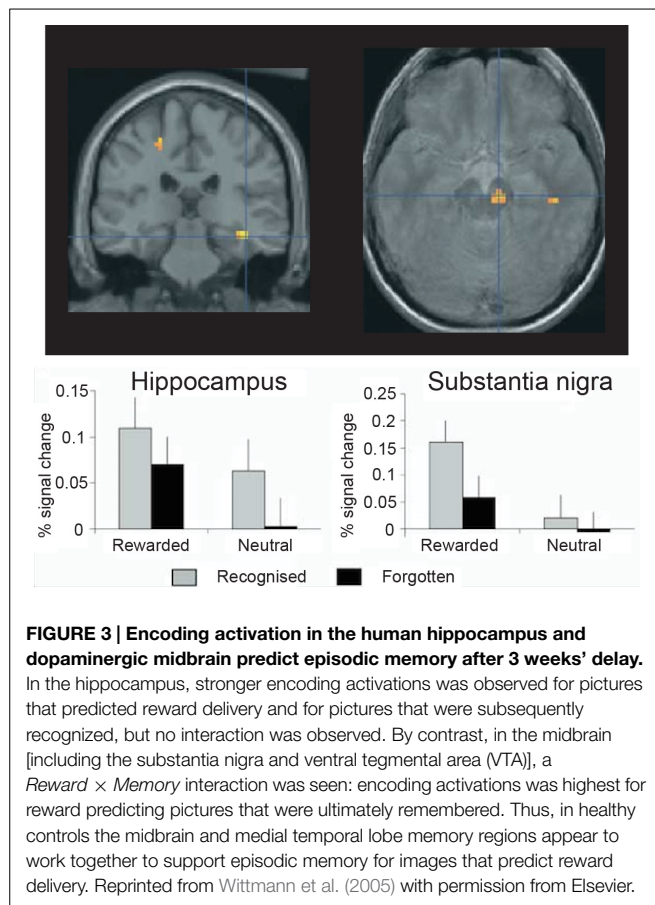
Furthermore, pharmacological work in humans has shown that dopaminergic agents can influence episodic memory. Specifically, enhancing dopamine transmission improves memory in a manner that appears consistent with STC. Chowdhury et al. (2012) administered L-DOPA to healthy older adults prior to an encoding session in which they viewed two categories of images, with one category reliably predicting delivery of monetary reward. Memory for half the items was tested after a 2-h delay, with memory for the remaining items tested after a 6-h delay. The study generated a striking finding—namely, a quadratic relationship between L-DOPA levels and delayed memory for neutral images (i.e., images that did not predict reward delivery), such that a moderate dose of L-DOPA improved 6-h memory for the neutral images relative to small or large doses. No such curve was evident after the shorter delay or for reward-predicting images at either delay. This is intriguing because 6-h memory for neutral images is the condition in which maximal forgetting would be expected, and thus it is also the condition in which heightened levels of dopamine could most easily rescue performance, much in the way that novelty exposure rescued 24-h memory in the 1-pellet encoding condition tested by Wang et al. (2010), as described earlier.

Additional pharmacological studies in humans are needed, as most of the human data pertinent to dopamine and memory come from task-based fMRI research (for review, see Shohamy and Adcock, 2010). Of course, fMRI cannot directly measure dopamine transmission and interpretations should be cautiously made (but for evidence of a relationship between dopamine receptor occupancy and fMRI signal from simultaneous fMRI/PET, see Mandeville et al., 2013). Nonetheless, the parallels with rodent data are striking. The earliest work on this topic demonstrated activation of the hippocampus and midbrain, including the VTA and substantia nigra (SN), in response to novel configurations of familiar images and during the encoding of successfully recalled words (Schott et al., 2004). Robust hippocampal activation was also observed, consistent with the hypothesis that these two regions form a functional unit in humans as well as in rodents.

The interpretation of these data is predicated on the hypothesis that the human mesolimbic dopamine network responds to novelty in much the same way as it responds to reward. A subsequent study from the same team provided compelling evidence for this hypothesis. The dopaminergic midbrain fires strongly to reward-predicting cues and to unexpected reward delivery (Schultz, 1998). To determine whether the human midbrain shows similar functionality with respect to novelty, Wittmann et al. (2007) presented participants with two colored squares that predicted the appearance of novel and familiar images, respectively. The predictions were accurate 75% of the time; on the remaining 25% of trials, unexpected novel or familiar pictures were presented. As hypothesized, fMRI data from the VTA/SN showed a strong response to the novelty-predicting cue and to the unexpected delivery of novel pictures following the familiarity-predicting cue. This is the pattern expected from the reward literature. Meanwhile, the bilateral hippocampus showed a strong response to the cue predicting novel (versus familiar) images, and activation of the VTA/SN and right hippocampus in response to the novelty-predicting cue was correlated across participants. Finally, memory was tested after a 24-h delay, and higher rates of recollection versus familiarity were observed for expected versus unexpected novel pictures. Taken together, the fMRI and behavioral data suggest that coactivation of the VTA/SN and hippocampus during novelty anticipation—prior to image presentation—facilitated encoding success.

Additional fMRI research has found evidence consistent with this interpretation, although it has used monetary reward rather than novel images to drive the VTA/SN (Figure 3). For example, in another study, Wittmann et al. (2005) presented images from two categories, only one of which reliably predicted a chance to win money, and then tested memory for the images at two times: immediately following encoding and again 3 weeks later. The fMRI data showed a strong VTA/SN response to the reward-predicting images, which were better remembered than the non-rewarded images when memory was tested after 3 weeks but not when memory was tested immediately. Moreover, when the experimenters probed brain regions whose encoding activation predicted 3-week memory, they found activation in the VTA/SN and the hippocampus. Furthermore, VTA/SN responses showed a *Reward* \times *Memory* interaction: activation was higher for subsequently remembered images that predicted reward delivery relative to both images that (1) did not predict reward and (2) predicted reward but were ultimately forgotten. These data indicate that long-term memory is supported by encoding activation in the hippocampus and the dopaminergic midbrain, and they highlight reward anticipation as a potent means for driving the human VTA/SN (see also Adcock et al., 2006; Wolosin et al., 2012, 2013).

Why does the prospect of earning a reward influence encoding? Murty and Adcock (2014) advance the following argument with respect to the incidental encoding of task-irrelevant but salient events: if you are pursuing a valuable goal and you notice something unusual, it may be worth remembering in case it bears on goal-attainment. To test this hypothesis, they showed participants high (\$2.00) and low (\$0.10) value cues before repeatedly presenting color versions of trial-unique “target”



images. The participants' task was to press a button when the target image changed from color to grayscale. On a subset of trials, a novel image was inserted into the series of targets, allowing Murty and Adcock to pose this question: Are participants more likely to remember novel images following presentation of the high versus the low-value cue? The authors predicted that this effect would emerge, based on the hypothesis that the hippocampus would signal expectancy violations (random presentation of a novel image in a repeating sequence) especially vigorously when a desired goal was at stake.

A memory test administered 30 min later confirmed expectations: memory was better for novel images presented after the high-value versus the low-value cue. Furthermore, only one brain region showed a stronger response to novel images following high- versus low-value cues—namely, the left hippocampus. Hippocampal activation was predicted by the VTA response to high-value cues, with subsequent analyses indicating that the VTA-to-hippocampal relationship was not direct, but was instead mediated by several cortical regions, including the visual cortex, medial PFC, ventrolateral PFC, and subgenual cingulate. In summary, this work showed that reward anticipation can serve as a context that facilitates incidental encoding, and it also provided a rationale for the existence of this mechanism: the brain is frugal, allocating memory space to unexpected events only if it seems like they might help the organism reap reward more effectively.

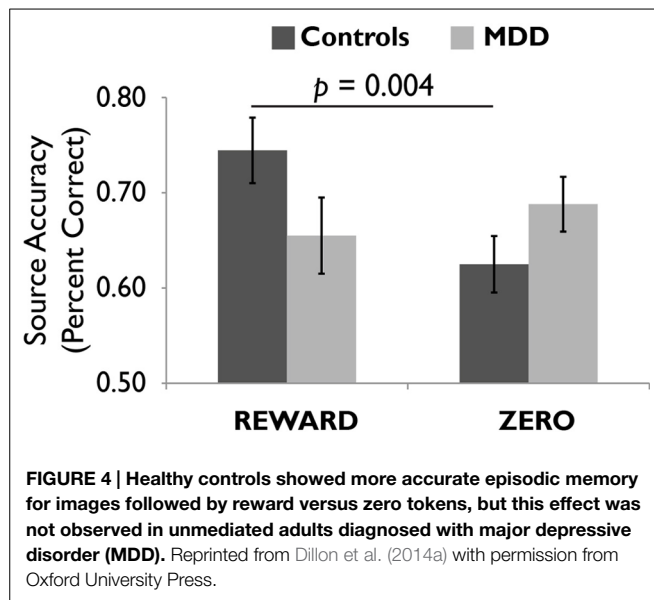
Overall, the human literature is consistent with studies in hippocampal slices and rodents. Novel configurations of familiar stimuli, stimuli presented when novelty is expected, and images shown during reward anticipation are all well-retained after a delay, although dissociating the effect of these manipulations on short- versus long-term has not received as much attention as in the rodent literature. Limited data from pharmacological studies indicate that dopamine transmission may drive these effects, and fMRI data are consistent with this hypothesis, as they show that coactivation of the hippocampus and dopaminergic midbrain supports memory in these tasks. Although more research is needed, the existing data indicate a similar role for dopamine vis-à-vis memory persistence in humans and rodents.

Depression

Based on the evidence reviewed thus far, disruption of the mesolimbic dopamine pathway should have significant, negative consequences for episodic memory in humans, with stronger negative effects on long-term versus short-term memory. Because reward delivery is a potent trigger of dopamine release, memory for positive (i.e., rewarding) events should be preferentially disrupted. The putative relationship between anhedonic depression and dysfunction in mesolimbic dopamine circuitry is now widely-known and has been the subject of numerous reviews (Eshel and Roiser, 2010; Treadway and Zald, 2011; Dillon et al., 2014b; Pizzagalli, 2014), to which the interested reader is directed. The novel question at hand is whether or not this dysfunction has the expected effect on memory for rewarded/positive material. Surprisingly, there is virtually no work on this topic. However, a recent study from our group yielded encouraging results.

We scanned healthy controls and unmedicated, depressed adults as they viewed drawings followed by reward and zero (non-reward) tokens (Dillon et al., 2014a). A source memory test administered directly after encoding revealed better memory for rewarded versus non-rewarded drawings in the controls, but this effect was absent in the depressed group (Figure 4), consistent with the loss of the positive memory advantage in depression (Burt et al., 1995). We also found a stronger response to the reward versus zero tokens in the VTA/SN and right parahippocampus in the controls, but not the depressed group (Figure 5). Finally, the memory advantage for rewarded versus non-rewarded stimuli was strongly correlated with encoding activation in the VTA/SN in the controls, but no correlation was observed in the depressed adults. Thus, we obtained evidence consistent with the hypothesis that disrupted activation of the VTA/SN and MTL memory regions compromised episodic memory for rewarded material in depressed adults relative to controls.

However, there is much work to be done. Our study—like many studies in the human literature on reward and memory—tested memory directly after encoding. Future studies in depression should test memory after a delay of at least 6 h, as this would give dopamine a chance to exert its effects on the transition from early- to late-LTP. Furthermore, an incidental encoding design—where memory testing is unannounced beforehand—would be preferable (participants in our study knew their memory would be tested prior to encoding). This is because intentional encoding



designs like the one we used invite group differences in encoding strategy, which make accurate interpretation challenging. Finally, although it is not possible to directly assess dopamine transmission using fMRI, the reinforcement learning literature has productively used computational modeling to extract a putative dopamine signal from fMRI data (O'Doherty et al., 2007), and there is reason to believe that a computational approach to psychiatric disorders will prove fruitful (Maia and Frank, 2011). In particular, computational modeling may be able to provide increasingly sensitive tests of the proposed role for dopamine abnormalities in human episodic memory failures (for theory on the role of prediction errors in memory, see Henson and Gagnepain, 2010). For example, one could look for a positive relationship between (positive) prediction errors at encoding and accuracy on delayed memory tests in healthy controls, since positive prediction errors are known to elicit burst firing in VTA dopamine neurons (Schultz, 1998). Obtaining such evidence would then allow a test of the hypothesis that this relationship is disrupted in unipolar depression, either because prediction errors are not generated appropriately in depressed adults, or because, once generated, they are not signaled effectively to medial temporal lobe memory regions.

At this point some readers may wonder whether this hypothesis, even if true, is clinically relevant. After all, the cardinal symptoms of major depression are excessive sadness and anhedonia, not memory deficits, and most people probably do not think of depression as a memory disorder. However, research with patients tells a different story. MacQueen et al. (2002) administered a survey to 100 outpatients and found that memory problems were rated as the third most troublesome aspect of depression, behind low libido and weight gain but ahead of sad mood, low energy, poor sleep, and many other symptoms that might be considered more characteristic of depression. The same study found reason to believe patient reports: patients' self-reported assessment of memory problems was correlated with performance on recollection-based memory tests, which

depend heavily on hippocampal function. In other words, the participants thought that their memories were failing, and they were right (but see Mowla et al., 2008, for evidence that depressed adults have limited insight into the extent of their memory deficits).

A new field focused on memory therapeutics for depression is emerging in response to findings like these (for review, see Dalgleish and Werner-Seidler, 2014; for a review of work directed at improving working memory in depressed adults, see (Becker et al., 2015). The field is oriented around the fact that autobiographical memory retrieval has several striking qualities in depressed adults. First, it is oriented toward emotionally negative material. Second, it is frequent overly general: given a cue and the explicit instruction to retrieve a specific, time-limited memory, depressed adults are instead prone to recall categorical memories that span several discrete events and that often have a negative theme. Third, autobiographical memory retrieval is apt to set off a downward spiral of rumination and self-recrimination, as the depressed individual perseverates on past failures and fails to see how events may play out differently in the future. The field of memory therapeutics takes these qualities of autobiographical memory retrieval as targets, training depressed adults to rapidly retrieve positive memories and elaborate upon them, being as specific as possible and avoiding cycles of negative rumination. Furthermore, when unintentional retrieval of negative memories does occur, clients are instructed to regard them with the non-judgmental, accepting perspective taught in mindfulness-based cognitive therapy (Teasdale et al., 2000), as a way to defuse the memories' emotional charge. Although the field of memory therapeutics is very new, there is already some evidence that these methods are clinically effective (e.g., Watkins et al., 2009, 2012; Neshat-Doost et al., 2012).

The proposal developed here is complementary to work on memory therapeutics because its focus is different. Memory therapeutics are primarily aimed at improving the precision and selectivity of retrieval, whereas the hypothesis advanced here proposes that dopamine dysfunction compromises memory formation. If this hypothesis proves true, there is no reason why a clinician could not target both encoding and retrieval for maximum benefit. One can imagine a scenario in which improved dopaminergic tone in mesolimbic circuits could enhance reward responses in depressed adults, boosting late-LTP and thus improving long-term memory for positive events. Simultaneously, a memory therapeutics approach focused on directing retrieval searches toward concrete, specific, positive material from a client's life could enhance positive mood while simultaneously decreasing the propensity to ruminate on overgeneral negative memories. Together, these two strategies could have a powerful effect on mood that would place clients on an upward spiral toward better outcomes.

Qualifications and a Role for Dopamine in Negative Memories

In order to advance the central argument of this article, I have glossed over several important points that deserve mention. First, this proposal should not be read as equating depression

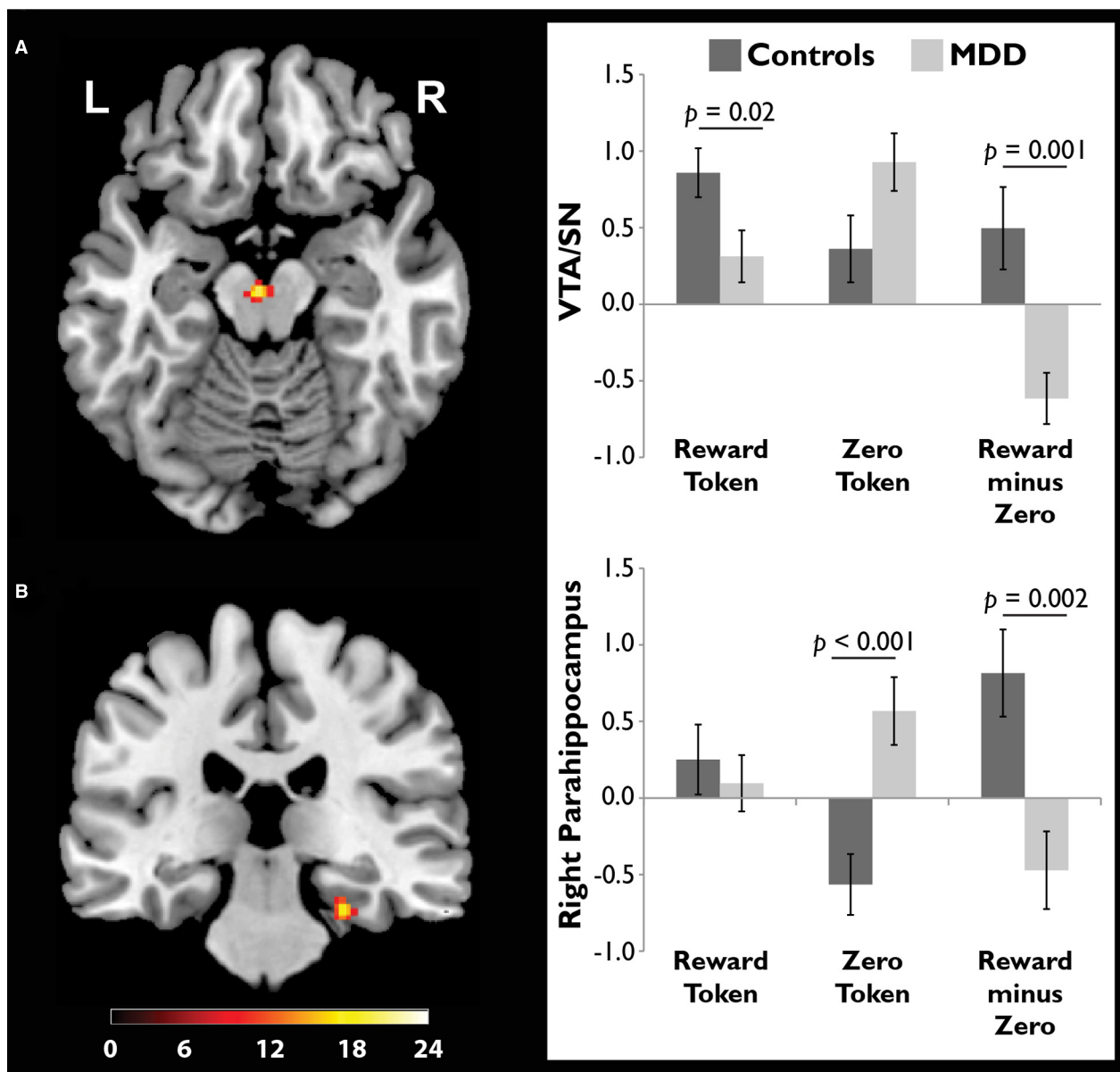


FIGURE 5 | Compared to adults diagnosed with major depressive disorder (MDD), healthy controls showed a stronger response to reward tokens in the ventral tegmental area/substantia nigra (VTA/SN). In both the (A) VTA/SN and (B) right parahippocampus, controls showed stronger activation in response to reward versus zero tokens, while the opposite pattern was observed in the MDD group. In controls but not depressed participants, the VTA/SN “reward minus zero” activation difference score was significantly, positively correlated with the memory advantage for rewarded images (data not shown). Reprinted from Dillon et al. (2014a) with permission from Oxford University Press.

with dopamine dysfunction. Depression is a heterogeneous condition and a diagnosis of MDD can reflect a wide range of symptoms, many of which have little (if anything) to do with dopamine. Thus, this proposal should be narrowly read: it is strictly about how dopamine dysfunction may compromise memory persistence in depression. Second, the mechanism proposed here is complementary to the hippocampal stress hypothesis, because stress is thought to cause dopaminergic abnormalities that are central to depression (Dillon et al., 2014b). In other words, hippocampal volume reductions and reward-

based memory deficits may both be downstream consequences of stress, and it may be useful to determine whether and how they interact with one another (i.e., do changes in D1/D5 receptor distributions figure in hippocampal volume reductions?), particularly because recent rodent studies indicate that antagonizing glucocorticoid receptors can disrupt the acquisition, retrieval, and reconsolidation of conditioned place preferences for rewarding events (Dong et al., 2006; Fan et al., 2013; Achterberg et al., 2014). Third, chronic stress models used to induce anhedonia in rodents result in tonically reduced

dopamine concentrations (Willner, 2005), but memory for individual events is probably influenced by phasic dopamine bursting. How tonic and phasic dopamine levels interact to influence memory retention is unclear and an important topic for future study (Shohamy and Adcock, 2010). Fourth, this proposal is focused on encoding and consolidation, but it will be important to examine memory retrieval in depression as well, as the discussion of memory therapeutics implies. Healthy adults retrieve positive memories to repair negative moods, and in doing so they activate the striatum and the medial PFC (Speer et al., 2014), two regions that figure prominently in mechanistic accounts of depression (Pizzagalli et al., 2009; Lemogne et al., 2012). Thus, poor memory for positive material in depression may frequently reflect problems with retrieval. Indeed, as retrieval depends heavily on PFC function (Dobbins et al., 2002) and depression is characterized by hypofrontality (e.g., Mayberg et al., 1994; Pizzagalli, 2011), problems mounting successful retrieval attempts may be a general issue in depressed adults, extending beyond memory for positive material. Fifth, although there is a wealth of evidence indicating that depression is typically associated with weak responses to rewarding stimuli and dysfunction in brain reward systems (Treadway and Zald, 2011; Pizzagalli, 2014; Whitton et al., 2015), evidence linking these findings to dopamine is usually indirect and often based on inference from studies in non-human animals, thus the link between anhedonic symptoms of depression and dopamine needs strengthening. Other neurochemical systems, including the opioid, endocannabinoid, serotonergic, and glutamatergic systems, are likely relevant, and of course humans experience a diverse range of positive emotions that may be more (e.g., excitement) or less (e.g., tranquility) tightly linked to phasic dopamine bursting. Similarly, the account of LTP offered earlier amounts to a thumbnail sketch, as the process of memory formation is remarkably complex and involves many factors at the molecular level. Nonetheless, dopamine appears to play a central role in anhedonic depression and the transition from early-to-late LTP, and thus it is an excellent starting point for a mechanistic account of positive memory deficits in depression.

Finally, it is very clear in the behavioral neuroscience literature that the positive effects of dopamine on retention are not limited to memory for positive events (e.g., Zweifel et al., 2011), although it may be most easily detected and studied in such cases. As an example, Sariñana et al. (2014) showed that D1 receptors in the dentate gyrus are crucial for contextual fear conditioning in mice. They created D1 and D5 receptor knockouts and administered shock in one box (context A). 24 h later, the mice were exposed to the original box and another, similar box (context B). D5 knockouts and control mice froze readily in context A but not context B, consistent with contextual fear conditioning, but D1 knockouts did not discriminate—they froze to a similar degree in both contexts. Similarly, markers of early gene activity (*c-fos* counts) in the dentate gyrus told a similar tale: D5 knockouts discriminated between their home cage, context A (where they had been shocked), and an entirely novel context, but no such differentiation was seen in the D1 knockouts. Consistent with many findings reviewed earlier, no deficit was seen in D1 knockouts when memory was tested one to 3 h

after acquisition, implying that the effect is specific to memory persistence, and no deficit was seen in amygdala-dependent cue-based fear conditioning, implying that the effect is dependent on D1 receptors in the hippocampus. In short, this study provided evidence that D1 receptors in the dentate gyrus are critical for contextual fear conditioning.

In the current context, I wish to use the technically elegant work of Sariñana et al. (2014) to make a simpler point: despite the strong link between dopamine and reward, this study underscores the fact that dopamine (and D1 receptors in particular) can be important for retaining memory for aversive experiences so long as they depend on hippocampal activation. Presumably these findings depend on the subset of midbrain dopaminergic neurons that respond to salient stimuli whether those stimuli are rewarding or punishing (Matsumoto and Hikosaka, 2009; Bromberg-Martin et al., 2011). Consequently, although studying interactions between dopamine networks and the MTL memory system may prove especially valuable for understanding positive memory deficits in depression, it may help explain poor memory more broadly. Along these lines, an intriguing study in rodents found that if phasic dopamine bursting is blocked, a simple light-shock fear conditioning paradigm can result in behavior consistent with generalized anxiety, presumably because the tight relationship between illumination of the light and shock delivery is not well-encoded, leading to overgeneralization of the fear response (Zweifel et al., 2011).

Conclusion

Depressed adults typically present with episodic memory deficits, and they rate these deficits as a particularly troublesome aspect of the illness (MacQueen et al., 2002). Furthermore, memory for positive material is especially impaired in depression but the neural mechanisms responsible for this deficit are not well characterized. I propose that poor memory for positive material in depression emerges because of anhedonia and its association with dysfunction in mesolimbic dopamine networks widely associated with reward processing.

Although there is little direct work on this topic, there is a compelling body of evidence, across several levels of analysis, implicating dopamine transmission in memory persistence. The STC hypothesis presents dopamine as the instigator of protein synthesis that cements the transition from early- to late-LTP in hippocampal neurons. Experiments probing episodic memory in rodents show that blocking D1/D5 receptors in the hippocampus during encoding has little effect on tests of immediate memory but exerts a powerfully negative effect on delayed tests. By contrast, administration of D1/D5 agonists and exposure to novelty reliably boost memory retention. A growing literature in healthy humans is consistent with findings in rats and hippocampal slices, as reward anticipation and novelty exposure enhance encoding and retention. These mechanisms are space-saving devices: the brain can only store so much material, and it gives privileged access to events that are proximal to dopamine release. In this way, episodes that culminate in reward delivery are well-retained, presumably to allow the organism to behave adaptively should similar circumstances arise in the future.

In sum, there is a widespread consensus that anhedonic depression is associated with dysfunction in brain reward circuitry, with an emphasis on stress-induced disruption of the mesolimbic dopamine system. Initial results suggest that this has consequences for memory, as one would predict based on the molecular, behavioral, and human neuroscience literatures. Given the potential to make a meaningful difference in the way memory problems in depression are understood and treated, and in light of the considerable supporting evidence marshaled here, thoroughly

testing and refining this proposal would constitute time and effort well-spent.

Acknowledgments

This author is indebted to Drs. Diego Pizzagalli, Randy Auerbach, and Michael Treadway for their helpful comments on drafts of this manuscript. Funding: The author is supported by a grant from the National Institute of Mental Health (K99/R00 MH094438).

References

- Abraham, W. C. (2003). How long will long-term potentiation last? Philosophical Transactions of the Royal Society of London. *Ser. B Biol. Sci.* 358, 735–744. doi: 10.1098/rstb.2002.1222
- Achterberg, E. J. M., Trezza, V., and Vanderschuren, L. J. M. J. (2014). Glucocorticoid receptor antagonism disrupts reconsolidation of social reward-related memories in rats. *Behav. Pharmacol.* 25, 216–225. doi: 10.1097/FBP.0000000000000039
- Adcock, R. A., Thangavel, A., Whitfield-Gabrieli, S., Knutson, B., and Gabrieli, J. D. E. (2006). Reward-motivated learning: mesolimbic activation precedes memory formation. *Neuron* 50, 507–517. doi: 10.1016/j.neuron.2006.03.036
- Allen, T. A., and Fortin, N. J. (2013). The evolution of episodic memory. *Proc. Natl. Acad. Sci. U.S.A.* 110(Suppl.) 10379–10386. doi: 10.1073/pnas.1301199110
- American Psychiatric Association. (2013). *Diagnostic and Statistical Manual of Mental Disorders*, 5th Edn. Washington, DC: American Psychiatric Association.
- Ballarín, F., Moncada, D., Martínez, M. C., Alen, N., and Viola, H. (2009). Behavioral tagging is a general mechanism of long-term memory formation. *Proc. Natl. Acad. Sci. U.S.A.* 106, 14599–14604. doi: 10.1073/pnas.0907078106
- Baudry, M., Bi, X., Gall, C., and Lynch, G. (2011). The biochemistry of memory: the 26 year journey of a “new and specific hypothesis” *Neurobiol. Learn. Mem.* 95, 125–133. doi: 10.1016/j.nlm.2010.11.015
- Beck, A. T., Rush, A. J., Shaw, B. F., and Emery, G. (1979). *Cognitive Therapy of Depression. Cognitive Therapy of Depression*. New York: Guilford Press.
- Becker, E. S., Vanderhasselt, M.-A., and Vrijen, J. N. (2015). Memory training in depression. *Curr. Opin. Psychol.* 4, 48–52. doi: 10.1016/j.copsyc.2015.02.004
- Bethus, I., Tse, D., and Morris, R. G. M. (2010). Dopamine and memory: modulation of the persistence of memory for novel hippocampal NMDA receptor-dependent paired associates. *J. Neurosci.* 30, 1610–1618. doi: 10.1523/JNEUROSCI.2721-09.2010
- Bliss, T. V., and Lomo, T. (1973). Long-lasting potentiation of synaptic transmission in the dentate area of the anaesthetized rabbit following stimulation of the perforant path. *J. Physiol.* 232, 331–356. doi: 10.1113/jphysiol.1973.sp010273
- Bower, G. H. (1981). Mood and memory. *Am. Psychol.* 36, 129–148. doi: 10.1037/0003-066X.36.2.129
- Bower, G. H. (1987). Commentary on mood and memory. *Behav. Res. Ther.* 25, 443–455. doi: 10.1016/0005-7967(87)90052-0
- Bromberg-Martin, E. S., Matsumoto, M., and Hikosaka, O. (2011). Dopamine in motivational control: rewarding, aversive, and alerting. *Neuron* 68, 815–834. doi: 10.1016/j.neuron.2010.11.022
- Burt, D. B., Zembar, M. J., and Niederehe, G. (1995). Depression and memory impairment: a meta-analysis of the association, its pattern, and specificity. *Psychol. Bull.* 117, 285–305. doi: 10.1037/0033-2909.117.2.285
- Chowdhury, R., Guitart-Masip, M., Bunzeck, N., Dolan, R. J., and Duzel, E. (2012). Dopamine modulates episodic memory persistence in old age. *J. Neurosci.* 32, 14193–14204. doi: 10.1523/JNEUROSCI.1278-12.2012
- Clayton, N. S., and Dickinson, A. (1998). Episodic-like memory during cache recovery by scrub jays. *Nature* 395, 272–274. doi: 10.1038/26216
- Dalgleish, T., and Werner-Seidler, A. (2014). Disruptions in autobiographical memory processing in depression and the emergence of memory therapeutics. *Trends Cogn. Sci.* 18, 596–604. doi: 10.1016/j.tics.2014.06.010
- Dillon, D. G., Dobbins, I. G., and Pizzagalli, D. A. (2014a). Weak reward source memory in depression reflects blunted activation of VTA/SN and parahippocampus. *Soc. Cogn. Affect. Neurosci.* 9, 1576–1583. doi: 10.1093/scan/nst155
- Dillon, D. G., Rosso, I. M., Pechtel, P., Killgore, W. D. S., Rauch, S. L., and Pizzagalli, D. A. (2014b). Peril and pleasure: an RDOC-inspired examination of threat responses and reward processing in anxiety and depression. *Depress. Anxiety* 31, 233–249. doi: 10.1002/da.22202
- Dobbins, I. G., Foley, H., Schacter, D. L., and Wagner, A. D. (2002). Executive control during episodic retrieval: multiple prefrontal processes subserve source memory. *Neuron* 35, 989–996. doi: 10.1016/S0896-6273(02)00858-9
- Dong, Z., Han, H., Wang, M., Xu, L., Hao, W., and Cao, J. (2006). Morphine conditioned place preference depends on glucocorticoid receptors in both hippocampus and nucleus accumbens. *Hippocampus* 16, 809–813. doi: 10.1002/hipo.20216
- Eshel, N., and Roiser, J. P. (2010). Reward and punishment processing in depression. *Biol. Psychiatry* 68, 118–124. doi: 10.1016/j.biopsych.2010.01.027
- Fan, Y.-D., Niu, H.-C., Huma, T., Li, L., Wang, G.-M., Xu, L.-Q., et al. (2013). Blockage of glucocorticoid receptors during memory acquisition, retrieval and reconsolidation prevents the expression of morphine-induced conditioned place preferences in mice. *Zool. Res.* 34, E26–E34. doi: 10.3724/SPJ.1141.2013.E01E26
- Frey, S., and Frey, J. U. (2008). “Synaptic tagging” and “cross-tagging” and related associative reinforcement processes of functional plasticity as the cellular basis for memory formation,” in *Progress in Brain Research*, Vol. 169, eds W. S. Sossin, J.-C. Lacaille, V. F. Castellucci, and S. Belleville (Elsevier: Amsterdam), 117–143. doi: 10.1016/S0079-6123(07)00007-6
- Frey, U., and Morris, R. G. (1997). Synaptic tagging and long-term potentiation. *Nature* 385, 533–536. doi: 10.1038/385533a0
- Frey, U., and Morris, R. G. (1998a). Synaptic tagging: implications for late maintenance of hippocampal long-term potentiation. *Trends Neurosci.* 21, 181–188. doi: 10.1016/S0166-2236(97)01189-2
- Frey, U., and Morris, R. G. (1998b). Weak before strong: dissociating synaptic tagging and plasticity-factor accounts of late-LTP. *Neuropharmacology* 37, 545–552. doi: 10.1016/S0028-3908(98)00040-9
- Frey, U., Schroeder, H., and Matthies, H. (1990). Dopaminergic antagonists prevent long-term maintenance of posttetanic LTP in the CA1 region of rat hippocampal slices. *Brain Res.* 522, 69–75. doi: 10.1016/0006-8993(90)91578-5
- Gotlib, I. H., Kasch, K. L., Traill, S., Joormann, J., Arnou, B. A., and Johnson, S. L. (2004). Coherence and specificity of information-processing biases in depression and social phobia. *J. Abnorm. Psychol.* 3, 386–398. doi: 10.1037/0021-843X.113.3.386
- Hamilton, J. P., and Gotlib, I. H. (2008). Neural substrates of increased memory sensitivity for negative stimuli in major depression. *Biol. Psychiatry* 63, 1155–1162. doi: 10.1016/j.biopsych.2007.12.015
- Henson, R. N., and Gagnepain, P. (2010). Predictive, interactive multiple memory systems. *Hippocampus* 20, 1315–1326. doi: 10.1002/hipo.20857
- Huang, Y., Coupland, N. J., Lebel, R. M., Carter, R., Seres, P., Wilman, A. H., et al. (2013). Structural changes in hippocampal subfields in major depressive disorder: a high-field magnetic resonance imaging study. *Biol. Psychiatry* 74, 62–68. doi: 10.1016/j.biopsych.2013.01.005
- Huang, Y. Y., and Kandel, E. R. (1995). D1/D5 receptor agonists induce a protein synthesis-dependent late potentiation in the CA1 region of the hippocampus. *Proc. Natl. Acad. Sci. U.S.A.* 92, 2446–2450. doi: 10.1073/pnas.92.7.2446
- Kahn, I., and Shohamy, D. (2013). Intrinsic connectivity between the hippocampus, nucleus accumbens, and ventral tegmental area in humans. *Hippocampus* 23, 187–192. doi: 10.1002/hipo.22077
- Kendler, K. S., Karkowski, L. M., and Prescott, C. A. (1999). Causal relationship between stressful life events and the onset of major depression. *Am. J. Psychiatry* 156, 837–841. doi: 10.1176/ajp.156.6.837

- Kusuki, T., Imahori, Y., Ueda, S., and Inokuchi, K. (1997). Dopaminergic modulation of LTP induction in the dentate gyrus of intact brain. *Neuroreport* 8, 2037–2040. doi: 10.1097/00001756-199705260-00046
- Lemogne, C., Delaveau, P., Freton, M., Guionnet, S., and Fossati, P. (2012). Medial prefrontal cortex and the self in major depression. *J. Affect. Disord.* 136, e1–e11. doi: 10.1016/j.jad.2010.11.034
- Lisman, J. E., and Grace, A. A. (2005). The hippocampal-VTA loop: controlling the entry of information into long-term memory. *Neuron* 46, 703–713. doi: 10.1016/j.neuron.2005.05.002
- Lisman, J., Grace, A. A., and Duzel, E. (2011). A neoHebbian framework for episodic memory; role of dopamine-dependent late LTP. *Trends Neurosci.* 34, 536–547. doi: 10.1016/j.tins.2011.07.006
- Lynch, M. A. (2004). Long-term potentiation and memory. *Physiol. Rev.* 84, 87–136. doi: 10.1152/physrev.00014.2003
- MacDonald, C. J., Carrow, S., Place, R., and Eichenbaum, H. (2013). Distinct hippocampal time cell sequences represent odor memories in immobilized rats. *J. Neurosci.* 33, 14607–14616. doi: 10.1523/JNEUROSCI.1537-13.2013
- MacQueen, G., and Frodl, T. (2011). The hippocampus in major depression: evidence for the convergence of the bench and bedside in psychiatric research? *Mol. Psychiatry* 16, 252–264. doi: 10.1038/mp.2010.80
- MacQueen, G. M., Campbell, S., McEwen, B. S., Macdonald, K., Amano, S., Joffe, R. T., et al. (2003). Course of illness, hippocampal function, and hippocampal volume in major depression. *Proc. Natl. Acad. Sci. U.S.A.* 100, 1387–1392. doi: 10.1073/pnas.0337481100
- MacQueen, G. M., Galway, T. M., Hay, J., Young, L. T., and Joffe, R. T. (2002). Recollection memory deficits in patients with major depressive disorder predicted by past depressions but not current mood state or treatment status. *Psychol. Med.* 32, 251–258. doi: 10.1017/S0033291701004834
- Maia, T. V., and Frank, M. J. (2011). From reinforcement learning models to psychiatric and neurological disorders. *Nat. Neurosci.* 14, 154–162. doi: 10.1038/nn.2723
- Malinow, R. (2003). AMPA receptor trafficking and long-term potentiation. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 358, 707–714. doi: 10.1098/rstb.2002.1233
- Mandeville, J. B., Sander, C. Y. M., Jenkins, B. G., Hooker, J. M., Catana, C., Vanduffel, W., et al. (2013). A receptor-based model for dopamine-induced fMRI signal. *Neuroimage* 75, 46–57. doi: 10.1016/j.neuroimage.2013.02.036
- Matsumoto, M., and Hikosaka, O. (2009). Two types of dopamine neuron distinctly convey positive and negative motivational signals. *Nature* 459, 837–841. doi: 10.1038/nature08028
- Matt, G. E., Vázquez, C., and Campbell, W. K. (1992). Mood-congruent recall of affectively toned stimuli: a meta-analytic review. *Clin. Psychol. Rev.* 12, 227–255. doi: 10.1016/0272-7358(92)90116-P
- Mayberg, H. S., Lewis, P. J., Regenold, W., and Wagner, H. N. (1994). Parahypoperfusion in unipolar depression. *J. Nucl. Med.* 35, 929–934.
- Moncada, D., and Viola, H. (2007). Induction of long-term memory by exposure to novelty requires protein synthesis: evidence for a behavioral tagging. *J. Neurosci.* 27, 7476–7481. doi: 10.1523/JNEUROSCI.1083-07.2007
- Monroe, S. M., and Harkness, K. L. (2005). Life stress, the “kindling” hypothesis, and the recurrence of depression: considerations from a life stress perspective. *Psychol. Rev.* 112, 417–445. doi: 10.1037/0033-295X.112.2.417
- Mowla, A., Ashkani, H., Ghanizadeh, A., Dehbozorgi, G. R., Sabayan, B., and Chohedri, A. H. (2008). Do memory complaints represent impaired memory performance in patients with major depressive disorder? *Depres. Anxiety* 25, E92–E96. doi: 10.1002/da.20343
- Murty, V. P., and Adcock, R. A. (2014). Enriched encoding: reward motivation organizes cortical networks for hippocampal detection of unexpected events. *Cereb. Cortex* 24, 2160–2168. doi: 10.1093/cercor/bht063
- Neshat-Doost, H. T., Dalgleish, T., Yule, W., Kalantari, M., Ahmadi, S. J., Dyregrov, A., et al. (2012). Enhancing autobiographical memory specificity through cognitive training: an intervention for depression translated from basic science. *Clin. Psychol. Sci.* 1, 84–92. doi: 10.1177/2167702612454613
- Nolen-Hoeksema, S. (1991). Responses to depression and their effects on the duration of depressive episodes. *J. Abnorm. Psychol.* 100, 569–582. doi: 10.1037/0021-843X.100.4.569
- O’Doherty, J. P., Hampton, A., and Kim, H. (2007). Model-based fMRI and its application to reward learning and decision making. *Ann. N. Y. Acad. Sci.* 1104, 35–53. doi: 10.1196/annals.1390.022
- Pakkenberg, B., Pelvig, D., Marner, L., Bundgaard, M. J., Gundersen, H. J. G., Nyengaard, J. R., et al. (2003). Aging and the human neocortex. *Exp. Gerontol.* 38, 95–99. doi: 10.1016/S0531-5565(02)00151-1
- Pizzagalli, D. A. (2011). Frontocingulate dysfunction in depression: toward biomarkers of treatment response. *Neuropsychopharmacology* 36, 183–206. doi: 10.1038/npp.2010.166
- Pizzagalli, D. A. (2014). Depression, stress, and anhedonia: toward a synthesis and integrated model. *Annu. Rev. Clin. Psychol.* 10, 393–423. doi: 10.1146/annurev-clinpsy-050212-185606
- Pizzagalli, D. A., Holmes, A. J., Dillon, D. G., Goetz, E. L., Birk, J. L., Bogdan, R., and et al. (2009). Reduced caudate and nucleus accumbens response to rewards in unmedicated subjects with major depressive disorder. *Am. J. Psychiatry* 166, 702–710. doi: 10.1176/appi.ajp.2008.08081201
- Rogerson, T., Cai, D. J., Frank, A., Sano, Y., Shobe, J., Lopez-Aranda, M. F., et al. (2014). Synaptic tagging during memory allocation. *Nat. Rev. Neurosci.* 15, 157–169. doi: 10.1038/nrn3667
- Sapolsky, R. M. (1996). Why stress is bad for your brain. *Science* 273, 749–750. doi: 10.1126/science.273.5276.749
- Sapolsky, R. M. (2000). Glucocorticoids and hippocampal atrophy in neuropsychiatric disorders. *Arch. Gen. Psychiatry* 57, 925–935. doi: 10.1001/archpsyc.57.10.925
- Sariñana, J., Kitamura, T., Künzler, P., Sultzman, L., and Tonegawa, S. (2014). Differential roles of the dopamine 1-class receptors, D1R and D5R, in hippocampal dependent memory. *Proc. Natl. Acad. Sci. U.S.A.* 111, 8245–8250. doi: 10.1073/pnas.1407395111
- Schott, B. H., Sellner, D. B., Lauer, C., Habib, R., Frey, J. U., Guderian, S., et al. (2004). Activation of midbrain structures by associative novelty and the formation of explicit memory in humans. *Learn. Mem.* 11, 383–387. doi: 10.1101/lm.75004
- Schultz, W. (1998). Predictive reward signal of dopamine neurons. *J. Neurophysiol.* 80, 1–27.
- Sheline, Y. I., Sanghavi, M., Mintun, M. A., and Gado, M. H. (1999). Depression duration but not age predicts hippocampal volume loss in medically healthy women with recurrent major depression. *J. Neurosci.* 19, 5034–5043.
- Shohamy, D., and Adcock, R. A. (2010). Dopamine and adaptive memory. *Trends Cogn. Sci.* 14, 464–472. doi: 10.1016/j.tics.2010.08.002
- Smith, W. B., Starck, S. R., Roberts, R. W., and Schuman, E. M. (2005). Dopaminergic stimulation of local protein synthesis enhances surface expression of GluR1 and synaptic transmission in hippocampal neurons. *Neuron* 45, 765–779. doi: 10.1016/j.neuron.2005.01.015
- Speer, M. E., Bhanji, J. P., and Delgado, M. R. (2014). Savoring the past: positive memories evoke value representations in the striatum. *Neuron* 84, 847–856. doi: 10.1016/j.neuron.2014.09.028
- Squire, L. R. (1992). Memory and the hippocampus: a synthesis from findings with rats, monkeys, and humans. *Psychol. Rev.* 99, 195–231. doi: 10.1037/0033-295X.99.2.195
- Stockmeier, C. A., Mahajan, G. J., Konick, L. C., Overholser, J. C., Jurjus, G. J., Meltzer, H. Y., et al. (2004). Cellular changes in the postmortem hippocampus in major depression. *Biol. Psychiatry* 56, 640–650. doi: 10.1016/j.biopsych.2004.08.022
- Teasdale, J. D., Segal, Z. V., Williams, J. M., Ridgeway, V. A., Soulsby, J. M., and Lau, M. A. (2000). Prevention of relapse/recurrence in major depression by mindfulness-based cognitive therapy. *J. Consult. Clin. Psychol.* 68, 615–623. doi: 10.1037/0022-006X.68.4.615
- Travis, S., Coupland, N. J., Silversone, P. H., Huang, Y., Fujiwara, E., Carter, R., et al. (2014). Dentate gyrus volume and memory performance in major depressive disorder. *J. Affect. Disord.* 172C, 159–164. doi: 10.1016/j.jad.2014.09.048
- Treadway, M. T., and Zald, D. H. (2011). Reconsidering anhedonia in depression: lessons from translational neuroscience. *Neurosci. Biobehav. Rev.* 35, 537–555. doi: 10.1016/j.neubiorev.2010.06.006
- Tse, D., Langston, R. F., Kakeyama, M., Bethus, I., Spooner, P. A., Wood, E. R., et al. (2007). Schemas and memory consolidation. *Science* 316, 76–82. doi: 10.1126/science.1135935
- Tulving, E. (1993). What is episodic memory? *Curr. Direct. Psychol. Sci.* 2, 67–70. doi: 10.1111/1467-8721.ep10770899
- Wang, Q., Van Heerikhuizen, J., Aronica, E., Kawata, M., Seress, L., Joels, M., et al. (2013). Glucocorticoid receptor protein expression in human

- hippocampus; stability with age. *Neurobiol. Aging* 34, 1662–1673. doi: 10.1016/j.neurobiolaging.2012.11.019
- Wang, S., Redondo, R. L., and Morris, R. G. (2010). Relevance of synaptic tagging and capture to the persistence of long-term potentiation and everyday spatial memory. *Proc. Natl. Acad. Sci. U.S.A.* 107, 19537–19542. doi: 10.1073/pnas.1008638107
- Watkins, E. R., Baeyens, C. B., and Read, R. (2009). Concreteness training reduces dysphoria: proof-of-principle for repeated cognitive bias modification in depression. *J. Abnorm. Psychol.* 118, 55–64. doi: 10.1037/a0013642
- Watkins, E. R., Taylor, R. S., Byng, R., Baeyens, C., Read, R., Pearson, K., et al. (2012). Guided self-help concreteness training as an intervention for major depression in primary care: a Phase II randomized controlled trial. *Psychol. Med.* 42, 1359–1371. doi: 10.1017/S0033291711002480
- Whitton, A. E., Treadway, M. T., and Pizzagalli, D. A. (2015). Reward processing dysfunction in major depression, bipolar disorder and schizophrenia. *Curr. Opin. Psychiatry* 28, 7–12. doi: 10.1097/YCO.0000000000000122
- Willner, P. (2005). Chronic mild stress (CMS) revisited: consistency and behavioural-neurobiological concordance in the effects of CMS. *Neuropsychobiology* 52, 90–110. doi: 10.1159/000087097
- Wittmann, B. C., Bunzeck, N., Dolan, R. J., and Düzel, E. (2007). Anticipation of novelty recruits reward system and hippocampus while promoting recollection. *Neuroimage* 38, 194–202. doi: 10.1016/j.neuroimage.2007.06.038
- Wittmann, B. C., Schott, B. H., Guderian, S., Frey, J. U., Heinze, H., and Düzel, E. (2005). Reward-related fMRI activation of dopaminergic midbrain is associated with enhanced hippocampus-dependent long-term memory formation. *Neuron* 45, 459–467. doi: 10.1016/j.neuron.2005.01.010
- Wolosin, S. M., Zeithamova, D., and Preston, A. R. (2012). Reward modulation of hippocampal subfield activation during successful associative encoding and retrieval. *J. Cogn. Neurosci.* 24, 1532–1547. doi: 10.1162/jocn_a_00237
- Wolosin, S. M., Zeithamova, D., and Preston, A. R. (2013). Distributed hippocampal patterns that discriminate reward context are associated with enhanced associative binding. *J. Exp. Psychol. Gen.* 142, 1264–1276. doi: 10.1037/a0033609
- Zakzanis, K. K., Leach, L., and Kaplan, E. (1998). On the nature and pattern of neurocognitive function in major depressive disorder. *Neuropsychiatry Neuropsychol. Behav. Neurol.* 11, 111–119.
- Zweifel, L. S., Fadok, J. P., Argilli, E., Garelick, M. G., Jones, G. L., Dickerson, T. M. K., et al. (2011). Activation of dopamine neurons is critical for aversive conditioning and prevention of generalized anxiety. *Nat. Neurosci.* 14, 620–626. doi: 10.1038/nn.2808

Conflict of Interest Statement: The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2015 Dillon. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Reduction in ventral striatal activity when anticipating a reward in depression and schizophrenia: a replicated cross-diagnostic finding

Gonzalo Arrondo¹, Nuria Segarra¹, Antonio Metastasio¹, Hisham Ziauddeen^{1,2,3}, Jennifer Spencer^{1,3}, Niels R. Reinders¹, Robert B. Dudas^{1,3,4}, Trevor W. Robbins^{5,6}, Paul C. Fletcher^{1,2,3,6} and Graham K. Murray^{1,3,6*}

OPEN ACCESS

Edited by:

Eduardo A. Garza-Villarreal,
National Institute of Psychiatry, Mexico

Reviewed by:

Sharna Jamadar,
Monash University, Australia
Kristine Rømer Thomsen,
Aarhus University, Denmark
Tracy Barbour,
Massachusetts General Hospital, USA

*Correspondence:

Graham K. Murray,
Department of Psychiatry, University of
Cambridge, Box 189 Addenbrooke's
Hospital, Cambridge, CB2 0QQ, UK
gm285@cam.ac.uk

Specialty section:

This article was submitted to
Psychology for Clinical Settings,
a section of the journal
Frontiers in Psychology

Received: 04 March 2015

Accepted: 11 August 2015

Published: 26 August 2015

Citation:

Arrondo G, Segarra N, Metastasio A,
Ziauddeen H, Spencer J, Reinders
NR, Dudas RB, Robbins TW, Fletcher
PC and Murray GK (2015) Reduction
in ventral striatal activity when
anticipating a reward in depression
and schizophrenia: a replicated
cross-diagnostic finding.
Front. Psychol. 6:1280.
doi: 10.3389/fpsyg.2015.01280

¹ Department of Psychiatry, University of Cambridge, Cambridge, UK, ² Wellcome Trust-MRC Institute of Metabolic Science, Cambridge, UK, ³ Cambridgeshire and Peterborough NHS Foundation Trust, UK, ⁴ Psychiatric Liaison Service, Ipswich Hospital, Norfolk and Suffolk NHS Foundation Trust, UK, ⁵ Department of Psychology, University of Cambridge, Cambridge, UK, ⁶ Behavioural and Clinical Neuroscience Institute, University of Cambridge, Cambridge, UK

In the research domain framework (RDoC), dysfunctional reward expectation has been proposed to be a cross-diagnostic domain in psychiatry, which may contribute to symptoms common to various neuropsychiatric conditions, such as anhedonia or apathy/avolition. We used a modified version of the Monetary Incentive Delay (MID) paradigm to obtain functional MRI images from 22 patients with schizophrenia, 24 with depression and 21 controls. Anhedonia and other symptoms of depression, and overall positive and negative symptomatology were also measured. We hypothesized that the two clinical groups would have a reduced activity in the ventral striatum when anticipating reward (compared to anticipation of a neutral outcome) and that striatal activation would correlate with clinical measures of motivational problems and anhedonia. Results were consistent with the first hypothesis: two clusters in both the left and right ventral striatum were found to differ between the groups in reward anticipation. *Post-hoc* analysis showed that this was due to higher activation in the controls compared to the schizophrenia and the depression groups in the right ventral striatum, with activation differences between depression and controls also seen in the left ventral striatum. No differences were found between the two patient groups, and there were no areas of abnormal cortical activation in either group that survived correction for multiple comparisons. Reduced ventral striatal activity was related to greater anhedonia and overall depressive symptoms in the schizophrenia group, but not in the participants with depression. Findings are discussed in relation to previous literature but overall are supporting evidence of reward system dysfunction across the neuropsychiatric continuum, even if the specific clinical relevance is still not fully understood. We also discuss how the RDoC approach may help to solve some of the replication problems in psychiatric fMRI research.

Keywords: reward system, ventral striatum, monetary incentive delay, depressive symptoms, research domain framework

Introduction

Current psychiatric diagnostic manuals divide psychopathology into separate diagnostic categories based in the co-occurrence of signs and symptoms rather than on the basis of underlying physiology. However, within each of these categories there is great heterogeneity in the type and severity of symptoms and, similarly, it is common to have overlap in symptoms between the different diagnoses (Lilienfeld, 2014). Recently the Research Domain Criteria (RDoC) framework has been proposed as an alternative framework in order to advance the understanding of mental disorders (Insel et al., 2010). The specific aim of the RDoC project is to increase research that validates new cross-diagnostic dimensions and biological and behavioral measures to carry out better classifications of mental problems. To achieve this improved classification, the RDoC group has proposed a set of five domains or functional systems that are typically affected in psychopathology, and seven units of analysis at which these five constructs can be studied, thus creating a 2-dimensional matrix that can guide research. Additionally, this matrix also includes a further column of “paradigms,” that is, tools that can be used to measure abnormalities in the domains (Cuthbert, 2014a).

Schizophrenia patients, as defined in current diagnostic manuals, typically suffer from symptoms of psychosis (delusions, hallucinations, and disordered behavior, thought, and speech), whereas depression patients are diagnosed on the basis of low mood, anhedonia, and accompanying physical symptoms such as reduced energy. However, psychotic symptoms can appear in the course of a major depressive episode and low mood, blunted affect, alogia, or anhedonia are also characteristic of schizophrenia (Bedwell et al., 2014). Moreover, the boundary between depression and schizophrenia is frequently unclear, and the diagnosis of schizoaffective disorder is often used in cases in which a patient has features of mood disorder and schizophrenia. Additionally, the existence of mild symptoms of psychosis in young people, previously thought to be confer a high risk for schizophrenia, has been shown to be a more general risk factor for different psychiatric disorders including depression (Murray and Jones, 2012; Hui et al., 2013).

At the cognitive-behavioral construct level, disrupted reward processing has been implicated in both schizophrenia and depression. Moreover, it has been suggested that whereas reward receipt may only be subtly affected in both disorders (Cohen and Minor, 2010; Arrondo et al., 2015), the anticipation of reward (Juckel et al., 2006b; Sherdell et al., 2012) and its motivational aspects (e.g., the effort that a subject is willing to make to get it) (Treadway et al., 2012; Barch et al., 2014; Gard et al., 2014) may be markedly dysfunctional in schizophrenia and in depression (for in-depth reviews of the issue see Barch and Dowd, 2010; Kring and Caponigro, 2010; Treadway and Zald, 2011; Der-Avakian and Markou, 2012; Argyropoulos and Nutt, 2013; Kring and Barch, 2014; Whitton et al., 2015). Abnormalities in reward prediction error signaling in the striatum in schizophrenia are also a well-known finding that could be involved in the pathogenesis of psychotic symptoms (Fletcher and Frith, 2009; Ziauddeen and Murray, 2010). Similar changes have also been found in depression (Kumar et al., 2008), and indeed when participants in both patient groups were studied within the

same prediction error learning paradigm researchers found brain activation differences between controls and the two groups of patients (Gradin et al., 2011).

Hence, consistent with the RDoC proposal, there appears to be dimensional continuity between schizophrenia and depression with respect to at least some aspects of reward processing. According to the RDoC matrix some of the stated similarities would fit in the Positive Valence systems domain, which is defined as involving “Systems primarily responsible for responses to positive motivational situations or contexts, such as reward seeking, consummatory behavior, and reward/habit learning,” and specifically the Expectancy/Reward Prediction Error component within the Approach Motivation subsystem. The ventral striatum/nucleus accumbens is considered a key area involved in the processing of reward anticipation, and the main neurotransmitter thought to be involved in predicting rewards and learning from them is dopamine. The Monetary Incentive Delay task (MID) is a paradigm that it is well-known to elicit strong striatal activations related to the expectation and salience of rewards (Knutson et al., 2001). It has been widely used both in the healthy population and in patients with neuropsychiatric symptoms, with several results pointing toward both schizophrenia and depression patients having a decreased activity in the striatum when anticipating rewards (Juckel et al., 2006a,b; Nielsen et al., 2012; Stoy et al., 2012). Moreover, reanalyzing data from a set of studies, Hägele and colleagues showed that depression, schizophrenia and alcohol disorders were all associated with reduced activity in the right ventral striatum with this effect correlated with depressive symptoms as measured by the Beck Depression Inventory (Hägele et al., 2015). However, given that patient groups were not matched to each other or to controls in age and gender in the study by Hägele and colleagues we sought to replicate and extend it using a specifically-designed study with matched groups; we also wished to relate neural responses to additional key symptoms, as previous work has indicated that ventral striatum (de)activation could also relate to other symptoms such as anhedonia (Simon et al., 2010; Stoy et al., 2012) and more generally, to positive (Nielsen et al., 2012), and negative (Juckel et al., 2006a; Waltz et al., 2009) symptoms.

In brief, we compared patients with depression and schizophrenia to healthy controls in a Monetary Incentive Delay task. The aim was to further understand how perturbation of the anticipation of reward relates to anhedonia, depression, and overall positive and negative symptoms. Consistent with the results of Hägele and colleagues, we hypothesized that the two clinical groups would have a reduced activity in the ventral striatum when anticipating reward and that the striatal activation would correlate with clinical measures (in a direction such that patients with the least activation would have the more pronounced psychopathology).

Materials and Methods

Participants

Sixty seven participants were recruited for the study, of whom 22 had a diagnosis of schizophrenia, 24 a diagnosis of major

depressive disorder and 21 were healthy volunteers. DSM-IV criteria were used for group classification. Inclusion criteria were to be aged between 18 and 65 and speak English proficiently. Exclusion criteria were any contraindication for entering a MRI scan and history of neurological disorder, physical illness, and alcohol or drug dependence. Participants from both clinical groups had subjective symptoms of anhedonia. Demographic and clinical details of participants are provided in **Table 1**. Patients with schizophrenia were recruited through psychiatric community services of the Cambridgeshire and Peterborough NHS Foundation Trust. Patients with depression were recruited through psychiatric and psychological community services of the Cambridgeshire and Peterborough NHS Foundation Trust, and through public advertisement. Diagnoses for patients from psychiatric services of the Cambridgeshire and Peterborough NHS Foundation Trust and suitability for the study were confirmed by the review of all available clinical and anamnestic information by each individual's psychiatrist (an experienced psychiatrist with several years of postgraduate experience who had passed the membership examination of the Royal College of Psychiatrists). Diagnoses for patients recruited from psychology services or advertisement were confirmed by a psychiatric interview including PANSS (Kay et al., 1987) and assessment with the Mini-International Psychiatric Inventory (Sheehan et al., 1998) the interview was conducted either by an experienced research psychiatrist with several years of postgraduate experience who had passed the membership examination of the Royal College of Psychiatrists or by a registered clinical psychologist with several years postgraduate experience.

Thirteen of the depressed participants were taking antidepressant medication: citalopram 30–60 mg daily, mirtazapine 30–45 mg, and venlafaxine 75–225 mg. All patients with schizophrenia were taking atypical antipsychotic medication (specifically clozapine, aripiprazole, risperidone, quetiapine, or olanzapine); two patients were taking a combination of typical and atypical medication. The mean chlorpromazine equivalent dose was 401.24 (*sd* 91.43) mg/day (Kroken et al., 2009). Eight patients with schizophrenia were additionally taking antidepressant medication: citalopram 20–40 mg, fluoxetine 20 mg, mirtazapine 45 mg, venlafaxine 150–225 mg.

The study was conducted at University of Cambridge (Wolfson Brain Imaging Centre and Department of Psychiatry). All participants were evaluated using the following clinical scales: Brief Psychiatric Rating Scale (BPRS, Overall and Gorham, 1962); Positive and Negative Syndrome Scale (PANSS, Kay et al., 1987); Scale for the Assessment of Negative Symptoms Beck Depression Inventory, (SANS, BDI, Beck et al., 1996); Snaith–Hamilton Pleasure Scale (SHAPS, Snaith et al., 1995); and the Temporal Experience of Pleasure scale–TEPS, (Gard et al., 2006). Scales were selected to measure constructs with a possible striatal neural substrate and also according to previous findings of significant correlations with ventral striatum activity during the MID. The Cattell Culture Fair Intelligence Test (CFIT) was used to measure IQ (Cattell et al., 1973).

The study was approved by the Cambridgeshire 3 National Health Service research ethics committee. Written informed consent was obtained from all participants prior to participation.

fMRI Paradigm

The fMRI paradigm was a variation of the Monetary Incentive Delay (MID) task (**Figure 1**). It used an event-related design in which stimuli served as cues signaling the subsequent outcome. Overall, there were 60 trials in the experiment, which was conducted in a single scanning session. There were two types of cues (after Kirsch et al., 2003): reward cue (an arrow pointing upwards; 30 events) or neutral (a horizontal bar with arrows in both extremes; 30 events), and the participants were instructed to press a button in a rapid manner when requested, after the cues disappeared but before the outcome was known. After a 1–4 s random interval showing a fixation cross, the image of a coin indicated the amount of reward (£1 in 70% of the events and 1 penny in 30%; 21 and 9 events) in the case of the win cue, whereas a yellow or orange circle (70 and 30% of events, respectively; 21 and 9 events) were shown after the neutral cue. Hence, despite our instruction to the participants (which was designed to help engagement with the task), rewards did not depend on the subject's performance while pressing the button. This alteration from the original MID task was intended to reduce the confounds of motor preparation and task-induced anxiety which have been proposed as possible reasons for the previously inconsistent results in depression using the MID (Treadway and Zald, 2011). The inter-trial interval, in which a black screen was shown, lasted between 2 and 6 s. Reward and neutral cues, as well as ensuing outcomes were pseudo-randomly presented. Thus, the design was optimized to detect differences between the two anticipation conditions. Behavioral information obtained from the task included reaction time and responses. Data on response times was lost for 8 participants due to programming problems that did not affect acquisition of other data.

MRI Acquisition and Preprocessing

A Siemens Trio Tim 3 T scanner with a 12 channel head coil was used for image acquisition. Functional images were obtained using a Gradient-echo T2*-weighted echo planar sequence and consisted in 32 non-contiguous oblique axial planes (in order to minimize signal drop-out in ventral regions, which was especially important according to our hypothesis). Other parameters included relaxation time = 2000 ms; echo time = 30 ms; flip angle = 78; voxel size = $3.14 \times 3.14 \times 3.75$ mm³, matrix size 64 × 64; bandwidth 2232 HZ/Px. The structural image was obtained using a high-resolution T1-weighted three-dimensional MP-RAGE sequence.

Imaging preprocessing and analysis was carried out using the FEAT v5.98 (fMRI Expert Analysis Tool) routine within the FSL program (FMRIB's Software Library, www.fmrib.ox.ac.uk/fsl). Functional time-series were sequentially realigned, coregistered to a whole brain echo-planar image and finally to the structural high resolution T1 image, and non-brain components were removed. Functional images were also spatially smoothed using a 6 mm at full width half-maximum (FWHM) Gaussian kernel and frequency filtered (130 s cut off). Images were normalized to the

TABLE 1 | Group demographics and clinical characteristics.

PARAMETRIC ONE-WAY ANOVA#									
	C Mean ± sd (n)	D Mean ± sd (n)	S Mean ± sd (n)	Test statistic (F)	p Statistic	p Levene	Post-hoc pair-wise comparisons		
							C vs. D	C vs. S	D vs. S
Age#	34.33 ± 10.11 (21)	33.08 ± 9.15 (24)	32.73 ± 7.62 (22)	0.19	0.828	0.294			
Culture Fair (IQ)#	114.24 ± 19.97 (21)	107.08 ± 16.6 (24)	99.38 ± 19.16 (21)	3.38	0.040	0.681	0.486	0.034	0.422
Education (years)#	14.85 ± 1.93 (20)	13.43 ± 2.21 (23)	13.62 ± 2.11 (21)	2.81	0.068	0.697			C>S
CHI SQUARE									
	C ratio	D ratio	S ratio	Test statistic (X ²)	p Statistic	p Levene	Post-hoc pair-wise comparisons		
							C vs. D	C vs. S	D vs. S
Gender (male/female)	17/4	17/7	19/3	1.71	0.425				
Handedness (right/left)	18/3	22/2	17/5	2.00	0.367				
White-British/other	17/4	20/2	17/5	0.04	0.978				
KRUSKAL-WALLIS ANOVA									
	C Median, IQR (n)	D Median, IQR (n)	S Median, IQR (n)	Test statistic (H)	p Statistic	p Levene	Post-hoc pair-wise comparisons		
							C vs. D	C vs. S	D vs. S
BPRS	24, 1 (21)	44, 11,25 (24)	40, 14,5 (22)	42.47	< 0.001		< 0.001	< 0.001	1
BDI	3, 8 (21)	32, 8 (24)	20, 11 (22)	47.37	< 0.001		< 0.001	0.001	0.005
SHAPS	24, 6 (21)	36, 10 (24)	29.5, 8 (22)	25.82	< 0.001		< 0.001	0.012	0.091
TEPS ant	43.5, 8.8 (20)	26.5, 11 (24)	35, 11.5 (22)	22.04	< 0.001		< 0.001	0.048	0.066
TEPS con	37, 8.75 (20)	44, 11,25 (24)	40, 14.5 (22)	16.00	< 0.001		< 0.001	0.012	1
TEPS total	80, 17 (20)	53.5, 21 (24)	62, 18,25 (22)	22.22	< 0.001		< 0.001	0.017	0.159
PANSS +	7, 0 (21)	7, 1	13, 11.5	26.15	< 0.001		0.372	< 0.001	0.001
PANSS -	7, 0 (21)	12.5, 9	14, 6	38.79	< 0.001		< 0.001	< 0.001	0.273
SANS	0, 0 (20)	2, 1	2, 1	42.49	< 0.001		< 0.001	< 0.001	0.495

Comparisons between groups consisted in parametric One-Way ANOVAs, chi square tests or non-parametric ANOVAs. Post-hoc comparisons were corrected for multiple comparisons. Any significant post-hoc test results for two-group (pair-wise) comparisons ($p < 0.05$) are also indicated by the use of "greater than" symbols (eg., $S\>D>C$ indicates that results in pair-wise comparisons S vs. C and D vs. C were significant). C is the control group, D depression, and S schizophrenia. "p statistic" is the p-value of the omnibus test; "p Levene" is the p-value of Levene's test for the inequality of variances. Standard deviations are denoted as \pm sd. IQR stands for interquartile range and n refers to the group sample size for each comparison. BPRS is Brief Psychiatric Rating Scale, BDI is Beck Depression Inventory, SHAPS is Snail-Hamilton Pleasure Scale, TEPS is the Temporal Experience of Pleasure Scale (ant, anticipatory subscale; con, consummatory subscale), PANSS is Positive and Negative Syndrome Scale (+, positive symptoms; -, negative symptoms), and SANS is Scale for the Assessment of Negative Symptoms. #Results of Kruskal-Wallis ANOVA for these variables can be found in Supplementary Material.

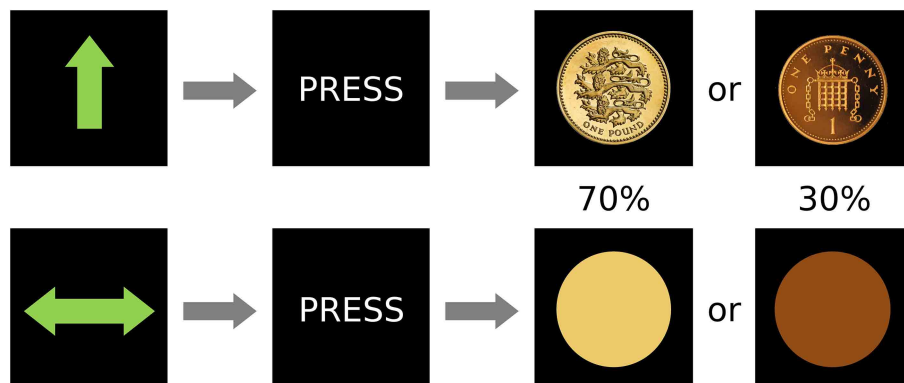


FIGURE 1 | Design of the paradigm. There were two conditions, reward (top) and neutral (bottom) that were signaled by an unambiguous cue and were followed by the subject pressing a button and then by two possible outcomes with different frequencies (0.7 or 0.3 probability of occurrence).

Montreal Neurological Institute (MNI) standard template and the first six volumes were discarded to allow for T1 equilibration effects.

Statistical Analysis

Non-imaging Data

Non-imaging comparisons were carried out in SPSS 21 (IBM, Armonk, NY, US). Results were considered significant if $p < 0.05$. Normality of all variables was initially evaluated through visual inspection of histograms, whisker plots, and Q-Q plots in the 3 groups.

Results from clinical scales did not follow a normal distribution in at least one of the groups. Hence, differences between groups were tested using Kruskal-Wallis ANOVA. To investigate differences in proportions (gender, handedness, and ethnicity) Chi-square comparisons were carried out. Finally, variables which a-priori were considered to be more likely to meet ANOVA assumptions, and for which the results of the initial inspection was less clear (age, intelligence, and education), were taken to a One-Way ANOVA, residuals saved, and explored through inspection of histograms, whisker plots, and Q-Q plots. In the case of these three variables there was some evidence for a non-fully normal distribution of the residuals. However, since ANOVA is considered to be robust to deviations from normality, parametric results are reported in the main article. Additionally, results from Kruskal-Wallis ANOVAs for these variables can be found in Supplementary Material.

Whenever an ANOVA was significant, we conducted *post-hoc* tests consisting of pair-wise comparisons corrected for multiple testing. In the case of Kruskal-Wallis ANOVA, adjusted significance levels were calculated by multiplying the unadjusted significance values by the number of comparisons with a maximum p -value of 1 (SPSS standard method). Gabriel (due to the slightly unequal sample sizes) or Games-Howell procedures were used for One-Way ANOVAs depending on the result of Levene's test on the inequality of variances; the latter was used if the test was significant.

In the case of response times, whisker plots representing data in the reward and neutral condition showed that a participant

with depression had a much higher response time in both conditions. However, the difference between the RT in both conditions was within normal parameters (as confirmed by the evaluation of the residuals of the RT differences between conditions), and therefore not likely to influence the results. Nevertheless, we carried out a repeated-measures ANOVA (within subjects effect: Two levels of condition, between subjects effect: group) with and without this participant.

Imaging Data

We used a single statistical linear regression model with 6 explanatory variables (reward cue, neutral cue; high reward outcome, low reward outcome, and the two neutral outcomes) and their temporal derivatives. Movement parameters from the realignment step were also included in the first-level model.

The a-priori contrast of interest was the anticipation of reward and consisted of the comparison of the BOLD levels during the reward cue and the neutral cue. The reward anticipation contrast uses all cue events in the experiment, with 30 neutral cues and 30 reward cues. Other possible contrasts included the comparison between outcomes, but it was decided not to investigate them at the group level due the reduced number of events that they involved, and because well-predicted rewards often evoke limited brain activity at the time of reward delivery (Berns et al., 2001).

The reward anticipation contrast (reward cue vs. neutral cue) from the first level was taken to the group-level analysis, where it was included in a one sample analysis (control group only, to illustrate this contrast in the healthy population), and One-Way between groups ANOVA (to investigate group differences). Differences were evaluated at the whole brain level and within an a-priori volume of interest mask of the ventral striatum previously used by our group (Bernacer et al., 2013). This region of interest (ROI) included the nucleus accumbens and ventral aspects of the caudate nucleus and putamen (blue regions in **Figure 4**). Comparisons at the whole brain level and within the ROI were cluster-thresholded using a family-wise error (FWE) correction of $p < 0.05$ after a strict initial cluster threshold of $Z > 3$ (Woo et al., 2014). Uncorrected results are also displayed in Supplementary Material as part of exploratory analyses that

may be of use in future hypothesis generation and meta-analyses, using the same $Z > 3$ threshold and a minimum cluster size of 10 across the whole brain.

We extracted the mean parameter estimates for all clusters of differential activation between groups in the imaging ANOVA analysis (using the FSL tool Featquery in the normalized individual images). Then, *post-hoc* pairwise comparisons (two-tailed *t*-tests, equality of variances was not assumed if Levene's test was statistically significant), aimed at exploring the group differences that were driving the significant ANOVA results, were carried out in SPSS. The same *post-hoc* analysis procedure was used for the significant clusters within the ROI comparison of the ventral striatum region.

Activation tables were created using Autoaq (Automatic atlas queries for fsl: <http://brainder.org/2012/07/30/automatic-atlas-queries-in-fsl>). Number of voxels, maximum voxel *Z*-value (*Z* max), MNI coordinates of the maximum peak (MAX X,Y,Z), anatomical label of the max peak and significant pairwise *post-hoc t*-test comparisons are reported within tables. Anatomical labels were the most probable location of the highest peak according to the Harvard-Oxford cortical and subcortical structural atlases included in FSL (Desikan et al., 2006). Images were created using MRIcron (C. Rorden; <http://www.mccauslandcenter.sc.edu/micron/micron/index.html>) and presented in neurological form (right in the image corresponds to the right hemisphere).

A secondary analysis consisted of carrying out Spearman correlations between the mean parameter estimates in the right ventral striatum and clinical symptoms in each of the groups separately. The right ventral striatum was selected as it was the region with reduced activation in schizophrenia and in depression.

Results

There were no group differences in age, gender, handedness, years of education, or ethnicity. As expected, participants with schizophrenia had a lower IQ and patients had greater psychiatric symptomatology than healthy controls (Table 1, Table S1).

Response times were shorter in the reward condition ($F = 36.71$, $p \leq 0.001$) but did not differ between groups ($F = 0.384$, $p = 0.683$). Figure 2 shows whisker plots for the three groups and the two conditions and Table 2 summarizes means and standard deviations. This difference in response time between conditions is characteristic of the paradigm hence indicating that the experimental manipulation was effective in the whole group of participants (Hägele et al., 2015). Results did not change when a subject with slow RTs was taken out (within subject factor $F = 40.348$, $p \leq 0.001$; between subject factor $F = 0.624$, $p = 0.540$).

Reward anticipation (contrast of reward cue vs. neutral cue) in the control group activated areas in the frontal lobe (medial frontal cortex, anterior cingulate cortex), striatum, and thalamus, and cerebellum (FWE cluster corrected across the whole brain $p < 0.05$; see Figure 3 and Table S2). When the three groups were compared using ANOVA, no clusters survived the multiple comparison correction at the whole brain level. However, the ROI analysis in the ventral striatum led to the appearance of two clusters (significant when corrected for multiple comparison)

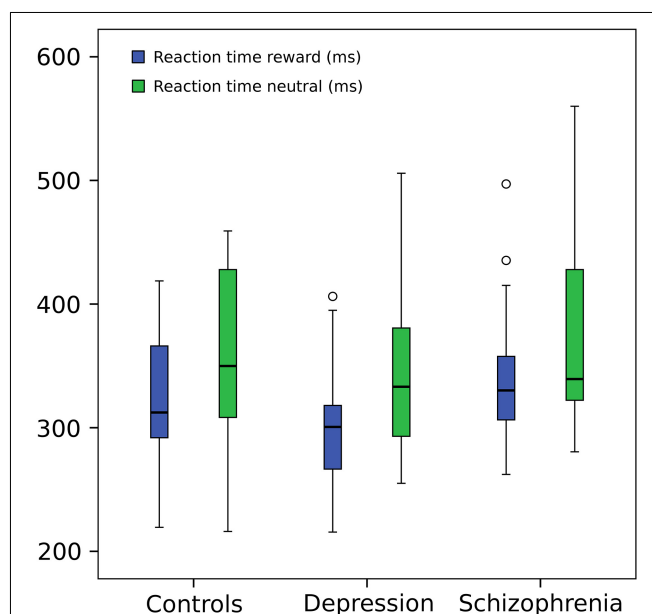


FIGURE 2 | Reaction time. Repeated measures ANOVA (within subject factor: condition, between subject factor: group): reaction times in reward trials were shorter ($F = 36.71$, $p = 0.001$) but did not differ between groups ($F = 0.384$, $p = 0.683$). For graphical purposes a subject from the depression group was eliminated from the image due to a much greater mean RT (around 800 ms) in both conditions, but included in the statistical analysis (However, its elimination from the ANOVA did not change significant results). Central line represents the median value (second quartile, Q2) and the box borders indicate the 1st (Q1) and 3rd quartile (Q3). Hence, the total length of the box is the interquartile range (IQR). Whiskers mark last value in the sample located between the $1.5 \times$ IQR below Q1 and $1.5 \times$ IQR above Q3. Circles represent data between the 1.5 and $3 \times$ IQR below Q1 or above Q3.

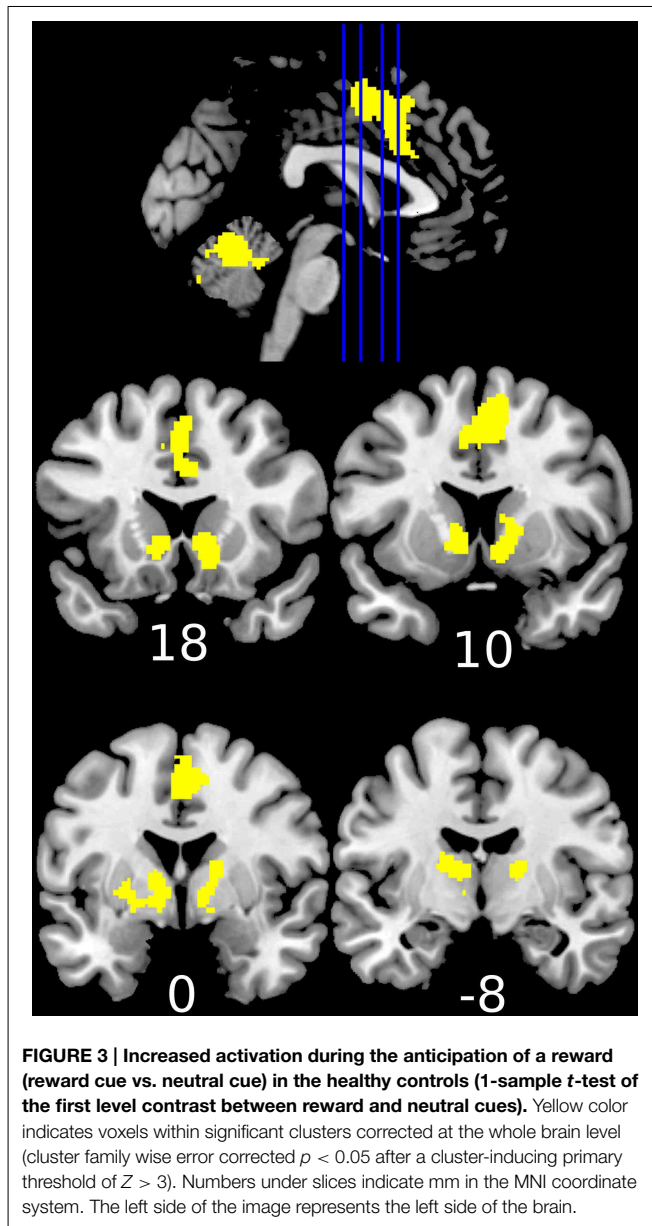
TABLE 2 | Response times by group and condition.

	C (n = 17) Mean \pm sd	D (n = 20) Mean \pm sd	S (n = 22) Mean \pm sd	p Levene
RT reward	0.321 \pm 0.057	0.331 \pm 0.135	0.342 \pm 0.058	0.322
RT neutral	0.356 \pm 0.066	0.367 \pm 0.126	0.368 \pm 0.068	0.427

C is the control group, D depression and S schizophrenia. p Levene is the *p*-value of Levene's test for the inequality of variances. Standard deviations are denoted as \pm sd.

bilaterally in the accumbens nuclei (Figure 4, Table 3). As ANOVA does not indicate the direction of group difference effects, we employed *post-hoc* tests. *Post-hoc* analysis showed activation was significantly greater in the controls than the depressed patients in the right and left accumbens; controls' activation was significantly higher in the right, but not left, accumbens when compared to schizophrenia patients; the two patient groups did not differ from each other (Figure 5 and Table 3).

There were two clusters with *Z*-values above 3 and a cluster size greater than 10 in the uncorrected whole-brain analysis (Table S3). One was located in the right accumbens, the other in the frontal pole. A cluster of eight voxels in the left accumbens was the next biggest cluster. The frontal pole result was derived



from a reduced activation in the depression group compared to the other two groups.

An analysis of correlations between anhedonia (SHAPS and TEPS scales), depression (BDI), psychiatric symptoms (BPRS), and positive and negative symptoms (PANSS and SANS scales) and the parameter estimates in the right ventral striatal cluster of significant differences in the ANOVA was carried out (Table 4 and Figures S1–S9 in Supplementary Material). A negative correlation between severity of depression and anhedonia symptoms and ventral striatum activity was found in the schizophrenia group. Regarding anhedonia, SHAPS and the total TEPS score were statistically significant, whereas TEPS' subscales showed a trend toward significance (marginal significance). The correlation with SANS was also close to significance. Significant results found in schizophrenia patients did not hold in the other

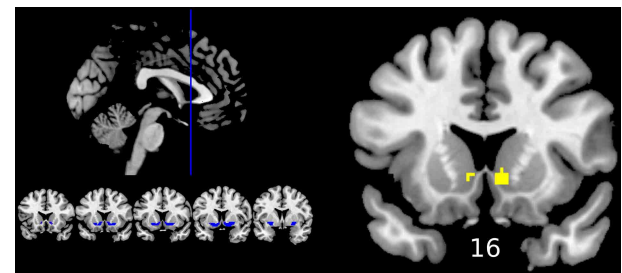


FIGURE 4 | Differences between groups in reward anticipation (One-Way ANOVA of the first level contrast between reward and neutral cues). Yellow color indicates voxels within significant clusters: cluster family wise error corrected $p < 0.05$ within the ventral striatum after a cluster-inducing primary threshold of $Z > 3$. The region of interest is shown in blue on the left of the image. For improved visualization cluster limits are not circumscribed to our ROI. Differences were driven by greater activations in the healthy controls group compared to both groups of patients (right ventral striatum) or only the depression group (left ventral striatum). Numbers under slices indicate mm in the MNI coordinate system. The left side of the image represents the left side of the brain.

groups and the direction of other correlations, such as the BPRS and PANSS negative symptoms within the depression group, was opposed to that expected.

Discussion

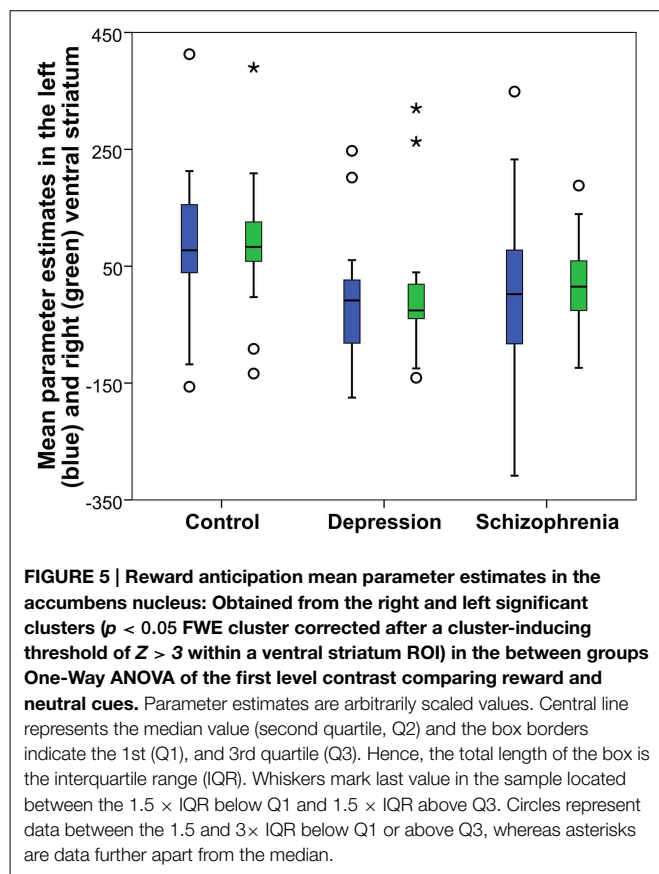
We used a Monetary Incentive Delay (MID) paradigm in a single experiment designed to examine activation during reward anticipation in controls, schizophrenia, and depression. Both clinical groups had a reduction in right ventral striatal activity when anticipating rewards as predicted; results in the left ventral striatum were reduced in depression but not definitively reduced in schizophrenia. We were not able to find a clear correlation between striatal activation and clinical symptoms of depression or anhedonia in the depression group, but such a correlation was present in the schizophrenia group. The design of the study is consistent with the Research Domain Criteria (RDoC) project: one of the designs that the RDoC framework has proposed is to include participants with disorders from different sections of the DSM/ICD diagnostic manuals with the aim of exploring an abnormal neurobehavioral construct to further understand its pathological mechanisms (Cuthbert, 2014a). The rationale underlying the study was the fact that neurocognitive domains of motivation and reward processing have been proposed to be abnormal in both disorders, and such abnormalities may be at the root of some of the common features between schizophrenia and depression.

Our results of reduced right ventral striatal activity during reward anticipation in both depression and schizophrenia can be considered a replication of the work of Hägele et al. which was recently published (2015). The combined evidence of both studies suggests that reduced BOLD signaling in the right nucleus accumbens is indeed a hallmark of pathological reward anticipation. The results in the left accumbens are more equivocal; unlike Hägele and colleagues, we demonstrated

TABLE 3 | Differences between groups in reward anticipation compared to the anticipation of a neutral outcome (ANOVA F test).

Voxels	Z MAX	MAX X (mm)	MAX Y (mm)	MAX Z (mm)	Label	Post-hoc pair-wise comparisons			
						C vs. D	C vs. S	D vs. S	Summary
17	3.92	8	16	-4	Right Accumbens	$t = 2.51, p = 0.016$	$t = 2.06, p = 0.046$	$t = -0.79, p = 0.431$	C>D&S
1	3.15	-8	16	-4	Left Accumbens	$t = 2.96, p = 0.005$	$t = 1.39, p = 0.171$	$t = -0.91, p = 0.371$	C>D

Corrected clusters ($p < 0.05$ FWE after a primary cluster inducing threshold of $Z > 3$ within the ventral striatum) are displayed. Number of voxels, maximum voxel Z-value (Z max), MNI coordinates of the maximum peak (MAX X,Y,Z), anatomical label of the max peak and post-hoc pair-wise t-tests are reported. C is the control group, D depression, and S schizophrenia. Any significant post-hoc test results for two-group (pair-wise) comparisons ($p < 0.05$) are also indicated by the use of "greater than" symbols (eg., S&D>C indicates that results in pair-wise comparisons S vs. C and D vs. C were significant).



reduced left accumbens activation in depression, but similar to Hägele and colleagues we failed to demonstrate a conclusive abnormality here in schizophrenia. Further study will be required to investigate whether there is any fundamental pathological importance in this small laterality effect or whether it relates to more trivial issues such as the precise sensitivity of the paradigm.

In the case of schizophrenia, our findings are not new, as previous studies had also found decreased activity in the basal ganglia and specifically in the ventral striatum (Nielsen et al., 2012), although other smaller studies did not find such differences (Walter et al., 2009; Waltz et al., 2010). On the other hand, results in the MID paradigm with depression patients have been surprisingly inconclusive. For example, the first study using the MID in depression ($n = 14$ patients and 12 controls)

showed no evidence of abnormality in the basal ganglia (Knutson et al., 2008). Our study and the work by Hägele et al. are among the ones with the biggest sample sizes and when taken into account in a combined way strongly suggest that abnormalities in the reward circuit are not limited to patients with a diagnosis of schizophrenia, but also relate to patients with depression symptoms. Other data such as a meta-analysis on reward and depression (Zhang et al., 2013) and a study by Pizzagalli et al. (2009), indicate that reduced basal ganglia activations may not only (or even primarily) be located in the ventral striatum, but could also involve other subcortical regions such as the caudate head or the posterior putamen. We were not able to confirm the results of the meta-analysis by Zhang and colleagues suggesting frontal and anterior cingulate over-activation in depression during reward anticipation.

One reason that has been put forward to account for the discrepancies between MID depression studies has been that reward is usually contingent on a speeded performance on the MID, which might be influencing some patients through a stress related, possibly dopaminergic, response (Treadway and Zald, 2011); these authors argue that the necessity to respond rapidly may enhance activation particularly in anxious patients, whose motivation may relate to hypersensitivity to perceived failure. To deal with this potential confound we modified the original task so reward did not depend on the speed of the response. This change may be related to the finding of reduced accumbens activation, which has also been shown in other experiments in which a speeded motor response was not involved (Smoski et al., 2009). It must be noted however that the paradigm used in the studies reported by Hägele et al. (2015) was closer to the original design of the MID and included the necessity of a fast response to obtain the reward. Similarly, we did not include a loss condition in our design, as it has been proposed that the strongest activations in healthy controls and when comparing them to patients come from contrasting the gain and neutral cues (Hägele et al., 2015). Hence, our design may be suitable and sensitive for psychiatric research when the main objective is to study reward processing. Although reward was delivered irrespective of the speed of button pressing, response time in the reward condition was shorter. Valence effects on response time can occur irrespective of a direct consequence of speeded responses (O'doherty et al., 2004; Pessiglione et al., 2006; Murray et al., 2008b). The phenomenon has been termed "reinforcement related speeding" (Cools et al., 2005; Murray et al., 2008a), and it

TABLE 4 | Correlation between symptoms and parameter estimates in the right nucleus accumbens.

	C		D		S	
	R	p	R	p	R	p
BPRS	0.110	0.664	0.584	0.003	−0.045	0.842
BDI	−0.030	0.898	0.285	0.176	−0.457	0.033
SHAPS	−0.405	0.069	0.163	0.447	−0.483	0.023
TEPS ant	0.237	0.314	−0.332	0.113	0.389	0.073
TEPS con	−0.074	0.757	−0.220	0.302	0.396	0.068
TEPS total	0.103	0.665	−0.367	0.078	0.432	0.045
PANSS +	0.111	0.633	0.304	0.149	−0.046	0.840
PANSS −	0.000	1.000	0.471	0.020	−0.300	0.174
SANS	−	−	0.227	0.286	−0.401	0.065

C is the control group, D depression, and S schizophrenia. Parameter estimates correspond to the mean of the parameter estimates in the clusters of differential activation between groups in the accumbens (obtained from the right significant cluster in the ANOVA comparing reward anticipation activity between groups). R is the Spearman correlation between the measures. BPRS is Brief Psychiatric Rating Scale, BDI is Beck Depression Inventory, TEPS is the Temporal Experience of Pleasure Scale (ant, anticipatory subscale; con, consummatory subscale), SHAPS is Snaith–Hamilton Pleasure Scale, PANSS is Positive and Negative Syndrome Scale (+, positive symptoms; −, negative symptoms), and SANS is Scale for the Assessment of Negative Symptoms.

is thought that a potential reward leads to enhanced motivation and hence faster responding (Crespi, 1942).

Whilst the right accumbens was a site of common underactivation in both patient groups compared to controls, no differences between the two patient groups were found. This result indicates that the two groups of psychiatric participants may be more similar to each other than when compared to healthy controls, which would be in accordance with the RDoC perspective of a common abnormal domain. Comparisons between patient groups were not reported in Hägele et al., but a qualitative analysis of their plotted results show them to be in line with ours; the two patient groups had a similar activity reduction in the striatum. Regarding the effects of the medications for psychosis, it is a limitation that all of the schizophrenia patients were taking antipsychotic medication. Thus, we cannot exclude the possibility that the results may in part be secondary to medication effects. However, we note that as a previous study from the Berlin group found that the striatal deactivation normalized when changing from typical to atypical antipsychotics (Schlagenhauf et al., 2008), and as all of our schizophrenia patients were taking atypical antipsychotic medication, it is unlikely that the results are solely due to medication. Nevertheless, it will be important, albeit challenging, to study medication free samples in future research.

The second hypothesis was that the BOLD signal in the striatum would negatively correlate with clinical symptoms of depression and anhedonia. Previous studies have only reported one or two clinical measures per study, but clinical constructs that correlated with the activity of the ventral striatum have included depression (Hägele et al., 2015) anhedonia and apathy (Simon et al., 2010; Stoy et al., 2012), positive symptoms (Nielsen et al., 2012) or negative (Juckel et al., 2006a; Waltz et al., 2009) symptoms, and severity of overall psychiatric symptoms measured by the BPRS (Waltz et al., 2009). In accordance with our cross-diagnostic approach and the aim of further

investigating the mechanisms of abnormal reward processing, we decided to include a broad range of clinical measures that encompassed most of the constructs previously reported to correlate with striatal activity in the MID. Results in this regard were mixed. On the one hand we were able to replicate Hägele's et al. results of reduced activity in the right accumbens nucleus of those schizophrenia participants with more depressive symptoms. We also extend the results of Hägele and colleagues to relationships between less activity and more anhedonia in schizophrenia, as measured by the TEPS and SHAPS scales. A limitation of our work is that we did not correct our correlation analyses for multiple comparisons, and considering the modest effect sizes observed, those significant correlations we do find may be vulnerable to Type I error. In contrast to results from Hägele et al., neither the BDI nor any of the other measures were associated with ventral striatal activation in both patient groups. However, the lack of brain-symptom associations demonstrated in the depression group, and the control group (as some of the scales were designed to assess patients only), cannot be considered as evidence of absence of brain-symptom relations. Some correlations such as the significant positive relationship between activation and the BPRS and PANSS (negative symptoms subscale) scales in the depression group were counterintuitive (and not maintained across groups). This may reflect a chance finding, or, as Treadway and Zald (2011) have speculated, the activation elicited by the MID task may be a composite of activation associated with reward anticipation and anticipatory anxiety about potential failure. We attempted to address this possibility by our modification of the task to dissociate reinforcement from performance but it remains possible that, especially in depression, some symptoms may relate to striatal overactivity (leading to greater activation when more psychopathology) and other symptoms may relate to underactivity. The relationship between psychopathology and striatal reward processing activation may be complex, and it is possible that striatal (dys)function may contribute

to symptom expression only through interactions with other regional dysfunction and other psychological processes. The concept of dysfunction of one psychological process in one brain area leading to expression of one symptom has the attraction of being a testable hypothesis but is necessarily an oversimplification.

An important challenge in assessing brain-symptom relationships is accurate symptom measurement. This is especially challenging when experts disagree about what constitutes a particular symptom. Anhedonia is defined in the DSM-IV-TR as a loss of interest or pleasure (American Psychiatric Association, 2000), which arguably reflects the consensus use of the term over the past 100 or so years (e.g., Myerson, 1922; see also Berrios, 1996). However, the DSM-5 contains a new definition within the schizophrenia (not depression) chapter, “the decreased ability to experience pleasure from positive stimuli or a degradation in the recollection of pleasure previously experienced,” and arguments continue as to whether a broad or narrow use of the term is more helpful (recently discussed by Treadway and Zald, 2011; Der-Avakian and Markou, 2012; Romer Thomsen et al., 2015). Consensus appears to be building that in depression and schizophrenia, anticipatory, and motivational aspects of reward are more compromised than consummatory (reward receipt) aspects, possibly related to dopaminergic abnormalities in both conditions (Argyropoulos and Nutt, 2013; Kring and Barch, 2014; Whitton et al., 2015); though see a recent study (Gard et al., 2014) documenting enhanced anticipation of pleasure in schizophrenia). However, it remains possible that this relative consensus may in part reflect the methods that have been recently used to investigate the issue. Variants of the monetary incentive delay task as we used here may be more sensitive to reward anticipation effects than reward delivery effects, as well-predicted rewards tend to evoke less strong brain responses than surprising rewards (e.g., Berns et al., 2001). In addition, given the limited temporal resolution of fMRI it can be hard to dissociate anticipatory and consummatory aspects of reward. Furthermore, most fMRI patient studies have, as we did, used monetary rewards, but processing of primary and secondary rewards may differ in important respects (Sescousse et al., 2013).

Strengths and Limitations

As noted, our work is similar to existing research, such as that of Hägele and colleagues. However, there are differences between the two studies. Our study was at higher field strength (we used a 3 Tesla magnet vs. a 1.5 Tesla of Hägele and all). Our study uses a matched control group whereas Hägele and colleagues, because theirs is a retrospective synthesis and re-analysis of previously published separate works, use a control group that is not matched to their patients in the basic features of age and gender. Our study uses different (slightly simpler) stimuli to Hägele and colleagues and, while both studies require a button press, in ours, the reinforcement is not actually contingent on the button press reaction time (as discussed above, this was suggested by Treadway and Zald (2011) as being advantageous in reducing anticipatory anxiety). Our study

includes a more detailed assessment of psychopathology (with the limitations that, as previously mentioned, when utilizing the psychopathology for correlation analyses we did not correct for multiple correlations, and that our sample size is modest for correlation analysis).

Conclusion

In summary, while a reduced activation in the ventral striatum when anticipating rewards is a common endophenotype in psychopathology, the mechanisms underpinning this finding and related symptoms are not completely clear. It will be crucial to further pinpoint the clinical relevance of this finding but it will require further studies and replications. Although some evidence from both the previous literature and our work points toward negative or depressive symptoms being more related to the reported finding, they require further confirmation. The future of this line of research fits nicely within the specifications laid out by the RDoC project, although it also faces similar challenges, such as the measurement error, and the biological and psychometric limitations of proposed endophenotypes and their relationship to behavior (Cuthbert, 2014b; Lilienfeld, 2014; Weinberger and Goldberg, 2014). Upcoming studies on reward and psychopathology will have to use bigger sample sizes and a broader range of clinical measurements in order to be able to obtain a compelling evidence of the relationship between brain activation and everyday behavior. As shown in our work, future studies could benefit from including participants with a range of diagnoses. This aim can be best achieved by a large-scale collaboration across different research groups. Moreover, the wide use of the MID task makes it a good candidate measure for such collaboration, although a common “official” version would be important. Our results indicate that a paradigm that does not base reward on performance might be better fitted for research with stress-prone participants.

Our overall findings are further evidence of reward system dysfunction across the neuropsychiatric continuum, even if the specific clinical relevance is still not fully understood. Studies on this line of fruitful research could provide new insights on the cross-diagnostic mechanisms of psychopathological symptoms, especially if conducted in a way that minimizes the challenges posed to the Research Domain Criteria approach.

Acknowledgments

Supported by the Wellcome Trust Institutional Strategic Support Fund [097814/Z/11], a MRC Clinician Scientist [G0701911], a Brain and Behavior Research Foundation Young Investigator, and an Isaac Newton Trust award to GM; an award to NS from the Secretary for Universities and Research of the Ministry of Economy and Knowledge of the Government of Catalonia and the European Union; by the University of Cambridge Behavioural and Clinical Neuroscience Institute, funded by a joint award from the Medical Research Council [G1000183] and Wellcome Trust [093875/Z/10/Z]; by awards from the Wellcome Trust [095692] and the Bernard Wolfe Health Neuroscience

Fund to PE, and by the Cambridge NIHR Biomedical Research Centre. The authors are grateful for the help of clinical staff in CAMEO, in the Cambridge Rehabilitation and Recovery service and Pathways, and in the Cambridge IAPT service, for help with participant recruitment.

References

- American Psychiatric Association. (2000). *Diagnostic and Statistical Manual of Mental Disorders*. Washington, DC: American Psychiatric Association.
- Argyropoulos, S. V., and Nutt, D. J. (2013). Anhedonia revisited: is there a role for dopamine-targeting drugs for depression? *J. Psychopharmacol.* 27, 869–877. doi: 10.1177/0269881113494104
- Arrondo, G., Murray, G. K., Hill, E., Szalma, B., Yathiraj, K., Denman, C., et al. (2015). Hedonic and disgust taste perception in borderline personality disorder and depression. *Br. J. Psychiatry* 207, 79–80. doi: 10.1192/bjp.bp.114.150433
- Barch, D. M., and Dowd, E. C. (2010). Goal representations and motivational drive in schizophrenia: the role of prefrontal-striatal interactions. *Schizophr. Bull.* 36, 919–934. doi: 10.1093/schbul/sbq068
- Barch, D. M., Treadway, M. T., and Schoen, N. (2014). Effort, anhedonia, and function in schizophrenia: reduced effort allocation predicts amotivation and functional impairment. *J. Abnorm. Psychol.* 123, 387–397. doi: 10.1037/a0036299
- Beck, A. T., Steer, R. A., Ball, R., and Ranieri, W. (1996). Comparison of Beck Depression Inventories -IA and -II in psychiatric outpatients. *J. Pers. Assess.* 67, 588–597. doi: 10.1207/s15327752jpa6703_13
- Bedwell, J. S., Gooding, D. C., Chan, C. C., and Trachik, B. J. (2014). Anhedonia in the age of RDoC. *Schizophr. Res.* 160, 226–227. doi: 10.1016/j.schres.2014.10.028
- Bernacer, J., Corlett, P. R., Ramachandra, P., McFarlane, B., Turner, D. C., Clark, L., et al. (2013). Methamphetamine-induced disruption of frontostriatal reward learning signals: relation to psychotic symptoms. *Am. J. Psychiatry* 170, 1326–1334. doi: 10.1176/appi.ajp.2013.12070978
- Berns, G. S., McClure, S. M., Pagnoni, G., and Montague, P. R. (2001). Predictability modulates human brain response to reward. *J. Neurosci.* 21, 2793–2798.
- Berrios, G. E. (1996). *The History of Mental Symptoms*. Cambridge: Cambridge University Press.
- Cattell, R. B., Krug, S. E., and Barton, K. (1973). *Technical Supplement for the Culture Fair Intelligence Tests, Scales 2 and 3*. Champaign, IL: Institute for Personality and Ability Testing.
- Cohen, A. S., and Minor, K. S. (2010). Emotional experience in patients with schizophrenia revisited: meta-analysis of laboratory studies. *Schizophr. Bull.* 36, 143–150. doi: 10.1093/schbul/sbn061
- Cools, R., Blackwell, A., Clark, L., Menzies, L., Cox, S., and Robbins, T. W. (2005). Tryptophan depletion disrupts the motivational guidance of goal-directed behavior as a function of trait impulsivity. *Neuropsychopharmacology* 30, 1362–1373. doi: 10.1038/sj.npp.1300704
- Crespi, L. P. (1942). Quantitative variation of incentive and performance in the white rat. *Am. J. Psychol.* 55, 467–517. doi: 10.2307/1417120
- Cuthbert, B. N. (2014a). The RDoC framework: facilitating transition from ICD/DSM to dimensional approaches that integrate neuroscience and psychopathology. *World Psychiatry* 13, 28–35. doi: 10.1002/wps.20087
- Cuthbert, B. N. (2014b). Translating intermediate phenotypes to psychopathology: the NIMH research domain criteria. *Psychophysiology* 51, 1205–1206. doi: 10.1111/psyp.12342
- Der-Avakian, A., and Markou, A. (2012). The neurobiology of anhedonia and other reward-related deficits. *Trends Neurosci.* 35, 68–77. doi: 10.1016/j.tins.2011.11.005
- Desikan, R. S., Ségonne, F., Fischl, B., Quinn, B. T., Dickerson, B. C., Blacker, D., et al. (2006). An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. *Neuroimage* 31, 968–980. doi: 10.1016/j.neuroimage.2006.01.021
- Fletcher, P. C., and Frith, C. D. (2009). Perceiving is believing: a Bayesian approach to explaining the positive symptoms of schizophrenia. *Nat. Rev. Neurosci.* 10, 48–58. doi: 10.1038/nrn2536
- Gard, D. E., Gard, M. G., Kring, A. M., and John, O. P. (2006). Anticipatory and consummatory components of the experience of pleasure: a scale development study. *J. Res. Pers.* 40, 1086–1102. doi: 10.1016/j.jrp.2005.11.001
- Gard, D. E., Sanchez, A. H., Cooper, K., Fisher, M., Garrett, C., and Vinogradov, S. (2014). Do people with schizophrenia have difficulty anticipating pleasure, engaging in effortful behavior, or both? *J. Abnorm. Psychol.* 123, 771–782. doi: 10.1037/abn0000005
- Gradin, V. B., Kumar, P., Waiter, G., Ahearn, T., Stickle, C., Milders, M., et al. (2011). Expected value and prediction error abnormalities in depression and schizophrenia. *Brain* 134, 1751–1764. doi: 10.1093/brain/awr059
- Hägele, C., Schlagenhauf, F., Rapp, M., Sterzer, P., Beck, A., Bermpohl, F., et al. (2015). Dimensional psychiatry: reward dysfunction and depressive mood across psychiatric disorders. *Psychopharmacology (Berl)* 232, 331–341. doi: 10.1007/s00213-014-3662-7
- Hui, C., Morcillo, C., Russo, D. A., Stochl, J., Shelley, G. F., Painter, M., et al. (2013). Psychiatric morbidity, functioning and quality of life in young people at clinical high risk for psychosis. *Schizophr. Res.* 148, 175–180. doi: 10.1016/j.schres.2013.05.026
- Insel, T., Cuthbert, B., Garvey, M., Heinssen, R., Pine, D. S., Quinn, K., et al. (2010). Research domain criteria (RDoC): toward a new classification framework for research on mental disorders. *Am. J. Psychiatry* 167, 748–751. doi: 10.1176/appi.ajp.2010.09091379
- Juckel, G., Schlagenhauf, F., Koslowski, M., Filonov, D., Wustenberg, T., Villringer, A., et al. (2006a). Dysfunction of ventral striatal reward prediction in schizophrenic patients treated with typical, not atypical, neuroleptics. *Psychopharmacology (Berl)* 187, 222–228. doi: 10.1007/s00213-006-0405-4
- Juckel, G., Schlagenhauf, F., Koslowski, M., Wustenberg, T., Villringer, A., Knutson, B., et al. (2006b). Dysfunction of ventral striatal reward prediction in schizophrenia. *Neuroimage* 29, 409–416. doi: 10.1016/j.neuroimage.2005.07.051
- Kay, S. R., Fiszbein, A., and Opler, L. A. (1987). The positive and negative syndrome scale (PANSS) for schizophrenia. *Schizophr. Bull.* 13, 261–276. doi: 10.1093/schbul/13.2.261
- Kirsch, P., Schienle, A., Stark, R., Sammer, G., Blecker, C., Walter, B., et al. (2003). Anticipation of reward in a nonaversive differential conditioning paradigm and the brain reward system: an event-related fMRI study. *Neuroimage* 20, 1086–1095. doi: 10.1016/S1053-8119(03)00381-1
- Knutson, B., Bhanji, J. P., Cooney, R. E., Atlas, L. Y., and Gotlib, I. H. (2008). Neural responses to monetary incentives in major depression. *Biol. Psychiatry* 63, 686–692. doi: 10.1016/j.biopsych.2007.07.023
- Knutson, B., Fong, G. W., Adams, C. M., Varner, J. L., and Hommer, D. (2001). Dissociation of reward anticipation and outcome with event-related fMRI. *Neuroreport* 12, 3683–3687. doi: 10.1097/00001756-200112040-00016
- Kring, A. M., and Barch, D. M. (2014). The motivation and pleasure dimension of negative symptoms: neural substrates and behavioral outputs. *Eur. Neuropsychopharmacol.* 24, 725–736. doi: 10.1016/j.euroneuro.2013.06.007
- Kring, A. M., and Caponigro, J. M. (2010). Emotion in schizophrenia: where feeling meets thinking. *Curr. Dir. Psychol. Sci.* 19, 255–259. doi: 10.1177/0963721410377599
- Kroken, R. A., Johnsen, E., Ruud, T., Wentzel-Larsen, T., and Jørgensen, H. A. (2009). Treatment of schizophrenia with antipsychotics in Norwegian emergency wards, a cross-sectional national study. *BMC Psychiatry* 9:24. doi: 10.1186/1471-244X-9-24

Supplementary Material

The Supplementary Material for this article can be found online at: <http://journal.frontiersin.org/article/10.3389/fpsyg.2015.01280>

- Kumar, P., Waiter, G., Ahearn, T., Milders, M., Reid, I., and Steele, J. D. (2008). Abnormal temporal difference reward-learning signals in major depression. *Brain* 131, 2084–2093. doi: 10.1093/brain/awn136
- Lilienfeld, S. O. (2014). The Research Domain Criteria (RDoC): an analysis of methodological and conceptual challenges. *Behav. Res. Ther.* 62, 129–139. doi: 10.1016/j.brat.2014.07.019
- Murray, G. K., Clark, L., Corlett, P. R., Blackwell, A. D., Cools, R., Jones, P. B., et al. (2008a). Incentive motivation in first-episode psychosis: a behavioural study. *BMC Psychiatry* 8:34. doi: 10.1186/1471-244X-8-34
- Murray, G. K., Corlett, P. R., Clark, L., Pessiglione, M., Blackwell, A. D., Honey, G., et al. (2008b). Substantia nigra/ventral tegmental reward prediction error disruption in psychosis. *Mol Psychiatry* 13, 239, 267–276. doi: 10.1038/sj.mp.4002058
- Murray, G. K., and Jones, P. B. (2012). Psychotic symptoms in young people without psychotic illness: mechanisms and meaning. *Br. J. Psychiatry* 201, 4–6. doi: 10.1192/bjp.bp.111.107789
- Myerson, A. (1922). Anhedonia. *Am. J. Psychiatry* 79, 87–103. doi: 10.1176/ajp.79.1.87
- Nielsen, M. Ø., Rostup, E., Wulff, S., Bak, N., Lublin, H., Kapur, S., et al. (2012). Alterations of the brain reward system in antipsychotic naive schizophrenia patients. *Biol. Psychiatry* 71, 898–905. doi: 10.1016/j.biopsych.2012.02.007
- O'doherty, J., Dayan, P., Schultz, J., Deichmann, R., Friston, K., and Dolan, R. J. (2004). Dissociable roles of ventral and dorsal striatum in instrumental conditioning. *Science* 304, 452–454. doi: 10.1126/science.1094285
- Overall, J. E., and Gorham, D. R. (1962). The brief psychiatric rating scale. *Psychol. Rep.* 10, 799–812. doi: 10.2466/pr0.1962.10.3.799
- Pessiglione, M., Seymour, B., Flandin, G., Dolan, R. J., and Frith, C. D. (2006). Dopamine-dependent prediction errors underpin reward-seeking behaviour in humans. *Nature* 442, 1042–1045. doi: 10.1038/nature05051
- Pizzagalli, D. A., Holmes, A. J., Dillon, D. G., Goetz, E. L., Birk, J. L., Bogdan, R., et al. (2009). Reduced caudate and nucleus accumbens response to rewards in unmedicated individuals with major depressive disorder. *Am. J. Psychiatry* 166, 702–710. doi: 10.1176/appi.ajp.2008.08081201
- Rømer Thomsen, K., Whybrow, P. C., and Kringelbach, M. L. (2015). Reconceptualizing anhedonia: novel perspectives on balancing the pleasure networks in the human brain. *Front. Behav. Neurosci.* 9:49. doi: 10.3389/fnbeh.2015.00049
- Schlagenhauf, F., Juckel, G., Koslowski, M., Kahnt, T., Knutson, B., Dembler, T., et al. (2008). Reward system activation in schizophrenic patients switched from typical neuroleptics to olanzapine. *Psychopharmacology (Berl)* 196, 673–684. doi: 10.1007/s00213-007-1016-4
- Sescousse, G., Caldú, X., Segura, B., and Dreher, J. C. (2013). Processing of primary and secondary rewards: a quantitative meta-analysis and review of human functional neuroimaging studies. *Neurosci. Biobehav. Rev.* 37, 681–696. doi: 10.1016/j.neubiorev.2013.02.002
- Sheehan, D. V., Lecrubier, Y., Sheehan, K. H., Amorim, P., Janavs, J., Weiller, E., et al. (1998). The Mini-International Neuropsychiatric Interview (M.I.N.I.): the development and validation of a structured diagnostic psychiatric interview for DSM-IV and ICD-10. *J. Clin. Psychiatry* 59(Suppl. 20), 22–33; quiz: 34–57.
- Sherdell, L., Waugh, C. E., and Gotlib, I. H. (2012). Anticipatory pleasure predicts motivation for reward in major depression. *J. Abnorm. Psychol.* 121, 51–60. doi: 10.1037/a0024945
- Simon, J. J., Biller, A., Walther, S., Roesch-Ely, D., Stippich, C., Weisbrod, M., et al. (2010). Neural correlates of reward processing in schizophrenia-relationship to apathy and depression. *Schizophr. Res.* 118, 154–161. doi: 10.1016/j.schres.2009.11.007
- Smoski, M. J., Felder, J., Bizzell, J., Green, S. R., Ernst, M., Lynch, T. R., et al. (2009). fMRI of alterations in reward selection, anticipation, and feedback in major depressive disorder. *J. Affect. Disord.* 118, 69–78. doi: 10.1016/j.jad.2009.01.034
- Snaith, R. P., Hamilton, M., Morley, S., Humayan, A., Hargreaves, D., and Trigwell, P. (1995). A scale for the assessment of hedonic tone the Snaith-Hamilton Pleasure Scale. *Br. J. Psychiatry* 167, 99–103. doi: 10.1192/bjp.167.1.99
- Stoy, M., Schlagenhauf, F., Sterzer, P., Bermpohl, F., Hägele, C., Suchotzki, K., et al. (2012). Hyporeactivity of ventral striatum towards incentive stimuli in unmedicated depressed patients normalizes after treatment with escitalopram. *J. Psychopharmacol.* 26, 677–688. doi: 10.1177/0269881111416686
- Treadway, M. T., Bossaller, N. A., Shelton, R. C., and Zald, D. H. (2012). Effort-based decision-making in major depressive disorder: a translational model of motivational anhedonia. *J. Abnorm. Psychol.* 121, 553–558. doi: 10.1037/a0028813
- Treadway, M. T., and Zald, D. H. (2011). Reconsidering anhedonia in depression: lessons from translational neuroscience. *Neurosci. Biobehav. Rev.* 35, 537–555. doi: 10.1016/j.neubiorev.2010.06.006
- Walter, H., Kammerer, H., Frasch, K., Spitzer, M., and Abler, B. (2009). Altered reward functions in patients on atypical antipsychotic medication in line with the revised dopamine hypothesis of schizophrenia. *Psychopharmacology (Berl)* 206, 121–132. doi: 10.1007/s00213-009-1586-4
- Waltz, J. A., Schweitzer, J. B., Gold, J. M., Kurup, P. K., Ross, T. J., Salmeron, B. J., et al. (2009). Patients with schizophrenia have a reduced neural response to both unpredictable and predictable primary reinforcers. *Neuropsychopharmacology* 34, 1567–1577. doi: 10.1038/npp.2008.214
- Waltz, J. A., Schweitzer, J. B., Ross, T. J., Kurup, P. K., Salmeron, B. J., Rose, E. J., et al. (2010). Abnormal responses to monetary outcomes in cortex, but not in the basal ganglia, in schizophrenia. *Neuropsychopharmacology* 35, 2427–2439. doi: 10.1038/npp.2010.126
- Weinberger, D. R., and Goldberg, T. E. (2014). RDOcs redux. *World Psychiatry* 13, 36–38. doi: 10.1002/wps.20096
- Whitton, A. E., Treadway, M. T., and Pizzagalli, D. A. (2015). Reward processing dysfunction in major depression, bipolar disorder and schizophrenia. *Curr. Opin. Psychiatry* 28, 7–12. doi: 10.1097/YCO.000000000000122
- Woo, C. W., Krishnan, A., and Wager, T. D. (2014). Cluster-extent based thresholding in fMRI analyses: pitfalls and recommendations. *Neuroimage* 91, 412–419. doi: 10.1016/j.neuroimage.2013.12.058
- Zhang, W. N., Chang, S. H., Guo, L. Y., Zhang, K. L., and Wang, J. (2013). The neural correlates of reward-related processing in major depressive disorder: a meta-analysis of functional magnetic resonance imaging studies. *J. Affect. Disord.* 151, 531–539. doi: 10.1016/j.jad.2013.06.039
- Ziauddeen, H., and Murray, G. K. (2010). The relevance of reward pathways for schizophrenia. *Curr. Opin. Psychiatry* 23, 91–96. doi: 10.1097/YCO.0b013e328336661b

Conflict of Interest Statement: Trevor W. Robbins has received research support from or served as a consultant to Cambridge Cognition, Eli Lilly, GlaxoSmithKline, and Lundbeck. Hisham Ziauddeen has been jointly funded by the Wellcome Trust and GlaxoSmithKline on the Translational Medicine and Therapeutics programme. Paul C. Fletcher has received funds from GlaxoSmithKline for consultation services and from Astra Zeneca for a lecture. Gonzalo Arrondo, Nuria Segarra, Antonio Metastasio, Jennifer Spencer, Robert B. Dudas, and Graham K. Murray and Niels R. Reinders have no financial interests.

Copyright © 2015 Arrondo, Segarra, Metastasio, Ziauddeen, Spencer, Reinders, Dudas, Robbins, Fletcher and Murray. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



The impact of threat of shock on the framing effect and temporal discounting: executive functions unperturbed by acute stress?

Oliver J. Robinson*, Rebecca L. Bond and Jonathan P. Roiser

Institute of Cognitive Neuroscience, University College London, London, UK

OPEN ACCESS

Edited by:

Nikolina Skandal, University of Cambridge, UK

Reviewed by:

Xochitl Angelica Ortiz, Universidad Autónoma de Nuevo León, Mexico

Scott A. Langenecker, The University of Illinois at Chicago, USA

*Correspondence:

Oliver J. Robinson, Institute of Cognitive Neuroscience, University College London, Alexandra House, 17 Queen Square, London, UK
oliver.j.robinson@gmail.com

Specialty section:

This article was submitted to Psychology for Clinical Settings, a section of the journal Frontiers in Psychology

Received: 22 May 2015

Accepted: 17 August 2015

Published: 31 August 2015

Citation:

Robinson OJ, Bond RL and Roiser JP (2015) The impact of threat of shock on the framing effect and temporal discounting: executive functions unperturbed by acute stress? *Front. Psychol.* 6:1315. doi: 10.3389/fpsyg.2015.01315

Anxiety and stress-related disorders constitute a large global health burden, but are still poorly understood. Prior work has demonstrated clear impacts of stress upon basic cognitive function: biasing attention toward unexpected and potentially threatening information and instantiating a negative affective bias. However, the impact that these changes have on higher-order, executive, decision-making processes is unclear. In this study, we examined the impact of a translational within-subjects stress induction (threat of unpredictable shock) on two well-established executive decision-making biases: the framing effect ($N = 83$), and temporal discounting ($N = 36$). In both studies, we demonstrate (a) clear subjective effects of stress, and (b) clear executive decision-making biases but (c) no impact of stress on these decision-making biases. Indeed, Bayes factor analyses confirmed substantial preference for decision-making models that *did not* include stress. We posit that while stress may induce subjective mood change and alter low-level perceptual and action processes (Robinson et al., 2013c), some higher-level executive processes remain unperturbed by these impacts. As such, although stress can induce a transient affective biases and altered mood, these need not result in poor financial decision-making.

Keywords: threat of shock, stress, temporal discounting, framing effect, anxiety, depression, executive function, Bayesian models

Introduction

Stress can significantly alter the way that we perceive and react to the world, promoting the processing of threatening and unexpected information (Robinson et al., 2013b,c). This threat bias can be adaptive – improving the ability to detect and avoid further sources of stress – but this bias also likely contributes, at least in part, to the facilitatory role that stress plays in the onset of mood and anxiety disorders (Kendler et al., 2004). Threat of unpredictable shock is a reliable within-subject method of inducing stress in both humans and experimental animals (Grillon, 2008; Davis et al., 2010). While the impact of threat of shock on basic perceptual processes is relatively well studied, its impact on higher-level executive processes such as decision-making is surprisingly poorly understood (Robinson et al., 2013c). Here, we explore the impact of threat of shock on two classic decision-making biases: the framing effect and temporal discounting.

Threat of shock is a translational stress-induction procedure in which an individual anticipates an unpredictable and unpleasant electrical shock (Schmitz and Grillon, 2012). In animals threat

of shock has been shown to engage neural circuitry distinct from that engaged during fear conditioning (Davis et al., 2010), another widely used but conceptually different aversive processing paradigm. More precisely, *anxiety* (or stress) is operationally defined as the prolonged apprehensive response to a context in which threats *may* occur, whereas *fear* is the acute response to a discrete, defined and *predictable* aversive stimulus or cue (Davis et al., 2010; Robinson et al., 2013c). Critically, stress induced by threat of shock has well documented psychological (Robinson et al., 2013c), psychophysiological (Grillon et al., 1991), and neural effects (Cornwell et al., 2007; Robinson et al., 2012, 2013b). Perhaps more importantly there is also emerging evidence that threat of shock evokes mechanisms related to those that participate in pathological anxiety, for example Generalized Anxiety Disorder (Robinson et al., 2013c, 2014). As such, it is hoped that exploring the impact of threat of shock on cognitive functions may also provide a window into the mechanisms which contribute to pathological stress-related disorders in healthy individuals, prior to disorder onset (Robinson et al., 2014).

Executive function is an umbrella term encompassing cognitive functions that do more than passively process information. Such non-automatic functions might integrate information from multiple sensory domains along with information stored in memory. Executive function therefore encompasses higher order processes, such as planning and decision-making. The impact of threat of shock on executive function has not been comprehensively studied. Here we explore one aspect of executive function: financial decision-making. We are aware of only two studies in which threat of shock was shown to alter financial choices. In the first, threat of shock promoted risk-avoidant decision-making (Clark et al., 2012). However, in this study, the threat cues were discrete, of short duration (5–5.5 s) and possibly more comparable to a fear cue than an anxiety/stress condition. In the second study (Robinson et al., 2015), threat of shock had no main effect on gambling choices, but did interact with trait anxiety, promoting ‘harm-avoidant’ (i.e., playing less ‘disadvantageous’ decks) under stress in those with the low anxiety symptoms. However, the Iowa gambling task used in this second study confounds a number of decision-making and basic cognitive processes, making the causes of this result unclear. In the small number of remaining studies that have addressed this question, the effects seemed to be largely restricted to reaction times (Murphy, 1959; Keinan, 1987; Engelmann et al., 2015), with stress having no impact on the decisions themselves.

Microeconomic theory has outlined a number of biases – or heuristics – which have been shown to guide individual financial decision-making behavior. In this study, we explore two well-established biases: the framing effect and temporal discounting. The framing effect describes the reliable propensity of individuals to alter their decisions, dependent on whether the same choice is ‘framed’ as a loss or a gain. Specifically, individuals tend to avoid risk when a choice is framed as a gain (e.g., keep £2 out of £10 vs. a 20% chance to win £10), and become risk-seeking when the exact same choice is framed as a loss (e.g., lose £8 out of £10 vs. a 20% chance to win £10). That is to say, all other things being equal, individuals are biased against certain outcomes

‘framed’ as losses (Kahneman and Tversky, 1979; De Martino et al., 2006). Temporal discounting is another such bias, in which an individual assigns less value to gains in the future relative to the present (Rachlin and Green, 1972; Berns et al., 2007). For instance, offered £10 today and £11 in a month, there is a bias toward accepting the lower value of £10 now. In other words, temporal distance causes devaluation of potential gains. In both paradigms, the subjective utility of financially identical options is biased by the context in which the options are presented. Given that both of these biases can be financially suboptimal and result in reduced gains/increased losses, it is plausible that they might be shifted by contexts such as stress-induced biases toward negative stimuli (Robinson et al., 2013c). Clinical support for this hypothesis comes from the observation of altered decision-making and negative biases in disorders associated with anxiety and negative affect such as major depression (Eshel and Roiser, 2010; and it should be noted that stress, negative mood, and anxiety are relatively diffuse, likely overlapping, concepts).

Therefore, we explored the impact of stress on the framing effect and temporal discounting. We predicted that threat of shock would induce a state of adaptive anxiety (Robinson et al., 2013c) and promote harm-avoidant decisions (Clark et al., 2012; Robinson et al., 2013c), thereby increasing both framing and temporal discounting. Specifically, in the context of uncertain threat an individual might be more loss averse, resulting in an increased framing effect (Porcelli and Delgado, 2009). At the same time in the context of an uncertain future, an individual might be biased toward immediate vs. future gains resulting in increased temporal discounting (Pulcu et al., 2014). This study explored these hypotheses with conventional significance testing as well as a Bayesian approach to enable a more nuanced comparison of different behavioral models.

Materials and Methods

Sample and Screening

Participants were recruited from the UCL Institute of Cognitive Neuroscience Subject Database, of which $N = 83$ (49 female: 34 male; mean age = 24, $SD = 5$) completed the framed gamble task and $N = 36$ (18 female: 18 male; mean age = 24, $SD = 6$) completed the temporal discounting task ($N = 35$ completed both). All participants completed a prior phone screen in which they reported no personal history of/treatment for psychiatric or neurological disorders or drug use (from a detailed specific checklist of all disorders), along with no cardiovascular problems, pacemakers, or cochlear implants. The demographics represented the naturalistic sample of individuals who responded to our call for participants and who passed screening. All subjects provided written informed consent (UCL Research Ethics Committee Project ID Number: 1764/001). Both decision tasks were presented on a desktop computer using the Cogent toolbox for Matlab (Wellcome Trust Centre for Neuroimaging and Institute of Cognitive Neuroscience, UCL, London, UK). To incentivise performance subjects were informed that additional compensation would be provided based upon task performance. Shocks were delivered to the non-dominant wrist using a DS7

stimulator (Digitimer Ltd, Welwyn Garden City, UK). Prior to testing, all subjects completed a shock work-up procedure in which shocks were titrated (over approximately 3–5 stimulations) to a level that was ‘unpleasant but not painful’ (Schmitz and Grillon, 2012).

Anxiety Measures

At the end of each block, participants indicated how anxious they had felt during each of the threat and safe conditions on a scale from 1 (“not at all”) to 10 (“very much so”) as a subjective manipulation check. Participants also provided self-report measures of depression (Beck Depression Inventory: BDI; Beck and Steer, 1987) and trait anxiety (State Trait Anxiety Inventory: STAI; Spielberger et al., 1970) at the end of the session.

Framed Gamble Task

This task was adapted from that used by De Martino et al. (2006). The task consisted of eight blocks (four safe, four threat), each comprising 14 trials. “YOU ARE NOW SAFE FROM SHOCK” or “YOU ARE NOW AT RISK OF SHOCK” was presented for 3 s at the beginning of each new block. Threat blocks had a red background, whilst safe blocks had a blue background. A single shock was delivered at a pseudorandom time in each threat block. Each trial began with a message “You receive £X” where, X = varying monetary amounts. Participants then had 4 s to choose between a certain option, which would leave them with a guaranteed portion of the total £X, or an option to gamble, which could lead to either winning the entire amount or winning nothing. The participant did not discover the outcome of any gambles, but was instructed to consider which option they would choose to maximize wins and minimize losses. In gain frames the participant would have the ‘sure’ option to “Keep £Y” a certain portion of the total whereas in loss frames participants were told they would “Lose £Z” (where $Z + Y = X$), implying that they would retain the rest of the total (i.e., £Y) – note that this represents precisely the same decision. Alternatively, participants could choose a gamble option which was presented with a pie chart indicating the probability of each keeping or losing the entire £X amount. In experimental trials, the expected values (sum of all possible outcomes weighted by their respective probabilities) of the gamble and sure options were matched (Figure 1A). Expected outcomes were 20, 40, 60, or 80% of the initial total £X, which was set as £25, £50, £75, or £100. Monetary parameters were counterbalanced across decision frames and between threat and safe blocks. ‘Catch’ trials were also included to verify that the participants were attending to and had understood the task. These trials were designed such that the expected outcome of one option was much larger than that of the other option, such that participants should always choose the option of higher value.

Each block consisted of 10 standard trials (five in each frame) and with four catch trials (one in each combination of frame and preferred option) in a random order, with the certain and gamble options randomized to the left or right of the screen. Choices were indicated by pressing the left or right arrow key. The chosen option was highlighted by a star for 1 s. Analysis was conducted on the proportion of trials on which participants chose to gamble,

and reaction times to choose each of the gamble and the certain options. Before the main task, six practice trials (one of each type of catch trial, plus a standard trial in each the gains and losses frames) were completed without a time limit or threat of shock to ensure participants understood the task. Task duration was approximately 15 min.

Temporal Discounting Task

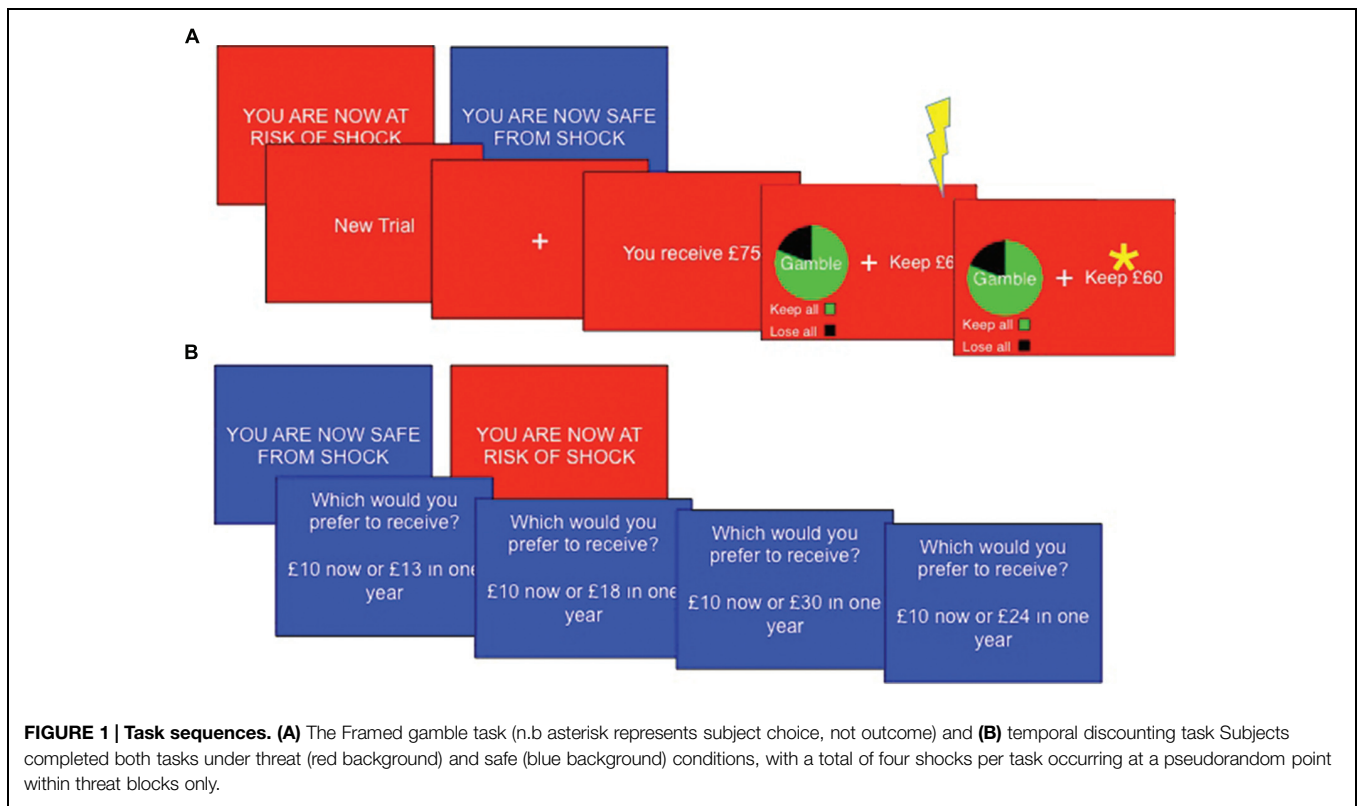
On each trial subjects were presented with a (self-paced) choice between an immediate reward (e.g., “£5.00 now”) and a delayed reward (e.g., “£10.00 in 25 years”). The value of the delayed reward was fixed whilst the immediate reward was adjusted based upon previous choices until an indifference point was reached. Adjustment was based on Yi et al. (2010) abbreviation of Johnson and Bickel’s (2002) algorithm. Indifference points were recorded for three delayed values (£10, £100, and £1000), three delays (1 day, 1 year, and 25 years) and for both gain and loss frames (“Which would you prefer to receive?” or “Which would you prefer to lose?”), yielding a total of 18 indifference points (Figure 1B). The task consisted of six blocks (three safe, three threat), during each of which six (randomly selected without replacement) indifference points were reached. As the algorithm adjusted the choices presented based upon previous responses, the number of trials per block varied (range ~60–80). The screen displayed “YOU ARE NOW SAFE FROM SHOCK” with a blue background at the start of safe blocks or “YOU ARE NOW AT RISK OF SHOCK” with a red background at the start of threat blocks. Shocks ($N = 4$) were presented in a pseudo-random order during threat blocks. Task duration was approximately 15 min.

Analysis

Conventional frequentist significance tests were run in SPSS version 22 (IBM Corp, Armonk, NY) whilst Bayesian analyses were run in JASP, employing the default prior (Rouder et al., 2012; Love et al., 2015; Morey and Rouder, 2015). Frequentist and Bayesian repeated-measures analysis of variance (ANOVA) models were constructed in exactly the same manner for all analyses (see below), with frequentist ANOVAs used to generate F-statistics, p-values and effect sizes for interactions of interest, and Bayesian ANOVAs used to generate log Bayes factors ($\log BF_{10}$)¹ for models of interest relative to a null model (main effect of subject).

In our Bayesian analyses, the ‘winning’ model was defined as the model with the highest BF_{10} relative to the null, and the relative predictive success of one model over another was computed by dividing the BF_{10} for one model by the other. Any value greater than zero indicates a model *better* than the comparison. Semantic labels were assigned to the magnitude of these comparisons to aid interpretation, ranging from anecdotal (1–3), to substantial (3–10), to strong (10–30) to decisive (> 100 ; Jeffreys, 1998). Where reported for interactions, the Bayes factors

¹Note that the ‘ BF_{10} ’ nomenclature in JASP refers to the Bayes factor for H1 vs. H0 (model relative to null); as distinct from ‘ BF_{01} ’, which is the Bayes factor for H0 vs. H1 (null relative to model). It does *not* refer to \log_{10} . Where we refer to ‘ $\log BF_{10}$ ’, we report the natural log of the BF_{10} values. This log is not required; it is simply used to make the frequently very large numbers more interpretable.



represent a model including the interaction plus the main effect of each component of the interaction.

Statistical Models

Manipulation efficacy was assessed using a paired *t*-test (and Bayesian equivalent) to compare retrospective ratings across stress conditions.

For the *framed gamble task*, the proportion of trials on which participants chose the gamble option was assessed using a repeated-measures ANOVA with stress condition (threat/safe) and decision frame (gains/losses domains) as within-subject factors. Reaction time was analyzed in a similar model, with the addition of choice (gamble/sure) as an additional within-subjects factor.

For the *temporal discounting task*, we analyzed normalized indifference points (indifference point/fixed delayed value) in a repeated-measures ANOVA with stress condition (threat/safe), decision valence (gain/lose), delayed value (£10/£100/£1000) and time (1 day, 1 year, 25 years) as within-subjects factors. Reaction times for choices across each indifference point (mean reaction time of all responses required to reach indifference) were analyzed in a separate model with the same factors.

Finally, for both tasks, additional exploratory between-subjects analyses were also run including measures of mood symptoms and order of threat-safe counterbalancing as covariates. These were separate models, run after the *a priori* within-subject models. BDI symptom data was not normally distributed, and was square-root transformed prior to analysis. Further exploratory analyses suggested during peer review were

also run: gender, age, and threat potentiated (threat minus safe) subjective ratings were included as additional between-subject factors.

Results

Data for these tasks are freely available for download².

Manipulation Check

Subjects reported feeling significantly more anxious in the stress relative to the safe condition in both the framed gamble [a mean rating (\pm SD) of $6 \pm 2/10$ relative to $2 \pm 1/10$; $t(82) = -16$, $p < 0.001$] and temporal discounting tasks [$6 \pm 2/10$ relative to $2 \pm 1/10$; $t(34) = -10$, $p < 0.001$]. Bayes factors indicated that models including stress conditions were decisively better than the null model for both the framed gamble ($\log BF_{10} = 57$) and temporal discounting tasks ($\log BF_{10} = 21$).

Framed Gamble Task

(A) Within-Subjects Effects

Choice behavior

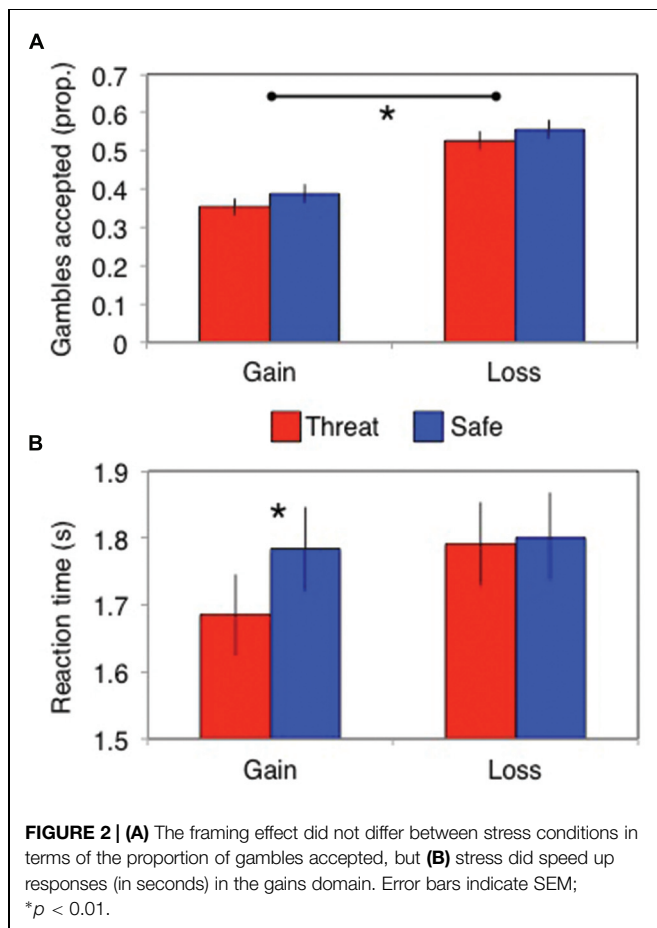
A significant framing effect was demonstrated. Specifically, participants gambled more in the losses frame (probability of gambling = 0.54 ± 0.2) than in the gains frame [probability of gambling = 0.37 ± 0.2 ; main effect of frame: $F(1,82) = 83$, $p < 0.001$, $\eta_p^2 = 0.5$]. However, this did not interact with threat

²<http://dx.doi.org/10.6084/m9.figshare.1423293>

of shock [stress \times frame interaction: $F(1,82) = 0.13$, $p = 0.72$, $\eta_p^2 = 0.002$; **Figure 2A**] and there was no main effect of stress [$F(1,82) = 3.7$, $p = 0.06$, $\eta_p^2 = 0.043$]. Bayes factor analysis revealed the winning model to be one including only a main effect of frame ($\log BF_{10} = 46$), which was substantially (7.6 times) better than a model additionally including the stress \times frame interaction ($\log BF_{10} = 44$).

Reaction times

Subjects (only $N = 74$ had RTs in all cells, since some subjects never chose at least one of the options) were significantly faster to choose in the loss than gain frame [main effect of frame: $F(1,74) = 7.5$, $p = 0.008$, $\eta_p^2 = 0.091$] and under threat of shock [main effect of stress: $F(1,74) = 5.0$, $p = 0.029$, $\eta_p^2 = 0.063$]. These effects were qualified by a significant stress \times frame interaction [$F(1,74) = 7.0$, $p = 0.01$, $\eta_p^2 = 0.086$]. Analyses of the simple main effects revealed that stress induced quicker responses in the gains domain [$F(1,74) = 12$, $p = 0.001$, $\eta_p^2 = 0.14$] but not in the losses domain [$F(1,74) = 0.11$, $p = 0.74$, $\eta_p^2 < 0.002$; **Figure 2B**]. Bayes factor analysis revealed that the winning model comprised individual main effects of stress and frame ($\log BF_{10} = 2.1$), without the interaction; but this was only anecdotally (1.1 times, Jeffreys, 1998) better than the model also including a stress \times frame interaction ($\log BF_{10} = 2.0$).



(B) Between-Subjects Effects

Choice behavior

Neither safe-threat order [frame \times order interaction: $F(1,81) = 0.20$, $p = 0.66$, $\eta_p^2 = 0.002$] nor baseline symptoms [frame \times STAI interaction: $F(1,81) = 1.4$, $p = 0.24$, $\eta_p^2 = 0.017$; frame \times BDI interaction: $F(1,80) = 0.28$, $p = 0.60$, $\eta_p^2 = 0.003$] interacted with any of the main effects of interest. There was no exploratory stress \times frame \times gender interaction [$F(1,81) = 0.07$, $p = 0.8$, $\eta_p^2 = 0.001$], stress \times frame \times age interaction [$F(1,81) = 0.62$, $p = 0.53$, $\eta_p^2 = 0.008$] or stress \times frame \times threat potentiated (threat minus safe) anxiety rating interaction [$F(1,80) = 0.40$, $p = 0.53$, $\eta_p^2 = 0.005$].

Reaction times

Neither safe-threat order [stress \times frame \times order interaction: $F(1,73) = 0.20$, $p = 0.65$, $\eta_p^2 = 0.003$] nor baseline symptoms [stress \times frame \times STAI interaction: $F(1,73) = 0.11$, $p = 0.74$, $\eta_p^2 = 0.001$; stress \times frame \times BDI interaction: $F(1,72) = 1.3$, $p = 0.26$, $\eta_p^2 = 0.018$] interacted with any of the interaction effects of interest.

Temporal Discounting Task

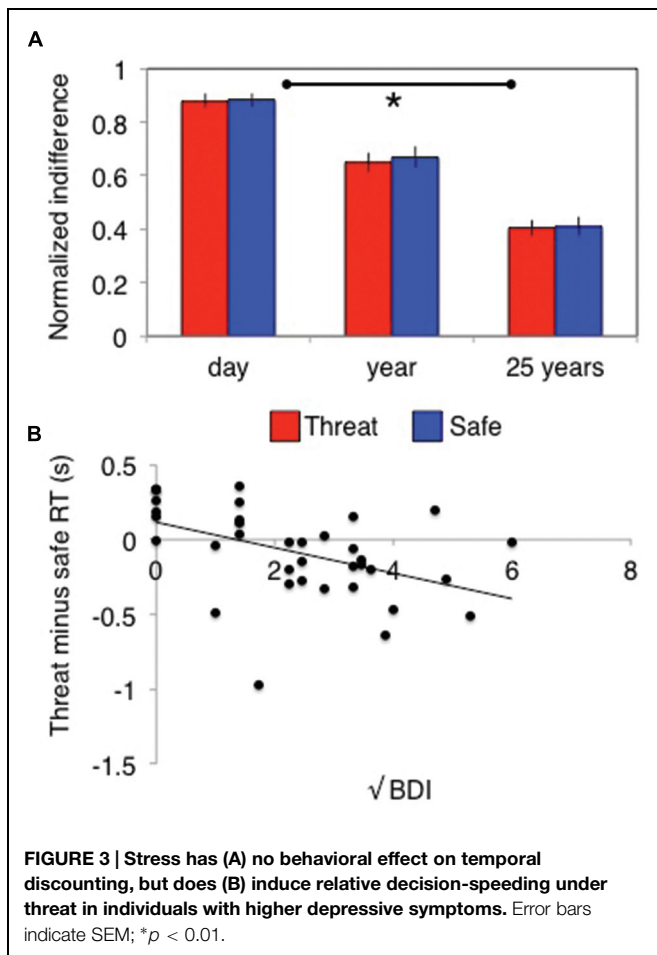
(A) Within-Subjects Effects

Choice behavior

Temporal discounting was demonstrated by a significant main effect of delay on indifference points [$F(2,70) = 79$, $p < 0.001$, $\eta_p^2 = 0.7$]. This varied depending upon whether subjects were asked about wins or losses [time \times valence interaction: $F(2,70) = 8$, $p = 0.001$, $\eta_p^2 = 0.2$] but did not differ across the different values [time \times value interaction: $F(4,140) = 1.3$, $p = 0.26$, $\eta_p^2 = 0.04$]. Critically, this also did not differ under stress [time \times stress interaction: $F(2,70) = 0.24$, $p = 0.79$, $\eta_p^2 = 0.007$, **Figure 3A**; main effect of stress: $F(1,35) = 0.8$, $p = 0.37$, $\eta_p^2 = 0.02$; time \times valence \times stress: $F(2,70) = 0.16$, $p = 0.86$, $\eta_p^2 = 0.004$; time \times valence \times value \times stress: $F(4,140) = 0.73$, $p = 0.58$, $\eta_p^2 = 0.02$]. Bayes factor analysis revealed a winning indifference point model comprising a time by valence interaction ($\log BF_{10} = 294$) that was decisively (> 150 times) better than model additionally including a stress by time interaction ($\log BF_{10} = 264$), a stress by valence by time model ($\log BF_{10} = 271$) or a time alone model ($\log BF_{10} = 271$).

Reaction times

There was a main effect of time on RTs [$F(2,70) = 9.9$, $p < 0.001$, $\eta_p^2 = 0.22$], a main effect of valence [$F(1,35) = 9.6$, $p = 0.004$, $\eta_p^2 = 0.22$] and a significant valence \times time interaction [$F(2,70) = 7.9$, $p < 0.001$, $\eta_p^2 = 0.18$]. Simple main effects analyses revealed that this interaction was driven by a main effect of time in the gain domain [subjects became progressively slower with increasing time: $F(2,34) = 15$, $p < 0.001$, $\eta_p^2 = 0.47$] but not the loss domain [subjects were always as slow as the slowest (i.e., 25 years) time point in the gains domain: $F(2,34) = 1.8$, $p = 0.19$, $\eta_p^2 = 0.09$]. There was no main effect of stress [$F(1,35) = 2.6$, $p = 0.12$, $\eta_p^2 = 0.07$] or stress by time interaction [$F(2,70) = 0.8$, $p = 0.4$, $\eta_p^2 = 0.02$]. There was a trend toward a



stress \times valence interaction [$F(1,35) = 4.1, p = 0.050, \eta_p^2 = 0.11$] but Bayes factor analysis revealed this model ($\log BF_{10} = 9$) to be decisively (>150 times) worse than the winning valence \times time model ($\log BF_{10} = 23$).

(B) Between-Subject Effects

Choice behavior

No effects of interest interacted with task order [for indifference points, time \times order interaction: $F(2,68) = 1.1, p = 0.34, \eta_p^2 = 0.03$] or baseline symptoms (for indifference points time \times STAI interaction: $F(2,68) = 0.51, p = 0.6, \eta_p^2 = 0.02$; time \times BDI interaction: $F(2,68) = 0.27, p = 0.8, \eta_p^2 = 0.008$). There was no exploratory time \times stress \times gender interaction [$F(2,68) = 0.26, p = 0.8, \eta_p^2 = 0.007$], time \times stress \times age interaction [$F(2,68) = 0.13, p = 0.88, \eta_p^2 = 0.004$] or time \times stress \times threat potentiated (threat minus safe) anxiety rating interaction [$F(2,66) = 1.4, p = 0.30, \eta_p^2 = 0.039$].

Reaction times

There was no time \times order interaction [$F(2,68) = 0.94, p = 0.4, \eta_p^2 = 0.03$] or time \times valence \times order interaction [$F(2,68) = 0.043, p = 0.96, \eta_p^2 = 0.001$], but there was a significant stress \times order interaction [$F(1,34) = 16, p < 0.001,$

$\eta_p^2 = 0.32$] driven by those who experienced the safe condition first responding significantly faster under threat [$F(1,34) = 17, p < 0.001, \eta_p^2 = 0.33$; no difference between conditions in those who received threat first: $F(1,34) = 2.5, p = 0.12, \eta_p^2 = 0.07$]. There was no interaction between trait anxiety and any of the effects of interest (all $p > 0.1$), but there was a significant interaction between stress and BDI scores [$F(1,34) = 9.1, p = 0.005, \eta_p^2 = 0.21$] driven by a negative correlation [$r(35) = -0.5$] between the difference between threat and safe RTs and BDI (correlation substantially better model than null: $\log BF_{10} = 2.3$; **Figure 3B**). In other words, the more depression symptoms an individual reported, the faster they responded under threat relative to safe conditions.

Discussion

In this study, we were able to replicate two well-established biases in decision-making: the framing effect and temporal discounting. Moreover, we demonstrated a clear impact of threat of shock on subjective mood and choice reaction times. However, contrary to our predictions, stress did *not* alter the observed decision-making biases, perhaps because these executive decision-making biases are traits that are impervious to, or are able to override, the lower-level state affective biases induced by stress (Robinson et al., 2013c).

We first replicated the well-established framing effect, in which there is a bias toward risky behavior in the losses domain, and toward risk-aversion in the gains domain (Kahneman and Tversky, 1979). However, this bias was not altered by threat of shock in this study. To the best of our knowledge there is no previous literature exploring the impact of threat of shock on this effect, but there are a number of studies exploring the impact of different manipulations on framing. Using the cold pressor task Porcelli and Delgado (2009) found that the framing effect was enhanced by stress relative to a non-stress control condition. Given our sample size ($N = 81$), we had 99.1% statistical power (with $\alpha = 0.05$; two-tailed) to replicate this interaction (with effect size $|d| = 0.487$; Porcelli and Delgado, 2009). One possible explanation for the discrepancy is simply that they used a different manipulation; very little, if any, work has directly compared threat of shock and cold pressor on stress responses (Robinson et al., 2013c). One key difference between paradigms, however, is that the cold pressor is generally completed *prior* to the task, since it requires the individual to submerge their hand in cold water. As such, it is plausible that it explores the impact of *recovery from* stress (Robinson et al., 2013c) rather than a current stressful context (which is a key advantage of the threat of shock technique used here). Another possibility highlighted by the authors is that their effect was a learning effect since the stress block always followed the no-stress block (Porcelli and Delgado, 2009).

Two further studies have explored the effects of another stress manipulation – the Trier social stressor test – on framing. Pabst et al found the *opposite* pattern to Porcelli and Delgado (2009); a *reduced* framing effect under stress (Pabst et al., 2013) albeit only in a subsample of their participants. Buckert et al. (2014)

also showed *reduced* framing under social stress on a game of dice task. This discrepancy across manipulations could be attributed to the specific type of stress; whilst the cold pressor is physically painful, the Trier task asks subjects to ready themselves for unprepared public speaking which is a social form of stress. It is possible that the specific domain of anxiety influences the outcome; e.g., social rewards might be particularly influenced by a social anxiety induction. Nevertheless, adding our null effect with a further stressor to these inconsistent effects of stress leaves the role of stress on framing unclear (**Table 1**). This lack of clarity also highlights the need for work directly comparing different stress manipulations across the same tasks.

In our second experiment we replicated the temporal discounting effect. Specifically, participants assigned less utility to outcomes in the future. Again, however, we failed to detect an impact of stress on this bias. Temporal discounting has not been assessed under threat of shock to our knowledge, but this null finding is consistent with at least three prior studies (Lempert et al., 2012; Haushofer et al., 2013; Jenks and Lawyer, 2015) utilizing the Trier social stressor test. In all three studies stress increased both subjective and/or hormonal indicators of stress, but had no effect on discounting (**Table 1**). Where anxiety *has* been associated with temporal discounting it is in studies exploring *between-subject* individual differences in social anxiety (Rounds et al., 2007) or clinically defined depression (Pulcu et al., 2014). In both cases this is consistent with temporal discounting being a stable, *trait* measure (Odum, 2011). Temporary or acute mood fluctuations such as stress or anxiety may be therefore be unable to influence such traits. That said, it should be noted that we fail to see an interaction with trait depression or anxiety symptoms in our sample and the Rounds effect cited above recently failed to replicate (Jenks and Lawyer, 2015) so firm conclusions are perhaps unwarranted.

We do, however, see some evidence of stress impacting reaction time across both tasks. This is critical because it suggests that the threat of shock manipulation was effective at instantiating behavioral change. In other words, in addition to the subjective reports of anxiety, subjects' responses *were* impacted by threat, even though these did not carry over into their decisions. Reaction time effects in the absence of decision effects are consistent with prior work (Murphy, 1959; Keinan, 1987; Engelmann et al., 2015) and perhaps reflect a bias toward making some decisions faster under conditions of stress. From an evolutionary perspective, such a mechanism could be adaptive: a faster decision about which direction to go when running away from a predator, for instance, may improve survival chances. It should be noted, moreover, that the reaction time correlation with depression symptoms we see in the temporal discounting

reaction time effect, somewhat mimics our previously reported effect in the Iowa Gambling Task (Robinson et al., 2015). In that study, we saw increased selection of disadvantageous decks under threat relative to safe in individuals with high anxiety or depression symptomatology (Robinson et al., 2015). Here we observed (in a partially overlapping sample) quicker responses under threat relative to safe conditions in those who reported greater depressive symptoms. It is plausible that these effects may represent some form of underlying vulnerability in individuals with subclinical depressive symptoms. Having said that, the effect in the present task was not also observed in an relationship with trait anxiety scores (unlike our prior report), which is surprising because trait anxiety and BDI scores are highly correlated in most samples (including this one: $R = 0.8$, $p < 0.001$). Whilst it is possible that the present effect is specific to depression vs. anxiety symptoms we feel that this conclusion would be unwarranted based on the current data, and it requires replication in a larger sample. Reaction time effects can, however, have multiple underlying causes; the effects could be driven by altered decision-making processes, but could also be driven by the time it takes to encode or instantiate a reaction toward a stimulus (Ratcliff and McKoon, 2008) and it is not possible to fully distinguish between these possibilities. Recent work has in fact highlighted the need for researchers to be extremely cautious when using reverse-inference to infer cognitive states from reaction time (Krajovich et al., 2015). Indeed, in general, the exact nature of the reaction time effects seen here were not predicted *a priori*, and in one instance the Bayesian and frequentist tests are partially discrepant (within-subject temporal discounting $p = 0.050$ vs. 150 times worse) and as such we do not wish to draw firm conclusions beyond observing that the effects indicate that the manipulation was having some effect during the tasks.

This raises the question as to why is there are clear affective biases under threat of shock (Robinson et al., 2013c), but that these biases do not impact decisions. One speculation is that this is because such 'bottom-up' biases do not influence some higher level, executive processes. Or, if the biases *are* processed later in the hierarchy, it may be that the executive *overrides* lower level biases. Evolutionarily, an ability of executive function to ignore or override lower level fear and stress responses might be adaptive in certain circumstances. Alternatively, these biases might reflect the use of highly efficient heuristics/rules of thumb which are robust to the effects of stress on affective processing. Lower level affective biases may therefore constitute *state* effects of mood disorders that change with symptoms, whilst the executive decision-making biases constitute stable *traits*. Such traits may contribute to stress-related disorder susceptibility (Pulcu et al., 2014), but not change with mood symptoms. Understanding the distinction between different levels of cognitive function that are impacted by stress might plausibly inform our treatments for stress-related disorders. Specifically, the focus might be on shifting lower-level *state* affective biases rather than *trait* executive biases. Either way, in contrast to our hypotheses the present study provides evidence for the proposition that certain higher order executive decision-making functions are impervious to stress induced by threat of shock.

TABLE 1 | Review of findings cited in this paper (=, null effect; N/A, not cited; ↑, increased effect; ↓, reduced effect).

	Framing	Temporal discounting
Cold pressor	↑	N/A
Social stress	↓	=
Threat of shock	=	=

Limitations

It should be noted that one explanation for our lack of effect is that our stress manipulation was not strong enough to elicit change. Perhaps under conditions of extreme threat (e.g., a warzone or high-stakes work environment) decision-making of this type can be shifted by threat. Alternatively given higher time pressure, higher financial gambles, or explicit feedback about the outcomes of gambles, individual's decisions would have been shifted by threat of shock. Moreover, as discussed above, there are many different 'stress' manipulations across social, pain, and other domains and it is unclear exactly how these overlap. The extent to which this non-significant effect of stress generalizes across stress manipulations is unclear. In addition, these findings do not rule out the impact of threat of shock on other types of decision-making such as those that involves working memory or inhibitory control [both of which have in fact been shown to be influenced by threat (Robinson et al., 2013a,c)]. Overall, a non-significant effect of this nature is difficult to prove as it may simply

be that we have failed to find the correct context in which stress impairs the executive functions explored here. A further limitation is that these individuals were not screened using a diagnostic interview. As such, some individuals may have previous diagnoses which they had forgotten, or may have been missed by using a checklist screening instrument. Finally, these findings do not of course rule out the possibility that positive mood might have effects on these sorts of decision making tasks. Indeed there is preliminary (albeit complex) data suggesting that positive mood can influence temporal discounting (Hirsh et al., 2010) and Iowa Gambling Task performance (De Vries et al., 2008).

Acknowledgments

A Medical Research Council Career Development Award to OR (MR/K024280/1) funded this research. We thank Benedetto De Martino for the original framing task code.

References

- Beck, A. T., and Steer, R. A. (1987). *BDI, Beck Depression Inventory: Manual*. New York, Psychological Corporation.
- Berns, G. S., Laibson, D., and Loewenstein, G. (2007). Intertemporal choice – toward an integrative framework. *Trends Cogn. Sci.* 11, 482–488. doi: 10.1016/j.tics.2007.08.011
- Buckert, M., Schwieler, C., Kudiella, B. M., and Fiebach, C. J. (2014). Acute stress affects risk taking but not ambiguity aversion. *Front. Neurosci.* 8:82. doi: 10.3389/fnins.2014.00082
- Clark, L., Li, R., Wright, C. M., Rome, F., Fairchild, G., Dunn, B. D., et al. (2012). Risk-avoidant decision making increased by threat of electric shock. *Psychophysiology* 49, 1436–1443. doi: 10.1111/j.1469-8986.2012.01454.x
- Cornwell, B. R., Baas, J. M. P., Johnson, L., Holroyd, T., Carver, F. W., Lissek, S., et al. (2007). Neural responses to auditory stimulus deviance under threat of electric shock revealed by spatially-filtered magnetoencephalography. *Neuroimage* 37, 282–289. doi: 10.1016/j.neuroimage.2007.04.055
- Davis, M., Walker, D. L., Miles, L., and Grillon, C. (2010). Phasic vs sustained fear in rats and humans: role of the extended amygdala in fear vs anxiety. *Neuropsychopharmacology* 35, 105–135. doi: 10.1038/npp.2009.109
- De Martino, B., Kumaran, D., Seymour, B., and Dolan, R. J. (2006). Frames, biases, and rational decision-making in the human brain. *Science* 313, 684–687. doi: 10.1126/science.1128356
- De Vries, M., Holland, R. W., and Witteman, C. L. (2008). In the winning mood: affect in the Iowa gambling task. *Judgm. Decis. Mak.* 3, 42–50.
- Engelmann, J. B., Meyer, F., Fehr, E., and Ruff, C. C. (2015). Anticipatory anxiety disrupts neural valuation during risky choice. *J. Neurosci.* 35, 3085–3099. doi: 10.1523/JNEUROSCI.2880-14.2015
- Eshel, N., and Roiser, J. P. (2010). Reward and punishment processing in depression. *Biol. Psychiatry* 68, 118–124. doi: 10.1016/j.biopsych.2010.01.027
- Grillon, C. (2008). Models and mechanisms of anxiety: evidence from startle studies. *Psychopharmacology* 199, 421–437. doi: 10.1007/s00213-007-1019-1
- Grillon, C., Ameli, R., Woods, S. W., Merikangas, K., and Davis, M. (1991). Fear-potentiated startle in humans: effects of anticipatory anxiety on the acoustic blink reflex. *Psychophysiology* 28, 588–595. doi: 10.1111/j.1469-8986.1991.tb01999.x
- Haushofer, J., Cornelisse, S., Seinstra, M., Fehr, E., Joëls, M., and Kalenscher, T. (2013). No effects of psychosocial stress on intertemporal choice. *PLoS ONE* 8:e78597. doi: 10.1371/journal.pone.0078597
- Hirsh, J. B., Guindon, A., Morisano, D., and Peterson, J. B. (2010). Positive mood effects on delay discounting. *Emotion* 10, 717–721. doi: 10.1037/a0019466
- Jeffreys, H. (1998). *The Theory of Probability*. Oxford, Oxford University Press.
- Jenks, C. W., and Lawyer, S. R. (2015). Using delay discounting to understand impulsive choice in socially anxious individuals: failure to replicate. *J. Behav. Ther. Exp. Psychiatry* 46, 198–201. doi: 10.1016/j.jbtep.2014.10.010
- Johnson, M. W., and Bickel, W. K. (2002). Within subject comparison of real and hypothetical money rewards in delay discounting. *J. Exp. Anal. Behav.* 77, 129–146. doi: 10.1901/jeab.2002.77-129
- Kahneman, D., and Tversky, A. (1979). Prospect theory: an analysis of decision under risk. *Econometrica* 47, 263–291. doi: 10.2307/1914185
- Keinan, G. (1987). Decision making under stress: scanning of alternatives under controllable and uncontrollable threats. *J. Pers. Soc. Psychol.* 52, 639–644. doi: 10.1037/0022-3514.52.3.639
- Kendler, K. S., Kuhn, J., and Prescott, C. A. (2004). The interrelationship of neuroticism, sex, and stressful life events in the prediction of episodes of major depression. *Am. J. Psychiatry* 161, 631–636. doi: 10.1176/appi.ajp.161.4.631
- Krajchich, I., Bartling, B., Hare, T., and Fehr, E. (2015). Rethinking fast and slow based on a critique of reaction-time reverse inference. *Nat. Commun.* 6:7455. doi: 10.1038/ncomms8455
- Lempert, K. M., Porcelli, A. J., Delgado, M. R., and Tricomi, E. (2012). Individual differences in delay discounting under acute stress: the role of trait perceived stress. *Front. Psychol.* 3:251. doi: 10.3389/fpsyg.2012.00251
- Love, J., Selker, R., Marsman, M., Jamil, T., Dropmann, D., Verhagen, A. J., et al. (2015). *JASP (Version 0.7.1) [Computer Software]*.
- Morey, R. D., and Rouder, J. N. (2015). *Bayes Factor (Version 0.9.11-3) [Computer Software]*.
- Murphy, R. E. (1959). Effects of threat of shock, distraction, and task design on performance. *J. Exp. Psychol.* 58, 134–141. doi: 10.1037/h0044233
- Odum, A. L. (2011). Delay discounting: trait variable? *Behav. Process.* 87, 1–9. doi: 10.1016/j.beproc.2011.02.007
- Pabst, S., Brand, M., and Wolf, O. T. (2013). Stress effects on framed decisions: there are differences for gains and losses. *Front. Behav. Neurosci.* 7:142. doi: 10.3389/fnbeh.2013.00142
- Porcelli, A. J., and Delgado, M. R. (2009). Acute stress modulates risk taking in financial decision making. *Psychol. Sci.* 20, 278–283. doi: 10.1111/j.1467-9280.2009.02288.x
- Pulcu, E., Trotter, P., Thomas, E., McFarquhar, M., Juhasz, G., Sahakian, B., et al. (2014). Temporal discounting in major depressive disorder. *Psychol. Med.* 44, 1825–1834. doi: 10.1017/S0033291713002584
- Rachlin, H., and Green, L. (1972). Commitment, choice and self-control 1. *J. Exp. Anal. Behav.* 17, 15–22. doi: 10.1901/jeab.1972.17-15
- Ratcliff, R., and McKoon, G. (2008). The diffusion decision model: theory and data for two-choice decision tasks. *Neural Comput.* 20, 873–922. doi: 10.1162/neco.2008.12-06-420

- Robinson, O. J., Bond, R. L., and Roiser, J. P. (2015). The impact of stress on financial decision-making varies as a function of depression and anxiety symptoms. *PeerJ* 3:e770. doi: 10.7717/peerj.770
- Robinson, O. J., Charney, D. R., Overstreet, C., Vytal, K., and Grillon, C. (2012). The adaptive threat bias in anxiety: amygdala–dorsomedial prefrontal cortex coupling and aversive amplification. *Neuroimage* 60, 523–529. doi: 10.1016/j.neuroimage.2011.11.096
- Robinson, O. J., Krimsky, M., and Grillon, C. (2013a). The impact of induced anxiety on response inhibition. *Front. Hum. Neurosci.* 7:69. doi: 10.3389/fnhum.2013.00069
- Robinson, O. J., Overstreet, C., Charney, D. S., Vytal, K., and Grillon, C. (2013b). Stress increases aversive prediction-error signal in the ventral striatum. *Proc. Natl. Acad. Sci. U.S.A.* 110, 4129–4133. doi: 10.1073/pnas.1213923110
- Robinson, O. J., Vytal, K., Cornwell, B. R., and Grillon, C. (2013c). The impact of anxiety upon cognition: perspectives from human threat of shock studies. *Front. Hum. Neurosci.* 7:203. doi: 10.3389/fnhum.2013.00203
- Robinson, O. J., Krimsky, M., Lieberman, L., Allen, P., Vytal, K., and Grillon, C. (2014). The dorsal medial prefrontal (anterior cingulate) cortex–amygdala aversive amplification circuit in unmedicated generalised and social anxiety disorders: an observational study. *The Lancet Psychiatry* 1, 294–302. doi: 10.1016/S2215-0366(14)70305-0
- Rouder, J. N., Morey, R. D., Speckman, P. L., and Province, J. M. (2012). Default Bayes factors for ANOVA designs. *J. Math. Psychol.* 56, 356–374. doi: 10.1016/j.jmp.2012.08.001
- Rounds, J. S., Beck, J. G., and Grant, D. M. (2007). Is the delay discounting paradigm useful in understanding social anxiety? *Behav. Res. Ther.* 45, 729–735. doi: 10.1016/j.brat.2006.06.007
- Schmitz, A., and Grillon, C. (2012). Assessing fear and anxiety in humans using the threat of predictable and unpredictable aversive events (the NPU-threat test). *Nat. Protocols* 7, 527–532. doi: 10.1038/nprot.2012.001
- Spielberger, C. D., Gorsuch, R. L., and Lushene, R. E. (1970). *The State-Trait Anxiety Inventory*. Palo Alto, CA: Consulting Psychologists Press Inc.
- Yi, R., Pitcock, J. A., Landes, R. D., and Bickel, W. K. (2010). The short of it: abbreviating the temporal discounting procedure. *Exp. Clin. Psychopharmacol.* 18, 366–374. doi: 10.1037/a0019904

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2015 Robinson, Bond and Roiser. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Assessment of Tobacco-Related Approach and Attentional Biases in Smokers, Cravers, Ex-Smokers, and Non-Smokers

Marcella L. Woud^{1*}, Joyce Maas², Reinout W. Wiers³, Eni S. Becker² and Mike Rinck^{1,2}

¹ Department of Psychology, Mental Health Research and Treatment Center, Ruhr-Universität Bochum, Bochum, Germany,

² Behavioural Science Institute, Radboud University Nijmegen, Nijmegen, Netherlands, ³ Addiction Development and Psychopathology Lab, University of Amsterdam, Amsterdam, Netherlands

OPEN ACCESS

Edited by:

Frank Ryan,
Imperial College, UK

Reviewed by:

Sally Adams,
University of Bath, UK
Charlotte Elisabeth Wittekind,
University Medical Center
Hamburg-Eppendorf, Germany

*Correspondence:

Marcella L. Woud
marcella.woud@rub.de
This study was conducted while
Marcella L. Woud held a position at
Behavioural Science Institute,
Radboud University Nijmegen,
Montessorilaan 3, 6525 HR Nijmegen,
Netherlands.

Specialty section:

This article was submitted to
Psychology for Clinical Settings,
a section of the journal
Frontiers in Psychology

Received: 07 May 2015

Accepted: 29 January 2016

Published: 26 February 2016

Citation:

Woud ML, Maas J, Wiers RW,
Becker ES and Rinck M (2016)
Assessment of Tobacco-Related
Approach and Attentional Biases in
Smokers, Cravers, Ex-Smokers, and
Non-Smokers. *Front. Psychol.* 7:172.
doi: 10.3389/fpsyg.2016.00172

According to theories of addictive behaviors, approach and attentional biases toward smoking-related cues play a crucial role in tobacco dependence. Several studies have investigated these biases by using various paradigms in different sample types. However, this heterogeneity makes it difficult to compare and evaluate the results. The present study aimed to address this problem, via (i) a structural comparison of different measures of approach-avoidance and a measure of smoking-related attentional biases, and (ii) using within one study different representative samples in the context of tobacco dependence. Three measures of approach-avoidance were employed: an Approach Avoidance Task (AAT), a Stimulus Response Compatibility Task (SRC), and a Single Target Implicit Association Test (ST-IAT). To assess attentional biases, a modified Stroop task including smoking-related words was administered. The study included four groups: $n = 58$ smokers, $n = 57$ non-smokers, $n = 52$ cravers, and $n = 54$ ex-smokers. We expected to find strong tobacco-related approach biases and attentional biases in smokers and cravers. However, the general pattern of results did not confirm these expectations. Approach responses assessed during the AAT and SRC did not differ between groups. Moreover, the Stroop did not show the expected interference effect. For the ST-IAT, cravers had stronger approach associations toward smoking-related cues, whereas non-smokers showed stronger avoidance associations. However, no such differences in approach-avoidance associations were found in smokers and ex-smokers. To conclude, these data do not provide evidence for a strong role of implicit approach and attentional biases toward smoking-related cues in tobacco dependency.

Keywords: tobacco dependence, approach-avoidance, attention, AAT, SRC, STIAT, Stroop

INTRODUCTION

*I need more cigarettes
Give me more cigarettes, I need
Gotta get more cigarettes
I want more cigarettes
Cigarettes
Cigarettes
(More Cigarettes–Replacements–2012)*

The inability to control drug use is a hallmark symptom of a drug addiction (Diagnostic Statistical Manual of Mental Disorders; DSM-5; 2013). Smoking, for example, represents such an addictive behavior, and it is considered one of the most difficult addictions to break. According to the World Health Organization (WHO) report in 2008, tobacco smoking causes 5.4 million deaths per year, and it remains the leading preventable cause of death worldwide (World Health Organization, 2008). Furthermore, research also showed that smoking increases the risk of engaging in other addictive behaviors (Merrill et al., 1999; Creemers et al., 2009). Hence, it is not surprising that there is a growing interest to elucidate the motivational and reward mechanisms underlying this destructive behavior.

According to dual process models of addiction (e.g., Deutsch and Strack, 2006; Wiers et al., 2007; Gladwin et al., 2011; for critical discussion see Gladwin and Figner, 2014), addictive behaviors can be understood best as the output of two distinct types of processes. On the one hand, reflective processes with limited (cognitive) capacity involve processes that are slower, more deliberate and explicit. On the other hand, impulsive processes do not require limited (cognitive) capacity and involve processes that are fast and automatic. It has been suggested that the latter processes are particularly involved in emotional and motivational aspects of behavior. Such dual process models of addiction posit that addictive behaviors are the result of an imbalance between these two processes, i.e., there is no cooperative interplay. There are easily activated, drug-oriented impulsive processes, in combination with relatively slow reflective processes that are not strong enough to control or regulate the impulsive process. Furthermore, and in line with the incentive-sensitization-theory of Robinson and Berridge (1993, 2003, 2008), dual process models of addiction hypothesize that the impulsive processes become sensitized with repeated drug use. Drug-related cues acquire incentive salience, which results in an activation of the mesolimbic dopamine system and an increase in dopamine levels. As a consequence of this neurological chain, the brain “interprets” drug-related cues as rewarding cues, and therefore prepares the corresponding motivational state, i.e., an approach action tendency, aimed at consuming the drug of interest. From an information processing perspective, this explains behavioral phenomena such as attentional and approach biases for drug-related cues: Due to the incentive salience of these cues, they automatically capture an individual’s attention and activate approach-related behaviors.

Over the last decades, there has been a surge of interest in tobacco-related information processing biases (for an overview and meta-analysis, see e.g., Waters and Sayette, 2006; Field and Cox, 2008; Rooke et al., 2008). Such investigations are important from a theoretical but also from a clinical perspective: On the one hand, such studies can test specific hypotheses derived from models of addiction, and on the other hand, these studies can advance our understanding of factors related to the high number of relapse in tobacco dependence. For example, Waters et al. (2003) found that smokers who showed a greater attentional bias for smoking-related words were more likely to lapse in the short-term.

Before summarizing studies investigating tobacco-related approach biases, an important distinction has to be made. This distinction concerns the operationalization of approach biases, namely whether they are operationalized as symbolic or actual motor responses. Regarding the assessment of symbolic motor responses, the Stimulus-Response Compatibility (SRC; Mogg et al., 2003) task has been used to assess symbolic tobacco-related approach biases. During the SRC, participants are instructed to move a manikin figure toward (approach) or away (avoidance) from, for example, smoking-related or neutral pictures. The time needed to initiate the manikin’s approach and avoidance movements serves as the dependent variable. Studies employing the SRC showed that smokers are faster to approach than to avoid smoking-related cues (e.g., Mogg et al., 2003, 2005; Bradley et al., 2004, 2008; Thewissen et al., 2007). Regarding the assessment of actual motor responses, the Approach-Avoidance Task (AAT; Rinck and Becker, 2007) is a suitable paradigm. Indeed, it also has been used to assess tobacco-related approach biases. During the AAT, participants are instructed to pull (approach) and to push (avoidance) a joystick in response to, for example, smoking-related or neutral pictures, that appear on the computer screen. Here, the time needed to execute the push and pull movements serve as the dependent variable. Most AATs apply an indirect task version. That is, the instructions do not ask participants to respond to the pictures’ content. Instead, participants are required to respond to an unrelated feature such as the pictures’ orientation or format. The advantage of such an indirect task version is that participants respond to a stimulus feature that is independent of the stimulus dimension that the task aims to assess, which disguises the research question and makes the use of response strategies less likely (Rinck and Becker, 2007). In the context of tobacco dependence, the AAT is a rather novel paradigm, and to the best of our knowledge, only three studies have employed the AAT so far (but for more AAT studies in the context of alcohol dependency, see e.g., Palfai and Ostafin, 2003; Wiers et al., 2010, 2011; Eberl et al., 2013; Kersbergen et al., 2015). The study by Wiers C. E. et al. (2013), examined tobacco-related approach biases in heavy smokers, non-smokers, and ex-smokers. Results showed that heavy smokers were faster to approach smoking-related pictures compared to non-smokers and ex-smokers. Moreover, this approach bias was correlated with levels of craving. The study by Machulska et al. (2015) compared smokers to non-smokers, and found that smokers, unlike non-smokers, exhibited an approach bias toward smoking-related pictures compared to food-related control pictures (see also Larsen et al., 2014). Finally, according to results of Watson et al. (2013), tobacco-related approach biases can also be conditional. They tested a group of deprived cigarette smokers and found that the bias assessed at baseline was associated with participants’ level of craving. After the baseline assessment, half of the participants were allowed to smoke a cigarette. These participants reported a reduction in craving but an increase in approach bias.

Beyond the studies examining actual and symbolic tobacco-related approach biases, there are also studies targeting tobacco-related approach associations. Word categorization tasks such as the Implicit Association Test (IAT; Greenwald et al., 1998)

have been employed in this type of research. During the IAT, participants simultaneously categorize target stimuli (e.g., smoking-related vs. control stimuli) and attribute words (e.g., approach- or avoidance-related words) as fast as possible into the appropriate superordinate category. The difference in reaction times between the possible combinations (e.g., smoking-related stimuli and approach attributes share the same response key, and neutral stimuli and avoidance-related words share the same response key) is assumed to reflect whether smoking is associated more strongly with either attribute category, with relatively fast responses reflecting relatively strong associations. The study by De Houwer et al. (2006) examined such associations and found that smokers indeed had stronger approach- than avoidance-related tobacco associations, respectively. However, most of the IAT studies compared general positive vs. negative smoking-related associations, and here the evidence is less clear (e.g., Swanson et al., 2001; Sherman et al., 2003; Huijding et al., 2005).

Regarding tobacco-related attentional biases, several studies found that smokers are slower to respond to smoking-related pictures (visual probe task) and words (Stroop task), compared to neutral pictures or words (visual probe task: e.g., Mogg et al., 2003, 2005; Bradley et al., 2008; Stroop task: e.g., Munafò et al., 2003, 2005; Larsen et al., 2014; and for a meta-analysis, see Cox et al., 2006). There is also evidence which further specifies these findings. Results of Mogg and Bradley (2002) showed a positive correlation between smoking-related attentional biases and daily cigarette consumption. Moreover, Wertz and Sayette (2001) found a greater Stroop interference in participants who were told that they were allowed to smoke during the study, compared to those who were told they were not allowed to smoke. Finally, smoking-related attentional biases seem to be related to levels of self-reported craving (Zack et al., 2001; Mogg and Bradley, 2002), and increase after participants have been deprived of smoking (Cox et al., 2006). For example, using a visual probe task, Field et al. (2004) found that deprived smokers maintained their gaze toward smoking-related cues compared to neutral cues.

In summary, there is evidence showing that tobacco dependency is characterized by smoking-related approach and attentional biases. Despite the importance of these findings, however, there are two significant limitations: First, within previous studies, only a limited number of groups have been compared. Second, previous studies employed only a limited number of tasks. This heterogeneity makes it difficult to compare and evaluate these studies, particularly in relation to the underlying theory. The present study aimed to address this problem via (i) a structural comparison of different measures of approach-avoidance and the most commonly used paradigm to assess attentional biases (i.e., the Stroop), and (ii) using within one study different representative samples in the context of tobacco dependence. To assess smoking-related approach avoidance biases, three different measures were used: the Approach Avoidance Task (AAT), the Stimulus Response Compatibility Task (SRC), and a Single Target Implicit Association Test (STIAT; Wigboldus et al., 2004). Given the fact that smoking does not have an inherently meaningful contrast category (such as, for example, alcohol vs. soft drinks), we chose to use a STIAT instead of an IAT. The AAT and SRC

used pictorial stimuli (smoking-related and matched control pictures), the STIAT used word stimuli (targets: smoking-related words, attributes: approach avoidance words). A modified Stroop including smoking-related words was administered to assess tobacco-related attentional biases. Finally, we also assessed explicit attitudes toward smoking and levels of craving over the course of the study. The study included four groups: smokers, cravers, ex-smokers, and non-smokers. Following the predictions of theories of addictive behaviors and the existing empirical evidence in this context, our main hypothesis was to find strong tobacco-related approach and attentional biases in smokers and cravers, compared to ex-smokers and non-smokers. Moreover, we expected that tobacco-related approach and attentional biases would be correlated positively across smokers, cravers, and ex-smokers.

METHODS

Participants

A total of 232 students from Radboud University (NL) were tested ($M_{\text{age}} = 22.36$, $SD = 3.2$, 158 females). Within this group, there were $n = 59$ smokers, $n = 59$ non-smokers, $n = 56$ cravers, and $n = 58$ ex-smokers. The selection criteria were as follows: Smokers were included if they were smoking at least six cigarettes a day for at least 2 months. The same criteria applied for cravers. In order to avoid craving, smokers were instructed to smoke a cigarette prior to the study. However, cravers were instructed to not smoke for 6 h prior to the study. The group of non-smokers included individuals who had never smoked a cigarette or a joint. Ex-smokers were included if they had stopped smoking at least 6 months earlier and had smoked a minimum of six cigarettes a day while actively smoking. Prior to the analyses, 11 participants were excluded: Two non-smokers (one because of technical problems during testing and another who was actually smoking once in a while), four cravers (one did not smoke six or more cigarettes a day, one was tested too early and thus did not crave for 6 h, and two did not comply with the rule to not smoke for 6 h prior to the study), four ex-smokers and one smoker (technical problems during testing), leaving a total sample of $N = 221$ ($n = 58$ smokers, $n = 57$ non-smokers, $n = 52$ cravers, and $n = 54$ ex-smokers)¹.

Materials

Self-Report Measures

Fagerström test for nicotine dependence (FTND)

The FTND (Heatherton et al., 1991) is a self-report measure assessing the degree of nicotine dependence. It contains six items, e.g., “How many cigarettes per day do you smoke?” “How soon after waking do you smoke your first cigarette?” The higher the FTND sum score, the higher participants’ level of dependence.

Explicit attitudes toward smoking

To assess explicit attitudes toward smoking, participants were asked to evaluate eight adjective pairs (e.g., smoking is

¹Please note that there are additional missing data for the self-report measures and reaction time data due to technical problems during testing.

“good–bad,” “sociable–unsociable,” “sexy–unsexy”) on a 7-point scale (see Huijding et al., 2005; Huijding and de Jong, 2006).

Pictorial stimuli

The pictorial stimuli included 20 smoking-related pictures and 20 matched control pictures (for examples, see Supplementary Material). These 40 pictures were divided across two sets, i.e., set A and B, each containing 10 smoking-related pictures and the corresponding 10 matched control pictures.

Approach avoidance task (AAT)

During the AAT (Rinck and Becker, 2007), participants responded to pictures presented on the computer screen by approaching and avoiding them using a joystick. The joystick was positioned in front of the computer screen, tightly fastened to the table. The instructions said that all pictures were tilted either slightly to the left or right, and that the tilt determined whether the pictures had to be pulled (approach movement) or pushed (avoidance movement; for a similar procedure, see Cousijn et al., 2011). Within each of the four participants groups, half of the participants pulled left-tilted and pushed right-tilted pictures, whereas the other half pushed left-tilted and pulled right-tilted pictures. Participants initiated each trial by pressing a button of the joystick with their index finger while holding the joystick in the central position. When the picture appeared, participants had to decide quickly whether the picture was tilted to the left or to the right, and had to respond according to their instructions. During pushing, the pictures became smaller, whereas they became larger during pulling. This zoom supported the approach-avoidance effect visually. Moreover, participants were instructed to “pull the joystick toward themselves,” and to “push it away from them.” Via these instructions, the movements’ reference point was the participant’s body. This disambiguated the movements, and labeled them as clear and unambiguous approach or avoidance movements. After pushing or pulling the joystick all the way into the right direction, participants had to bring it back to the central position and start the next trial. Pictures disappeared only when the joystick was pulled or pushed in the correct direction and when the joystick was pulled or pushed by an angle of 30 degrees.

The AAT started with a practice block during which two practice pictures were pushed and pulled 10 times each. After that, 160 assessment trials followed, including 10 smoking-related pictures and 10 matched control pictures. The assessment was divided into two blocks of 80 trials each. Within each block, the smoking-related pictures and matched control pictures were pushed and pulled four times each [i.e., $(4 \times 10) + (4 \times 10) = 80 \times 2 = 160$ trials in total].

Stimulus response compatibility task (SRC)

In each trial of the SRC task (Mogg et al., 2003), a picture appeared in the center of the screen. In addition, a manikin figure was displayed either below or above the picture. Participants were instructed to move the manikin figure either toward or away from the picture by making use of the keys “2” (manikin moved downwards) and “8” (manikin moved upwards) on the numeric part of the keyboard. There were two blocks with two different stimulus-response assignments: One block

required participants to move the manikin toward smoking-related pictures (approach movement) and to move the manikin away from control pictures (avoidance movement), whereas the other block required participants to move the manikin away from smoking-related pictures and toward control pictures. For the sake of brevity, the following terms will be used to describe these two different stimulus-response assignments: compatible block: manikin approaches smoking-related pictures and avoids control pictures; incompatible block: manikin avoids smoking-related pictures and approaches control pictures. The latency between picture onset and the participant’s response served as the dependent variable. All participants completed both blocks. However, the order of blocks was counterbalanced: Within each of the four participants groups, half of the participants started with the compatible block and then completed the incompatible block, whereas the other half started with the incompatible block and then completed the compatible block. Within each block, the manikin appeared below the picture in 50% of the trials, and above the picture in the other 50%. When the manikin appeared below the picture, 50% of the trials required a down response, whereas the other 50% required an up response, and the same was true when the manikin appeared above the picture. The manikin position and picture type varied randomly over trials.

The SRC started with a practice block during which the manikin approached one picture four times and also avoided one picture four times. After that, 160 assessment trials followed, including 10 smoking-related pictures and 10 matched control pictures. The assessment was divided in two blocks of 80 trials each. Within each block, the smoking-related pictures and matched control pictures were approached and avoided four times each [i.e., $(4 \times 10) + (4 \times 10) = 80 \times 2 = 160$ trials in total].

Single target implicit association test (STIAT)

The STIAT (Wigboldus et al., 2004) consisted of a complete sequence of five blocks: (a) attribute discrimination, (b) practice combined block, (c) first combined block, (d) practice reversed combined block, and (e) reversed combined block. Each block started with instructions describing the discrimination category and the assignment of the response keys (left vs. right). The procedure started with (a) the attribute discrimination block, in which participants had to sort words that belonged to two categories, namely approach or avoidance. Participants were asked to press one key in response to approach-related words, and the other key in response to avoidance-related words (i.e., either key “A” on the very left part of the keyboard or key “6” on the numeric part of the keyboard). The stimuli in this block consisted of six approach-related attribute words and six avoidance-related attribute words. Words were presented one after another in a fixed random order. In the second block, six smoking-related target words were also presented. Participants therefore practiced the combined block (b). There were two different response assignments: one assignment required participants to categorize smoking-related target words with the same key as approach-related attribute words. The other assignment required participants to categorize smoking-related target words with the same key as avoidance-related attribute words. For the sake of brevity, the following terms will be used

to describe these two response assignments: compatible block: smoking-related targets and approach-related attributes shared the same response key; incompatible block: smoking-related targets and avoidance-related attributes shared the same response key. The combined practice block included 24 trials: The six target words were presented once, the six attribute words that required the same response were also shown once, and the six words that required the opposite response key were presented 12 times. Because targets were assigned to only one key during combined blocks, there are fewer responses on the opposite key. Hence, to balance this mismatch of responses by the left and right key, attributes assigned to the opposite side of the targets were presented twice as often, resulting in an equal number of left and right key responses in each of the combined blocks. The key-assignment was counterbalanced within each of the four participant groups: Half of the participants were told to press the left key (“A” key) in response to all targets, and the other half were told to press the right key (“6” key) in response to all targets. Moreover, we controlled for the sequence of the combined blocks: Within each of the four participants groups, half of the participants started with the compatible block and then completed the incompatible block, whereas the other half started with the incompatible block and then completed the compatible block.

After the practice trials, the actual combined block followed (c). This block included 72 trials: The six target words were presented three times (18 trials), the six attribute words which required the same response were also shown three times (18 trials), and the six words which required the opposite response were presented six times each (36 trials). Next, participants practiced the reversal of the response assignment for target words (d). That is, participants who had pressed the approach key in response to smoking-related targets now had to respond with the avoidance key, the other half of the participants vice versa. This combined reversed practice block also consisted of 24 trials: The six target words were presented once, the six attribute words that required the same response were also shown once, and the six words that required the opposite response were presented 12 times each. Finally, the actual reversed combined block followed (e). This block included 72 trials again: The six target words were presented three times, the six attribute words which required the same response were also shown three times, and the six words which required the opposite response were presented six times. During each trial, reminder labels (appropriate category names positioned in the top left and top right corner of the screen) remained visible. Within each block, stimuli appeared in the same fixed random order for each participant. After incorrect responses, a red “X” appeared in the center of the screen. Given the high numbers of German students at Radboud University, we had two STIATs; a Dutch and a German version.

Emotional stroop

During this task, participants categorized word stimuli according to their print color. The stimuli were presented on cards. There were five print colors: white, blue, red, green, and yellow. There were three types of cards. All participants started with the practice card. Here, meaningless colored strings of “XXX” were presented. After that, the smoke card or the neutral card was presented

(randomized). On the smoke card, eight smoking-related words were shown (e.g., cigarette, smoke, cigar). These words differed from the smoking-related words used during the STIAT. On the neutral card, eight household-related words were shown (e.g., towel, broom, spoon). The order of the smoke and the neutral card was random. All cards contained 40 stimuli each, i.e., eight stimuli distributed across five columns. Each card appeared on the screen after a mouse click initiated by the experimenter. As soon as the participant had named the last word’s print color, the experimenter clicked again and the card disappeared. Reaction times were saved on the computer, for each card separately, and these reaction times were used in the analyses. Participants’ errors were recorded by the experimenter, who was blind to the type of card that was presented. Given the high numbers of German students at Radboud University, we used a German and a Dutch Stroop version.

Procedure

Participants were tested individually in separate testing cubicles. After having signed informed consent, participants’ level of carbon monoxide (CO) was assessed by means of the piCO+ smokelyzer (Bedfont Scientific, Kent, England). For smokers, CO levels were assessed 10 min after they had smoked their cigarette. Next, smokers, cravers and ex-smokers answered a question about their level of craving (“How strong is your urge to smoke a cigarette right now?”) using a scale from 0 (= no urge) to 100 (= strong urge). Moreover, smokers and cravers had to indicate how many cigarettes they would smoke on a normal day. Ex-smokers were asked to indicate this for the time they were still smoking. Then, the four computer tasks followed. There were two orders and this was counterbalanced: Within each of the four participants groups, half of the participants received order one (STIAT, SRC, Stroop, AAT), the other half order two (AAT, Stroop, SRC, STIAT). The tasks’ order was linked to the picture set (A or B, example: if a participant started with the AAT, picture set A was used for the AAT, and picture set B was used for the SRC, and vice versa if a participant started with the SRC). That is, for task order one, the AAT always included picture set A and the SRC included picture set B. For task order two, the AAT always included picture set B and the SRC included picture set A. After the computer tasks, smokers, cravers and ex-smokers completed a second craving question and the FTND. Ex-smokers received an adapted version of the FTND that was related to their past smoking behavior. The smoking attitude rating was then completed by all participants. Finally, cravers were asked to smoke a cigarette and after 10 min, their CO levels were assessed a second time. This second assessment served as an extra check for the cravers’ temporal abstinence, i.e., we expected their CO value to be higher than their CO value assessed before the start of the study. The present study had the necessary ethical approvals via the Behavioural Science Institute.

RESULTS

Participant Characteristics

Table 1 gives an overview of the samples’ characteristics and the means and standard deviations of the following measures: average of daily smoked cigarettes, levels of carbon monoxide

(CO) pre study, craving pre and post-study, and scores on the Fagerström Test for Nicotine Dependence (FTND). A chi-square test revealed that the four groups did not differ concerning gender, $\chi^2_{(3)} = 4.84$, $p = 0.18$. Univariate ANOVAs were conducted to examine the following baseline measures (please note that not all groups were involved in all comparisons): Age, $F_{(3, 217)} = 7.81$, $p < 0.001$, $\eta^2 = 0.1$; Average of daily smoked cigarettes, $F_{(2, 161)} = 0.51$, $p = 0.6$; CO levels pre study, $F_{(3, 202)} = 94.91$, $p < 0.001$, $\eta^2 = 0.59$; Craving pre study, $F_{(2, 160)} = 91.2$, $p < 0.001$, $\eta^2 = 0.53$; FTND scores, $F_{(2, 161)} = 7.84$, $p < 0.01$, $\eta^2 = 0.09$.

These outcomes were treated as follows: For age, we repeated the main analyses (i.e., for the AAT, SRC, STIAT, and Stroop) including age as a covariate. This did not change the results, and thus for clarity and given the lack of specific hypotheses regarding age, we report unadjusted analyses without this covariate, and did not analyse this baseline imbalance further. Hence, for the sake of brevity, we report all analyses without this factor. Moreover, we did not further analyze FTND scores, given the fact that the ex-smokers' score is a retrospectively assessed score and thus not an optimal measure. However, we did further examine the findings concerning the CO levels and craving scores pre study. Regarding the pre study CO levels, Bonferroni *post-hoc* tests including all four groups (i.e., smokers, non-smokers, cravers, ex-smokers) revealed that all group comparisons were significant (p 's < 0.002), except for the non-smokers vs. ex-smokers comparison ($p = 1$). Regarding the craving scores pre study, Bonferroni *post-hoc* tests including smokers, cravers, and ex-smokers revealed significant differences for all three comparisons, p 's < 0.03 .

Craving Over the Course of the Study

We also assessed participants' level of craving over the course of the study. A repeated-measures ANOVA including the between-subjects factor Group (smokers, cravers, ex-smokers) and the within-subjects factor Time (craving pre, craving post) revealed a significant main effect of Time, $F_{(1, 158)} = 23.88$, $p < 0.001$, $\eta^2 = 0.13$, and Group, $F_{(2, 158)} = 99.91$, $p < 0.001$, $\eta^2 = 0.56$. Moreover, there was a marginally significant Time \times Group interaction; $F_{(2, 158)} = 2.79$, $p = 0.065$, $\eta^2 = 0.03$. This interaction was further examined by three paired-samples t -tests, i.e., one for each group comparing craving scores pre vs. post: smokers: $t_{(56)} = 3.86$, $p < 0.001$; cravers: $t_{(51)} = 2.76$, $p < 0.01$; ex-smokers: $t_{(51)} = 1.92$, $p = 0.06$. Following this, smokers' and cravers' level of craving significantly increased over the course of the study. In the group of ex-smokers, this increase was marginally significant, although it would no longer be after application of a Bonferroni correction (for means and standard deviations, see Table 1).

Analyses Approach-Avoidance Biases

For the analyses of the AAT, SRC, and STIAT, the effects of potential outliers were corrected by computing the median reaction time (RT) of each participant. Thus, the means reported below are means of medians.

Approach Avoidance Task (AAT)

The analysis included only trials during which a participant pushed or pulled the joystick all the way into the right direction within one movement. As a first step, we examined the groups' error trials by means of a univariate ANOVA. Results showed that the groups did not differ here: $F_{(3, 217)} = 1.64$, $p = 0.18$ (smokers: $M = 0.05$, $SD = 0.04$; non-smokers: $M = 0.05$, $SD = 0.05$; cravers: $M = 0.07$, $SD = 0.09$; ex-smokers: $M = 0.05$, $SD = 0.06$).

Next, a difference score per participant was calculated. As a first step, RTs of pull movements were subtracted from RTs of push movements, for both picture types (i.e., smoking and control). As such, a positive difference score reflects an approach bias. After that, we subtracted the control pictures' difference score from that of smoking-related pictures. Here, a positive difference score indicates a stronger approach bias toward smoking-related pictures. Finally, participants with an error percentage greater than 20% were excluded from the analysis (non-smokers $n = 2$, cravers $n = 3$, ex-smokers $n = 2$).

To analyze the AAT data, a univariate ANOVA was conducted with Group (smokers, non-smokers, cravers, ex-smokers) and Order Tasks (one, two) as between-subjects factor, and the overall difference score as dependent variable. Of most interest was the main effect of Group. However, this effect did not reach significance, $F_{(3, 206)} = 0.9$, $p = 0.44$. As such, the groups did not differ in their approach-avoidance responses toward smoking-related and control pictures (for an overview of means, standard deviations and n 's per group, see Table 2)².

Stimulus Response Compatibility (SRC) task

As a first step, we examined the groups' error scores by means of a Univariate ANOVA. Results showed that the groups did not differ here: $F_{(3, 216)} = 0.41$, $p = 0.75$ (smokers: $M = 0.06$, $SD = 0.04$; non-smokers: $M = 0.06$, $SD = 0.04$; cravers: $M = 0.06$, $SD = 0.04$; ex-smokers: $M = 0.07$, $SD = 0.05$). Based on this error check, we excluded two participants from the analysis because their error percentage was greater than 20% (smokers: $n = 1$, ex-smokers: $n = 1$). To analyze the RT data, we subtracted RTs of the compatible block from RTs of the incompatible block. As such, a positive difference score indicates faster approach of smoking-related pictures. Next, we conducted a univariate ANOVA with

²For the sake of clarity, we only report the outcome of main interest in the text. Hence, please find the additional outcomes here (i.e., main effects and interactions) of the analyses of the AAT, SRC, STIAT and Stroop data: AAT: Order Tasks, $F_{(1, 206)} = 0.25$, $p = 0.62$, Group \times Order Tasks, $F_{(3, 206)} = 0.89$, $p = 0.45$. SRC: Order Tasks, $F_{(1, 202)} = 0.49$, $p = 0.49$, Order SRC, $F_{(1, 202)} = 0.41$, $p = 0.52$, Group \times Order Tasks, $F_{(3, 202)} = 1.59$, $p = 0.19$, Group \times SRC order, $F_{(3, 202)} = 2.88$, $p = 0.04$, $\eta^2 = 0.04$, Order Tasks \times SRC Order, $F_{(1, 202)} = 0.02$, $p = 0.9$, Group \times Order Tasks \times SRC Order, $F_{(3, 202)} = 0.59$, $p = 0.63$. STIAT: Order Tasks, $F_{(1, 202)} = 4.1$, $p = 0.04$, $\eta^2 = 0.02$, STIAT Order, $F_{(1, 202)} = 1.1$, $p = 0.3$, Group \times Order Tasks, $F_{(3, 202)} = 0.64$, $p = 0.59$, Group \times STIAT Order, $F_{(3, 202)} = 4.43$, $p < 0.01$, $\eta^2 = 0.06$, Order Tasks \times STIAT Order, $F_{(1, 202)} = 18.62$, $p < 0.001$, $\eta^2 = 0.08$, Group \times Order Tasks \times STIAT Order, $F_{(3, 202)} = 4.13$, $p < 0.01$, $\eta^2 = 0.06$. Stroop: Order Tasks, $F_{(1, 192)} = 0.56$, $p = 0.45$, Order Stroop, $F_{(1, 192)} = 14.32$, $p < 0.001$, $\eta^2 = 0.07$, Group \times Order Tasks, $F_{(3, 192)} = 0.53$, $p = 0.67$, Group \times Order Stroop, $F_{(3, 192)} = 3.87$, $p = 0.01$, $\eta^2 = 0.06$, Order Tasks \times Order Stroop, $F_{(1, 192)} = 0.48$, $p = 0.49$, Group \times Order Tasks \times Order Stroop, $F_{(3, 192)} = 2.51$, $p = 0.6$, $\eta^2 = 0.04$.

TABLE 1 | Descriptives of the four groups.

	Gender: f/m	Age	Average daily smoking	CO levels pre	CO levels post	Craving pre study	Craving post-study	FTND
Group		M (SD, n)	M (SD, n)	M (SD, n)	M (SD, n)	M (SD, n)	M (SD, n)	M (SD, n)
Smokers	39/19	22.28 (2.55, 58)	12.59 (4.83, 58)	17.02 (8.65, 58)		20.37 (23.97, 57)	31.75 (25.5, 57)	5.26 (1.63, 58)
Cravers	31/21	22.88 (3.72, 52)	12.94 (6.2, 52)	7.02 (5.3, 52)	11.16 (6.01, 49)	58.98 (20.83, 52)	65.63 (22.28, 52)	5.38 (1.33, 52)
Ex-smokers	37/17	23.57 (3.59, 54)	13.70 (6.69, 54)	2.41 (1.12, 39)		10.1 (11.09, 52)	13.08 (17.41, 52)	4.43 (1.06, 54)
Non-smokers	45/12	20.88 (2.23, 57)		1.77 (0.82, 57)				

CO levels pre: Levels of carbon monoxide (CO) assessed before the study (please note Footnote 1); CO levels post: Levels of carbon monoxide (CO) assessed again in cravers after the study; Average daily smoking: Average of daily smoked cigarettes for smokers, cravers and ex-smokers (retrospective); Craving pre study: Levels of cigarette craving in smokers, cravers and ex-smokers before the study; Craving post-study: Levels of cigarette craving in smokers, cravers and ex-smokers after the study; FTND: Mean sum score of the Fagerström Test for Nicotine Dependence (FTND) in smokers, cravers, and ex-smokers (retrospective).

TABLE 2 | Means and standard deviations for all four computer tasks per group.

Task	Group	Score 1 M (SD)	Score 2 M (SD)	Overall difference score M (SD)	n
AAT	Smokers	−17 (78)	−14 (65)	−3 (76)	58
	Cravers	−0.7 (54)	−19 (48)	19 (61)	49
	Ex-smokers	−13 (69)	−17 (60)	4 (69)	52
	Non-smokers	−14 (76)	−18 (60)	4 (83)	55
SRC	Smokers	760 (146)	843 (169)	83 (114)	56
	Cravers	796 (119)	901 (221)	106 (147)	52
	Ex-smokers	838 (179)	887 (222)	49 (143)	53
	Non-smokers	770 (121)	823 (156)	52 (104)	57
STIAT	Smokers	588 (89)	594 (74)	6 (92)	57
	Cravers	594 (87)	641 (100)	47 (86)	52
	Ex-smokers	614 (85)	623 (79)	9 (103)	53
	Non-smokers	594 (78)	577 (64)	−17 (60)	56
Stroop	Smokers	27898 (4829)	26411 (4434)	1487 (2634)	57
	Cravers	31644 (7757)	29116 (4398)	2528 (6965)	46
	Ex-smokers	28224 (4858)	26437 (4853)	1787 (2634)	50
	Non-smokers	31601 (12206)	29537 (9761)	2064 (15419)	55

AAT, Approach Avoidance Task; Score 1, RTs push movements—RTs pull movements smoke pictures; Score 2, RTs push movements—RTs pull movements control pictures; Overall difference score, Score 1—Score 2, i.e., a positive difference score indicates a stronger approach bias toward smoking-related pictures. SRC: Score 1, Compatible block (manikin approaches smoking-related pictures and avoid control pictures); Score 2, Incompatible block (manikin avoids smoking-related pictures and approaches control pictures); Overall difference score, Incompatible block—compatible block, i.e., a positive difference score indicates faster approach of smoking-related pictures. STIAT: Score 1, Compatible block (smoking-related targets and approach-related attributes shared the same response key); Score 2, Incompatible block (smoking-related targets and avoidance-related attributes shared the same response key); Overall difference score, Incompatible block—compatible block, i.e., a positive difference score indicates faster approach-related associations toward smoking-related pictures. Stroop: Score 1, RTs smoke card; Score 2, RTs neutral card; Overall difference score, Score 1—Score 2, i.e., a positive difference score indicates a greater interference for smoking-related stimuli.

Group (smokers, non-smokers, cravers, ex-smokers), Order SRC (compatible-incompatible, incompatible-compatible) and Order Tasks (one, two) as between-subjects factor, and the difference score as dependent variable. Of most interest here was the main effect of Group. Results showed that this effect was marginally significant, $F_{(3, 202)} = 2.24$, $p = 0.085$, $\eta^2 = 0.03$. However, *post-hoc* Bonferroni tests revealed that none of the between-group comparisons were significant (p 's > 0.05). As such, the groups did not differ in their approach-avoidance responses toward smoking-related and control pictures (for an overview of means, standard deviations and n 's per group, see **Table 2**).

Single Target Implicit Association Test (STIAT)

As a first step, we examined the groups' error score by means of a Univariate ANOVA. Results showed that the groups did not differ here: $F_{(3, 216)} = 1.31$, $p = 0.27$ (smokers: $M = 0.04$, $SD = 0.03$; non-smokers: $M = 0.05$, $SD = 0.04$; cravers: $M = 0.04$, $SD = 0.04$; ex-smokers: $M = 0.04$, $SD = 0.04$). Based on this error check, we excluded two participants from the analysis because their error percentage was greater than 20% (non-smokers: $n = 1$, ex-smokers: $n = 1$). To analyze the RT data, we subtracted RTs of the compatible block from RTs of the incompatible block. As such, a positive difference score

indicates faster approach-related associations toward smoking-related pictures. Next, we conducted a univariate ANOVA with Group (smokers, non-smokers, cravers, ex-smokers), Order STIAT (compatible-incompatible, incompatible-compatible) and Order Tasks (one, two) as between-subjects factor, and the difference score as dependent variable. Of most interest here was the main effect of Group. Results showed that this effect was significant, $F_{(3, 202)} = 5.74$, $p < 0.01$, $\eta^2 = 0.08$. *Post-hoc* Bonferroni tests revealed the following: cravers vs. non-smokers: $p < 0.001$, cravers vs. smokers: $p = 0.051$, cravers vs. ex-smokers: $p = 0.094$. None of the other comparison reached significance ($p's > 0.05$). Following this, cravers had a stronger approach-bias toward smoking-related cues than smokers, non-smokers and ex-smokers (for an overview of means, standard deviations and n 's per group, see **Table 2**).

Attentional Bias Emotional Stroop

Prior to the analysis, a difference score was calculated per participant. Here, we only used RTs of the neutral card and the smoke card, not the practice card including the meaningless colored "XXX" strings. More precisely, RTs of the neutral card were subtracted from RTs of the smoke card. As such, a positive difference score indicates greater interference for smoking-related stimuli. We excluded 6 participants because their difference score deviated more than 3 SD from their group's mean difference score (non-smokers $n = 2$, cravers $n = 4$). We conducted a univariate ANOVA including the between-subjects factor Group (smokers, non-smokers, cravers, ex-smokers), Order Stroop Card (smoke-neutral, neutral-smoke) and Order Tasks (one, two), and the difference score as dependent variable. Of main interest here was the main effect of Group. However, results showed that this effect was not significant, $F_{(3, 192)} = 0.4$, $p = 0.75$. As such, there were no group differences concerning the interference of smoking-related vs. neutral stimuli (for an overview of means, standard deviations and n 's per group, see **Table 2**).

Ratings Explicit Attitudes Toward Smoking

Prior to analysis, scores on the eight adjective pairs were collapsed into one overall score. Here, higher sum scores signal a more negative attitude toward smoking. To investigate whether the groups differed concerning their explicit attitude toward smoking, a univariate ANOVA was conducted with Group (smokers, non-smokers, cravers, ex-smokers) as between-subjects factor and the collapsed attitude sum score as dependent variable. Results showed a significant main effect of Group, $F_{(3, 217)} = 52.32$, $p < 0.001$, $\eta^2 = 0.42$. *Post-hoc* Bonferroni tests revealed significant results for all group comparisons ($p's < 0.01$), except for the smoker vs. craver comparison, $p = 0.1$ (smokers: $M = 33.66$, $SD = 4.47$; non-smokers: $M = 45.7$, $SD = 6.65$; cravers: $M = 33.82$, $SD = 5.68$; ex-smokers: $M = 38.11$, $SD = 6.41$). Moreover, one sample t -tests showed that all four group means deviated significant from zero, $p's < 0.001$. This result pattern generally shows that non-smokers have the most negative attitude toward smoking, followed by ex-smokers, smokers, and cravers.

Correlations

Across the group of smokers, cravers, and ex-smokers, correlations were calculated for the following measures: AAT, SRC, STIAT, Stroop, explicit attitudes toward smoking, daily smoking, FTND scores, urge pre study, urge post-study. **Table 3** gives an overview of these findings. We particularly expected to find positive correlations between tobacco-related approach and attentional biases. However, there were only two marginally significant correlations, i.e., between the AAT and the STIAT ($r = 0.15$), showing that the stronger the approach bias toward smoking-related pictures on the AAT, the stronger the approach-associations toward smoking-related words on the STIAT; and between the SRC and the Stroop ($r = 0.15$), showing that the stronger the approach bias toward smoking-related pictures on the SRC, the greater the smoking-related attentional bias on the Stroop. When looking at the correlations including explicit attitudes toward smoking, daily smoking, FTND scores, urge pre study, urge post-study, we found the following: Both the AAT and the SRC correlated significantly with explicit smoking attitudes ($r = -0.15$, marginally significant, and $r = -0.18$), showing that the stronger the approach bias toward smoking-related pictures, the less negative participants' attitude toward smoking was. Moreover, there was a marginally significant correlation between the Stroop and FTND scores ($r = 0.16$), indicating that the greater the smoking-related attentional bias on the Stroop, the higher participants' levels of nicotine dependence. Urge assessed before and after the computer tasks correlated with the STIAT (urge pre: $r = 0.27$, urge post: $r = 0.25$), and the Stroop (urge post: $r = 0.17$), showing that the higher levels of urge, the stronger the approach-associations toward smoking-related words on the STIAT, and the stronger the smoking-related attentional bias on the Stroop. Please note, however, that most of the correlations would not remain significant after controlling for multiple comparisons.

DISCUSSION

The present study examined the role of approach and attentional biases in tobacco dependence. We tested four groups, namely smokers, cravers, ex-smokers, and non-smokers. The following tasks were employed: To assess approach-related biases, we used the Approach Avoidance Task (AAT), the Stimulus Response Compatibility Task (SRC), and a Single Target Implicit Association Test (STIAT). A modified Stroop including smoking-related words was administered to assess attentional biases. Moreover, we assessed explicit attitudes toward smoking and levels of craving over the course of the study. We expected to find strong tobacco-related approach and attentional biases in smokers and cravers compared to ex-smokers and non-smokers. However, the general pattern of results did not confirm these expectations. Approach responses assessed during the AAT and SRC did not differ between groups. Moreover, the Stroop did not show the expected interference effect. Regarding the data of the STIAT, results were partly in line with our expectations: Cravers showed stronger approach associations toward smoking-related cues, whereas non-smokers showed stronger avoidance associations. However, no such differences

TABLE 3 | Correlations between AAT, SRC, STIAT, Stroop, explicit attitudes, daily smoking, FTND, urge pre-study, and urge post-study in smokers, cravers, and ex-smokers (N = 226).

	1	2	3	4	5	6	7	8
1. AAT	—							
2. SRC	−0.05	—						
3. STIAT	0.15 [#]	0.07	—					
4. Stroop	−0.04	0.15 [#]	−0.02	—				
5. Explicit attitudes	−0.15 [#]	−0.18*	−0.13	−0.07	—			
6. Daily smoking	0.06	−0.02	0.05	−0.05	0.14 [#]	—		
7. FTND	−0.1	0.00	0.02	0.16 [#]	−0.00	0.02	—	
8. Urge pre-study	0.07	0.09	0.27***	0.13	−0.19*	0.14 [#]	0.24**	—
9. Urge post-study	0.12	0.1	0.25**	0.17*	−0.29***	0.07	0.28**	0.82**

AAT, Approach-Avoidance Task (difference score smoke-related pictures—difference score control pictures); SRC, Stimulus Response Compatibility (incompatible—compatible); STIAT, Single Target Implicit Association Test (incompatible—compatible); Daily smoking, cigarettes per day; FTND, Fagerström Test for Nicotine Dependence.

[#] $p < 0.100$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

in approach-avoidance associations were found in smokers and ex-smokers. Generally, correlational analyses did not reveal the expected positive correlations between tobacco-related approach and attentional biases among smokers, cravers, and ex-smokers. However, we did find some patterns that are in line with the theory, e.g., the stronger the approach bias toward smoking-related pictures, the stronger the smoking-related approach associations. Regarding the assessment of participants' explicit smoking-related attitudes, results were generally indicative of a negative attitude toward smoking, with non-smokers and ex-smokers having the most negative attitudes toward smoking. Finally, results showed that smokers' and cravers' level of craving significantly increased over the course of the study. In the group of ex-smokers, this increase was marginally significant.

To summarize, our data do not provide strong evidence for the role of approach and attentional biases in tobacco dependency, except for findings on the STIAT. Given the large sample size of each group, a lack of statistical power does not seem to be a likely explanation. Hence, a closer inspection of the tested groups, the tasks and their stimuli could help to understand these null-findings. Regarding the groups, their average smoking behavior is the first index to check and compare. Across our groups, smokers, cravers, and ex-smokers smoked 12–13 cigarettes a day. These scores are rather low when comparing them, for example, with the samples tested by Wiers C. E. et al. (2013): In that study, smokers and ex-smokers reported an average between 22 and 24 cigarettes a day. Thus, one could argue that our groups were not “smoking enough” in order to show tobacco-related approach and attentional biases. However, other studies found such biases in samples that exhibited a similar smoking behavior than ours (e.g., Munafò et al., 2003; Bradley et al., 2004; Mogg et al., 2005). Moreover, in our study, smokers and cravers were supposed to be active smokers for at least two years. Another index, i.e., the groups' score on the Fagerström Test of Nicotine Dependence (FTND), is also rather inconclusive. Our groups scored around five on the FTND, which does not deviate much from the values in other studies (e.g., Munafò et al., 2003; Bradley et al., 2008; Wiers C. E. et al., 2013). To conclude, the sample's general smoking-related characteristics match with

other studies, and thus do not provide a sufficient explanation for the null-findings. Regarding the tasks we employed, we used well-established tasks in the context of tobacco-related approach and attentional biases (i.e., AAT, SRC; STIAT, and emotional Stroop). The AAT is a rather novel task for this specific type of addictive behavior. However, it has been proven successful in the assessment of alcohol-related approach biases (e.g., Palfai and Ostafin, 2003; Wiers et al., 2010, 2011; Eberl et al., 2013; Kersbergen et al., 2015). Therefore, given the fact we tapped into similar processes (i.e., approach biases), in combination with the successful results reported by the three previous studies (Wiers C. E. et al., 2013; Larsen et al., 2014; Machulska et al., 2015), the AAT seemed a promising instrument. Only the STIAT provided results that partly supported our predictions. That is, we found stronger approach associations toward smoking-related cues in cravers, whereas stronger avoidance associations were found in non-smokers. From a theoretical perspective, this is in line with assumptions put forward by dual process models of addiction (e.g., Deutsch and Strack, 2006; Wiers et al., 2007) and the incentive-sensitization model (Robinson and Berridge, 1993, 2003, 2008): For cravers who were deprived of smoking, smoking-related cues had a high incentive salience, which in turn automatically elicited an approach association. For non-smokers, in contrast, for whom smoking-related cues did not have incentive salience and were rather associated with negativity and unpleasantness, smoking-related cues automatically elicited an avoidance association. Interestingly, however, our STIAT version slightly deviated from that of other STIATs as it included a high number of trials. To summarize, the details of the specific tasks used cannot explain the present null-results. Finally, the choice of stimulus material needs to be analyzed. The AAT and SRC included pictures that depicted clear smoking-related scenes or attributes, and the corresponding matched control picture. A problem with such matched control pictures could be that they were in fact “too good.” That is, given their high similarity with the smoking-related pictures, they were possibly not distinctive enough. This, in combination with participants' instruction to react to the pictures as quickly as possible, could be partly responsible for not finding any differences in approach-related

response within the tested groups. Moreover, some of the pictures contained food-related objects, so the control pictures could have elicited approach tendencies too. In this context, Machulska et al. (2015) suggest that it might be beneficial to use pictures that depict the commencement of smoking behavior (following findings by Stippekohl et al., 2012, and for similar reason when using pleasant vs. unpleasant smoking-related pictures, see Bradley et al., 2008). Our picture set included only three of such pictures, which could partly explain the null-findings of the AAT and SRC. Finally, there are three additional limitations that could partly explain the present findings. A first limitation is the low reliability of some of the tasks we applied. Second, we did not use baseline CO levels as an inclusion criterion. Third, we cannot rule out that the smokers who were asked to smoke a cigarette prior to testing experienced a smoking-related priming effect during testing. Especially the latter two issues could have affected the results in an unfortunate manner.

To conclude, although the study has some limitations as highlighted above, the present findings remain rather puzzling. Our results neither replicate earlier findings, nor support predictions of dual process models of addiction (e.g., Deutsch and Strack, 2006; Wiers et al., 2007) or the incentive-sensation model (Robinson and Berridge, 1993, 2003, 2008). Following this, our findings do not provide support for studies aiming to re-train approach and attentional biases, a development which has revealed promising findings in the area of alcohol addiction. Here, results showed that computerized trainings, i.e., procedures derived from “Cognitive Bias Modification” techniques (cf. Koster et al., 2009; Woud and Becker, 2014), are able to reduce alcohol-related approach biases (e.g., via Alcohol-AAT-Training, AAATT). Most important, however, results showed that such trainings improve treatment outcomes even at one-year follow-up (Wiers et al., 2010, 2011; Eberl et al., 2013; Gladwin et al., 2014; and for an overview of CBM-related results in addiction, see Wiers R. W. et al., 2013). In fact, CBM training could be

also quite useful in the context of tobacco-related biases, as they operate comparably to those reported in the alcohol literature. Indeed, one published study applied a computerized re-training in the context of tobacco dependence (Wittekind et al., 2014). This study found that tobacco avoidance training reduced levels of cigarette consumption and dependence. However, the study is only a pilot study without a control group. Hence, these data should be interpreted with caution.

Following our null-findings, we suggest that future research should address a number of issues. To start with, studies should further examine the exact conditions of tobacco-related approach and avoidance biases. Moreover, this has to be examined among various relevant groups, e.g., smokers, cravers, ex-smokers, and non-smokers, while taking craving into account (Watson et al., 2013). Although the empirical evidence is rather supportive of the existence of tobacco-related biases, a few studies also found results that do not support a strong role of such biases in smoking. To illustrate, Munafò et al. (2003) and Mogg and Bradley (2002) did not find differences in information processing biases between abstinent smokers and non-abstinent smokers, whereas Larsen et al. (2014) did not find differences in biases between smokers and non-smokers. Hence, it might be possible that there are a number of subtle, boundary conditions, which are not yet fully understood.

Taken together, we did not find the expected tobacco-related approach and attentional biases, and therefore encourage future research to advance our understanding of the nature of these phenomena.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <http://journal.frontiersin.org/article/10.3389/fpsyg.2016.00172>

REFERENCES

- Bradley, B. P., Field, M., Mogg, K., and De Houwer, J. (2004). Attentional and evaluative biases for smoking cues in nicotine dependence: component processes of biases in visual orienting. *Behav. Pharmacol.* 15, 29–36. doi: 10.1097/00008877-200402000-00004
- Bradley, B. P., Field, M., Healy, H., and Mogg, K. (2008). Do the affective properties of smoking-related cues influence attentional and approach biases in cigarette smokers? *J. Psychopharmacol.* 22, 737–745. doi: 10.1177/0269881107083844
- Cousijn, J., Goudriaan, A. E., and Wiers, R. W. (2011). Reaching out towards cannabis: approach-bias in heavy cannabis users predicts changes in cannabis use. *Addiction* 106, 1667–1674. doi: 10.1111/j.1360-0443.2011.03475.x
- Cox, W. M., Fadardi, J. S., and Pothos, E. M. (2006). The addiction Stroop test: theoretical considerations and procedural recommendations. *Psychol. Bull.* 132, 443–476. doi: 10.1037/0033-2909.132.3.443
- Creemers, H. E., Korhonen, T., Kaprio, J., Vollebbergh, W. A., Ormel, J., Verhulst, F. C., et al. (2009). The role of temperament in the relationship between early onset of tobacco and cannabis use. The TRAIL Study. *Drug Alcohol Depend.* 104, 113–118. doi: 10.1016/j.drugalcdep.2009.04.010
- De Houwer, J., Custers, C., and de Clercq, A. (2006). Do smokers have a negative implicit attitude towards smoking? *Cogn. Emot.* 20, 1274–1284. doi: 10.1080/02699930500484506
- Deutsch, R., and Strack, F. (2006). Duality-models in social psychology: from opposing processes to interacting systems. *Psychol. Inq.* 17, 166–172. doi: 10.1207/s15327965pli1703_2
- Eberl, C., Wiers, R. W., Pawelzack, S., Rinck, M., Becker, E. S., and Lindenmeyer, J. (2013). Approach bias modification in alcohol dependence: do clinical effects replicate and for whom does it work best? *Dev. Cogn. Neurosci.* 4, 38–51. doi: 10.1016/j.dcn.2012.11.002
- Field, M., and Cox, W. M. (2008). Attentional bias in addictive behaviors: a review of its development, causes, and consequences. *Drug Alcohol Depend.* 97, 1–20. doi: 10.1016/j.drugalcdep.2008.03.030
- Field, M., Mogg, K., and Bradley, B. P. (2004). Eye movements to smoking-related cues: effects of nicotine deprivation. *Psychopharmacology* 173, 116–123. doi: 10.1007/s00213-003-1689-2
- Greenwald, A. G., McGhee, D. E., and Schwartz, J. K. L. (1998). Measuring individual differences in implicit cognition: the Implicit Association Test. *J. Pers. Soc. Psychol.* 74, 1464–1480. doi: 10.1037/0022-3514.74.6.1464
- Gladwin, T. E., Figner, B., Crone, E. A., and Wiers, R. W. (2011). Addiction, adolescence, and the integration of control and motivation. *Dev. Cogn. Neurosci.* 1, 364–376. doi: 10.1016/j.dcn.2011.06.008
- Gladwin, T. E., and Figner, B. (2014). “Hot” cognition and dual systems: introduction, criticisms, and ways forward,” in *Frontiers of Cognitive Psychology Series: Neuroeconomics, Judgment, and Decision Making*, eds E. Wilhelms and V. F. Reyna (New York, NY: Psychological Press), 157–180.

- Gladwin, T. E., Rinck, M., Eberl, C., Becker, E. S., Lindenmeyer, J., and Wiers, R. W. (2014). Mediation of cognitive bias modification for alcohol addiction via stimulus-specific alcohol avoidance association. *Alcoholism* 39, 101–107. doi: 10.1111/acer.12602
- Heatherington, T. F., Kozlowski, L. T., Frecker, R. C., and Fagerström, K. O. (1991). The fagerström test for nicotine dependence: a revision of the fagerström tolerance questionnaire. *Br. J. Addict.* 86, 1119–1127. doi: 10.1111/j.1360-0443.1991.tb01879.x
- Huijding, J., and de Jong, P. J. (2006). Automatic associations with the sensory aspects of smoking: positive in habitual smokers but negative in non-smokers. *Addict. Behav.* 31, 182–186. doi: 10.1016/j.addbeh.2005.04.014
- Huijding, J., de Jong, P. J., Wiers, R. W., and Verkooijen, K. (2005). Implicit and explicit attitudes toward smoking in a smoking and a non-smoking setting. *Addict. Behav.* 30, 949–961. doi: 10.1016/j.addbeh.2004.09.014
- Kersbergen, I., Woud, M. L., and Field, M. (2015). The validity of different measures of automatic alcohol action tendencies. *Psychol. Addict. Behav.* 29, 225–230. doi: 10.1037/adb0000009
- Koster, E. H. W., MacLeod, C., and Fox, E. (2009). Introduction to the special section on cognitive bias modification in emotional disorders. *J. Abnorm. Psychol.* 118, 1–4. doi: 10.1037/a0014379
- Larsen, H., Kong, G., Becker, D., Cousijn, J., Boendermaker, W., Cavallo, D., et al. (2014). Implicit motivational processes underlying smoking in American and Dutch adolescents. *Front. Psychiatry* 5:51. doi: 10.3389/fpsy.2014.00051
- Machulska, A., Zlomuzica, A., Adolph, A., Rinck, M., and Margraf, J. (2015). “A cigarette a day keeps the goodies away”: smokers show automatic approach tendencies for smoking—but not for food-related stimuli. *PLoS ONE* 10:e0116464. doi: 10.1371/journal.pone.0116464
- Merrill, J. C., Kleber, H. D., Shwartz, M., Liu, H., and Lewis, S. R. (1999). Cigarettes, alcohol, marijuana, other risk behaviors, and American youth. *Drug Alcohol Depend.* 56, 205–212. doi: 10.1016/S0376-8716(99)00034-4
- Mogg, K., and Bradley, B. P. (2002). Selective processing of smoking-related cues in smokers: manipulation of deprivation level and comparison of three measures of processing bias. *J. Psychopharmacol.* 16, 385–392. doi: 10.1177/026988110201600416
- Mogg, K., Bradley, B. P., Field, M., and De Houwer, J. (2003). Biases in eye movements to smoking-related cues in smokers: relationship to affective and motivational valence. *Addiction* 98, 825–836. doi: 10.1046/j.1360-0443.2003.00392.x
- Mogg, K., Field, M., and Bradley, B. P. (2005). Attentional and approach biases for smoking cues in smokers. An investigation of competing theoretical views of addiction. *Psychopharmacology* 180, 333–341. doi: 10.1007/s00213-005-2158-x
- Munafò, M. R., Johnstone, E. C., and Mackintosh, B. (2005). Association of serotonin transporter genotype with selective processing of smoking-related stimuli in current smokers and ex-smokers. *Nicot. Tob. Res.* 7, 773–778. doi: 10.1080/14622200500259861
- Munafò, M. R., Mogg, K., Roberts, S., Bradley, B., and Murphy, M. F. G. (2003). Selective processing of smoking-related cues in current smokers, ex-smokers and never-smokers on the modified Stroop task. *J. Psychopharmacol.* 17, 311–317. doi: 10.1177/02698811030173013
- Palfai, T. P., and Ostafin, B. D. (2003). Alcohol-related motivational tendencies in hazardous drinkers: assessing implicit response tendencies using the modified IAT. *Behav. Res. Ther.* 41, 1149–1162. doi: 10.1016/S0005-7967(03)00018-4
- Rinck, M., and Becker, E. S. (2007). Approach and avoidance in fear of spiders. *Behav. Res. Ther.* 38, 105–120. doi: 10.1016/j.jbtep.2006.10.001
- Robinson, T. E., and Berridge, K. C. (1993). The neural basis of drug craving: an incentive-sensitization theory of addiction. *Brain Res. Rev.* 18, 247–291. doi: 10.1016/0165-0173(93)90013-P
- Robinson, T. E., and Berridge, K. C. (2003). Addiction. *Annu. Rev. Psychol.* 54, 25–53. doi: 10.1146/annurev.psych.54.101601.145237
- Robinson, T. E., and Berridge, K. C. (2008). The incentive sensitization theory of addiction: some current issues. *Philos. Trans. R. Soc. B* 363, 3137–3146. doi: 10.1098/rstb.2008.0093
- Rooke, S. E., Hine, D. W., and Thorsteinsson, E. B. (2008). Implicit cognition and substance Use: a meta-analysis. *Addict. Behav.* 33, 1314–1328. doi: 10.1016/j.addbeh.2008.06.009
- Sherman, S. J., Presson, C. C., Chasin, L., Rose, J. S., and Koch, K. (2003). Implicit and explicit attitudes towards cigarette smoking: the effects of context and motivation. *J. Soc. Clin. Psychol.* 22, 13–39. doi: 10.1521/jscp.22.1.13.22766
- Stippeckohl, B., Walter, B., Winkler, M. H., Mucha, R. F., Pauli, P., Vaitl, D., et al. (2012). An early attentional bias to BEGIN-stimuli of the smoking ritual is accompanied with mesocorticolimbic deactivations in smokers. *Psychopharmacology* 222, 593–607. doi: 10.1007/s00213-012-2670-8
- Swanson, J. E., Rudman, L. A., and Greenwald, A. G. (2001). Using the Implicit Association Test to investigate attitude-behavior consistency for stigmatized behavior. *Cogn. Emot.* 15, 207–230. doi: 10.1080/02699930125706
- Thewissen, R., Havermans, R. C., Geschwind, N., van den Hout, M., and Jansen, A. (2007). Pavlovian conditioning of an approach bias in lowdependent smokers. *Psychopharmacology* 194, 33–39. doi: 10.1007/s00213-007-0819-7
- Waters, A. J., and Sayette, M. A. (2006). “Implicit cognition and tobacco addiction,” in *Handbook of Implicit Cognition and Addiction*, eds R. W. Wiers and A. W. Stacy (London: Sage), 309–338.
- Waters, A. J., Shiffman, S., Sayette, M. A., Paty, J. A., Gwaltney, C. J., and Balabanis, M. H. (2003). Attentional bias predicts outcome in smoking cessation. *Health Psychol.* 22, 378–387. doi: 10.1037/0278-6133.22.4.378
- Watson, P., de Wit, S., Cousijn, J., Hommel, B., and Wiers, R. W. (2013). Motivational mechanisms underlying the approach bias to cigarettes. *J. Exp. Psychopathol.* 4, 250–262. doi: 10.5127/jep.030512
- Wertz, J. M., and Sayette, M. A. (2001). Effects of smoking opportunity on attentional bias in smokers. *Psychol. Addict. Behav.* 15, 268–271. doi: 10.1037/0893-164X.15.3.268
- Wiers, C. E., Kühn, S., Javadi, A. H., Korucuoglu, O., Wiers, R. W., Walter, H., et al. (2013). Automatic approach bias toward smoking cues is present in smokers but not in ex-smokers. *Psychopharmacology* 229, 187–197. doi: 10.1007/s00213-013-3098-5
- Wiers, R. W., Bartholow, B. D., van den Wildenberg, E., Thush, C., Engels, R. C. M. E., Sher, K. J., et al. (2007). Automatic and controlled processes and the development of addictive behaviors in adolescents: a review and a model. *Pharmacol. Biochem. Behav.* 86, 263–283. doi: 10.1016/j.pbb.2006.09.021
- Wiers, R. W., Eberl, C., Rinck, M., Becker, E. S., and Lindenmeyer, J. (2011). Retraining Automatic action tendencies changes alcoholic patients’ approach bias for alcohol and improves treatment outcome. *Psychol. Sci.* 22, 490–497. doi: 10.1177/0956797611400615
- Wiers, R. W., Gladwin, T. E., Hofmann, W., Salemink, E., and Ridderinkhof, K. R. (2013). Cognitive bias modification and control training in addiction and related psychopathology: mechanisms, clinical perspectives and ways forward. *Clin. Psychol. Sci.* 1, 192–212. doi: 10.1177/2167702612466547
- Wiers, R. W., Rinck, M., Kordts, R., Houben, K., and Strack, F. (2010). Re-training automatic action-tendencies to approach alcohol in hazardous drinkers. *Addiction* 105, 279–287. doi: 10.1111/j.1360-0443.2009.02775.x
- Wigboldus, D. H. J., Holland, R. W., and van Knippenberg, A. (2004). *Single Target Implicit Associations*. Unpublished manuscript.
- Wittekind, C. E., Feist, A., Schneider, B. C., Moritz, S., and Fritsche, A. (2014). The Approach-Avoidance Task as a self-help intervention in cigarette smoking: a pilot study. *J. Behav. Ther. Exp. Psychiatry* 46, 115–120. doi: 10.1016/j.jbtep.2014.08.006
- World Health Organization (2008). *WHO Report on the Global Tobacco Epidemic 2008*. Available online at: http://www.who.int/tobacco/mpower/mpower_report_full_2008.pdf
- Woud, M. L., and Becker, E. S. (2014). Editorial for the special issue on cognitive bias modification techniques: an introduction to a time traveller’s tale. *Cogn. Ther. Res.* 38, 83–88. doi: 10.1007/s10608-014-9605-0
- Zack, M., Belsito, L., Scher, R., Eissenberg, T., and Corrigan, W. A. (2001). Effects of abstinence and smoking on information processing in adolescent smokers. *Psychopharmacology* 153, 249–257. doi: 10.1007/s002130000552

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2016 Woud, Maas, Wiers, Becker and Rinck. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Priming of conflicting motivational orientations in heavy drinkers: robust effects on self-report but not implicit measures

Lisa C. G. Di Lemma^{1,2}, Joanne M. Dickson¹, Pawel Jedras¹, Anne Roefs³ and Matt Field^{1,2*}

¹ Department of Psychological Sciences, University of Liverpool, Liverpool, UK, ² UK Centre for Tobacco and Alcohol Studies, Liverpool, UK, ³ Clinical Psychological Science, Maastricht University, Maastricht, Netherlands

OPEN ACCESS

Edited by:

Frank Ryan,
Imperial College, UK

Reviewed by:

Thomas Edward Gladwin,
Ministry of Defense, Netherlands
Kim Mitchell Caudwell,
Curtin University, Australia

*Correspondence:

Matt Field,
Department of Psychological
Sciences, University of Liverpool,
Bedford Street South,
Liverpool L69 7ZA, UK
mfield@liv.ac.uk

Specialty section:

This article was submitted to
Psychology for Clinical Settings,
a section of the journal
Frontiers in Psychology

Received: 19 June 2015

Accepted: 14 September 2015

Published: 02 October 2015

Citation:

Di Lemma LCG, Dickson JM,
Jedras P, Roefs A and Field M (2015)
Priming of conflicting motivational
orientations in heavy drinkers: robust
effects on self-report but not implicit
measures. *Front. Psychol.* 6:1465.
doi: 10.3389/fpsyg.2015.01465

We report results from three experimental studies that investigated the independence of approach and avoidance motivational orientations for alcohol, both of which operate within controlled and automatic cognitive processes. In order to prime their approach or avoidance motivational orientations, participants watched brief videos, the content of which (positive or negative depictions of alcohol, or neutral) varied by experimental group. Immediately after watching the videos, participants completed self-report (Approach and Avoidance of Alcohol Questionnaire; all studies) and implicit (visual probe task in study 1, stimulus-response compatibility task in studies 2 and 3) measures of alcohol-related approach and avoidance. In study 3, we incorporated an additional experimental manipulation of thought suppression in an attempt to maximize the influence of the videos on implicit measures. Findings were consistent across all three studies: increases in self-reported approach inclinations were mirrored by decreases in avoidance inclinations, and vice versa. However, a combined analysis of data from all studies demonstrated that changes in approach inclinations were partially independent of changes in avoidance inclinations. There were no effects on implicit alcohol-related processing biases, although methodological issues may partially account for these findings. Our findings demonstrate that subjective approach and avoidance inclinations for alcohol tend to fluctuate in parallel, but changes in approach inclinations may be partially independent from changes in avoidance inclinations. We discuss methodological issues that may partially account for our findings.

Keywords: alcohol, ambivalence, approach, automatic, avoidance, implicit, thought suppression

Introduction

According to the ambivalence model of craving (Breiner et al., 1999; McEvoy et al., 2004), the decision to consume alcohol is determined by the balance between motivational inclinations to indulge ("approach") and to abstain ("avoidance"). Approach and avoidance inclinations might arise from the desire for intoxication or the wish to keep a clear head for the next day, respectively. Motivational conflict (or ambivalence), which plays an important role in alcohol use disorders and their treatment (Hettema et al., 2005), arises when a person has the motivation to drink

and to abstain at the same time. Importantly, these motivational orientations can operate in both controlled (or explicit) and automatic (or implicit) cognitive processes. Controlled processes are rule-based and reflective, they operate within conscious awareness and they can be assessed with self-report measures. Automatic processes are activated spontaneously and they are typically assessed with indirect tasks such as computerized measures of attentional bias and automatic approach tendencies (Stacy and Wiers, 2010). One theoretical model proposed that subjective craving (a controlled process) and attentional bias (an automatic process) have reciprocal causal influences on each other (Field and Cox, 2008), although an alternative account is that automatic and controlled processes are both outputs of underlying processes that cannot be measured directly (motivational orientations; see Christiansen et al., 2015). In this paper we report results from three studies in which we experimentally manipulated motivational orientations for alcohol in order to thoroughly investigate the independence of approach and avoidance in both controlled and automatic processing.

Regarding controlled processes, the Approach and Avoidance of Alcohol Questionnaire (AAAQ; McEvoy et al., 2004) was developed to capture the strength of approach and avoidance inclinations for alcohol. The initial factor analysis of non-dependent drinkers' responses on the AAAQ yielded three subscales, two representing approach inclinations (inclined-indulgent and obsessed-compelled subscales, corresponding to mild and strong inclinations, respectively), and one representing inclinations to avoid drinking alcohol (the resolved-regulated subscale). Subsequent studies employed the AAAQ with different populations of drinkers and performed factor analysis on participants' responses. Each of these studies confirmed that approach and avoidance represent distinct underlying factors, although some studies with alcohol dependent patients (Klein et al., 2007; Schlauch et al., 2013c; but see Klein and Anker, 2013) identified a single underlying factor to approach inclinations rather than the qualitative distinction between mild and strong approach that was reported in the initial study (McEvoy et al., 2004). Many of these studies demonstrated that both approach and avoidance inclinations are independently associated with drinking-related variables. For example, approach and avoidance inclinations account for unique variance in quantity and frequency indices of alcohol consumption in both non-dependent (McEvoy et al., 2004) and alcohol dependent (Klein et al., 2007) drinkers. Approach and avoidance inclinations also have differential predictive validity in alcohol-dependent patients: following treatment, relapse to drinking is predicted by the strength of approach inclinations, but avoidance inclinations are not predictive (Schlauch et al., 2012; Klein and Anker, 2013; see also Schlauch et al., 2013c). On the other hand, avoidance inclinations (but not approach inclinations) predict the likelihood of entering into and engaging with treatment (Schlauch et al., 2012). Taken together, these findings provide support for the ambivalence model of craving (Breiner et al., 1999) because they demonstrate that self-reported approach and avoidance inclinations for alcohol are separable constructs that are uniquely associated with past and future

drinking behavior (see also Curtin et al., 2005; Schlauch et al., 2013a,b).

Regarding automatic processes, there is evidence for co-existence of appetitive (approach) and aversive (avoidance) alcohol-related processing biases in problem drinkers in a variety of sub-domains, including affective associations (Dickson et al., 2013), attentional bias (Stormark et al., 1997), and approach and avoidance tendencies (Barkby et al., 2012). Regarding attentional bias, heavy drinkers who are not seeking treatment have an attentional bias for alcohol cues (Townshend and Duka, 2001; Field et al., 2004). The strength of this attentional bias is reliably associated with the strength of subjective craving (Field et al., 2009) and is potentiated by experimental manipulations that increase the motivation to drink, such as induction of negative mood and exposure to alcohol-related cues (see Field and Cox, 2008). By contrast, alcohol-dependent patients who are tested in treatment contexts show initial attentional bias that is quickly followed by attentional avoidance (Stormark et al., 1997; Noël et al., 2006; Townshend and Duka, 2007; Vollstädt-Klein et al., 2009; Field et al., 2013). The latter pattern of attentional bias may reflect ambivalence, with appetitive motivational processes mapped to the initial attentional bias and aversive motivational processes mapped on to the subsequent attentional avoidance (see Field et al., 2013, for discussion). Consistent with this interpretation, a recent eye tracking study demonstrated that heavy drinkers who were identified as ambivalent (as assessed with the AAAQ) had an approach-avoidance pattern of attentional bias for alcohol cues (i.e., the initial attentional bias quickly followed by attentional avoidance that is characteristic of alcohol-dependent patients), whereas heavy drinkers who were not ambivalent maintained their attentional bias for alcohol cues (Lee et al., 2014).

Automatic approach and avoidance tendencies evoked by alcohol-related cues have been assessed with the alcohol-related stimulus-response compatibility (SRC) task (Field et al., 2008) and related tasks (Wiers et al., 2009). These tasks reveal that in heavy drinkers who are not seeking treatment, alcohol cues evoke automatic approach tendencies (Field et al., 2008, 2011; Wiers et al., 2009; Christiansen et al., 2012; Sharbanee et al., 2013a,b; Kersbergen et al., 2015), and in some studies the strength of these approach tendencies was associated with the strength of subjective craving (Field et al., 2005, 2008). A different pattern is seen in alcohol-dependent patients: one study reported no reliable tendency to approach or avoidance (Barkby et al., 2012) whereas another study found an automatic avoidance tendency, the strength of which was predictive of subsequent relapse (Spruyt et al., 2013). One explanation for these findings is that the standard version of the SRC task yields an index of automatic approach that is *relative* to avoidance. This means that the pattern that is observed in heavy drinkers who are not seeking treatment (Field et al., 2008, 2011; Christiansen et al., 2012; Kersbergen et al., 2015) could be attributed to strong automatic approach, weak automatic avoidance, or a combination of the two. Among alcohol-dependent patients, if alcohol cues simultaneously evoke strong automatic approach at the same time as strong automatic avoidance, this may explain why this population display either no overall bias (Barkby et al., 2012) or an avoidance bias (Spruyt

et al., 2013) depending on the strength of their motivational orientations to avoid alcohol at the time of testing.

Findings from the cross-sectional and prospective studies described above are consistent with the ambivalence model (Breiner et al., 1999) because they suggest that approach and avoidance motivational orientations for alcohol may exist independently of each other, rather than lying at opposite ends of a single continuum. More compelling evidence for the independence of approach and avoidance can be derived from experimental studies that attempt to influence one motivational orientation (approach or avoidance) in order to investigate if the opposing motivational orientation is (un)affected. Regarding automatic processes, we recently demonstrated that subliminal priming of approach or avoidance motivational orientations for alcohol had no effect on attentional biases or automatic approach or avoidance tendencies, although methodological issues complicated interpretation of those findings (Baker et al., 2014). Regarding controlled processes, several studies investigated the effects of exposure to alcohol cues on self-reported approach and avoidance inclinations for alcohol, and all reported findings that were suggestive of partially independent approach and avoidance responses to those cues (Curtin et al., 2005; Jones et al., 2013; Schlauch et al., 2013a,b). For example, in one study exposure to alcohol cues (pouring, holding, and sniffing a beer) led to increases in approach inclinations (AAAQ inclined-indulgent and obsessed-compelled subscales), but avoidance inclinations (AAAQ resolved-regulated subscale) were unaffected (Jones et al., 2013).

Although these studies are informative, a more rigorous experimental test of the independence of approach and avoidance would be to contrast the effects of experimental manipulations that are intended to increase approach or avoidance motivational orientations for alcohol. To achieve this, we were inspired by methods used in a previous study (Roefs et al., 2006) in which participants' automatic processing of food-related words was assessed in contexts that were intended to activate either approach (focusing on the preparation of a tasty meal) or avoidance (focusing on the importance of a healthy diet, and therefore avoiding unhealthy foods). In the present studies, participants viewed short videos that depicted either the positive or negative aspects of alcohol consumption, which should in principle activate approach or avoidance, respectively. Control groups of participants viewed videos that were unrelated to alcohol consumption. Immediately after watching the videos, participants completed the AAAQ (all studies) followed by computerized measures of attentional bias (study 1) and automatic approach and avoidance tendencies (studies 2 and 3). In addition, in study 3 we investigated if thought suppression (see Moss et al., 2015) would moderate the influence of videos on implicit measures.

Our general hypotheses were that the video depicting the positive consequences of alcohol consumption would increase self-reported approach (inclined-indulgent and obsessed-compelled subscales of the AAAQ) and indices of automatic approach (attentional bias in study 1, automatic approach tendencies in studies 2 and 3), but would not influence self-reported and automatic avoidance, as assessed by the

resolved-regulated subscale of the AAAQ and attentional avoidance (study 1) and automatic avoidance tendencies (studies 2 and 3), respectively. By contrast, the video depicting the negative consequences of alcohol consumption would increase self-reported and automatic indices of avoidance, but indices of approach would be unaffected.

Study 1

The alcohol-related visual probe task (see Field et al., 2004) is a computerized measure of attentional bias that can distinguish between attentional bias toward and attentional bias away from alcohol-related pictorial stimuli (hereafter referred to as *attentional avoidance*). In each trial of the task, an alcohol-related picture and a matched neutral picture are briefly presented on opposite sides of a computer screen before a visual probe replaces one of the pictures. Participants' manual reaction times to probes are used to infer biases in the allocation of visuospatial attention. An attentional bias for alcohol cues is inferred if the participant is faster to react to probes that replace alcohol pictures (congruent trials), rather than probes that replace neutral pictures (incongruent trials). If, however, this pattern is reversed (i.e., if the participant is faster to respond on incongruent trials), this is interpreted as attentional avoidance of alcohol cues. Biases in automatic attentional capture or delayed disengagement of attention can be inferred by comparing reaction times on these trials with those on other trials in which only neutral pictures are presented (Koster et al., 2004; see Baker et al., 2014). Although the literature on group differences is inconsistent (see Field and Cox, 2008), several studies demonstrated that heavy drinkers who are not seeking treatment have an attentional bias for alcohol cues when those cues are presented for 500 ms or longer (Townshend and Duka, 2001; Field et al., 2004; Baker et al., 2014), and this has been corroborated by studies of eye movements toward those cues (Lee et al., 2014). Conversely, alcohol-dependent patients who are tested in treatment settings show initial attentional bias for briefly-presented alcohol cues (50–100 ms), that is followed by attentional avoidance when those cues are presented for longer periods (upwards of 500 ms; Stormark et al., 1997; Noël et al., 2006; Townshend and Duka, 2007; Vollstädt-Klein et al., 2009; Field et al., 2013).

In the present study, participants watched a brief video that depicted either the positive consequences of alcohol consumption (alcohol-positive group), the negative consequences of alcohol consumption (alcohol-negative group), or that had no alcohol-related content (control group). Immediately after watching the video, participants completed the AAAQ and an alcohol-related visual probe task in which picture pairs were presented for 50 or 500 ms. We hypothesized that, relative to the control group, participants in the alcohol-positive group would have elevated scores on the inclined-indulgent and obsessed-compelled subscales of the AAAQ, and elevated attentional bias for alcohol cues presented for both 50 and 500 ms; however, scores on the resolved-regulated subscale of the AAAQ would not differ between alcohol-positive and control groups. By contrast, compared to the control group, participants in the alcohol-negative group would have elevated scores on the

resolved-regulated subscale of the AAAQ and would exhibit an “approach-avoidance” pattern of attentional bias on the visual probe task, with bias toward alcohol cues presented for 50 ms followed by attentional avoidance of those cues presented for 500 ms; however, scores on the inclined-indulgent and obsessed-compelled subscales of the AAAQ would not differ between the alcohol-negative and control groups.

Methods

Participants

Ninety participants (69 Female, mean age 21.70, $SD = 5.04$) were recruited from the students and staff at the University of Liverpool via online and poster advertising. Inclusion criteria included fluency in English, age between 18 and 45, normal or corrected-to-normal vision and self-reported alcohol consumption in excess of the current UK government guidelines for safe drinking (these are 14 units per week for females and 21 units per week for males, where 1 unit equals 8 g of alcohol). Exclusion criteria included any history of alcohol use disorders. Participants who had taken part in studies 2 or 3 were ineligible to participate. All participants provided informed consent before taking part in the study, which was approved by the University of Liverpool Research Ethics Committee.

Materials

Self-report measures

Timeline followback drinking diary (Sobell and Sobell, 1992)

Participants indicated their alcohol consumption over the previous 2 weeks. From this, we were able to calculate the total amount of alcohol consumed in standard UK units.

Alcohol use disorders identification test (AUDIT; Saunders et al., 1993)

This 10-item self-report questionnaire contains questions about frequency and quantity of alcohol consumption, and alcohol-related problems and harms. It yields a total score ranging between 0 and 40, with scores of 8 or above indicative of hazardous drinking.

Approach and avoidance of alcohol questionnaire, right now version (AAAQ; McEvoy et al., 2004)

This 14-item questionnaire assesses subjective tendencies to approach or avoid drinking at that moment in time. Respondents are asked to rate how strongly they agree with each item on a 9-point Likert scale, from 0 (not at all) to 8 (very strong). There are three underlying sub-scales: “Inclined-Indulgent” (mild approach, akin to desire to drink) “Obsessed-Compelled” (strong approach, akin to obsessive thoughts about drinking); and “Resolved-Regulated” (motivation to avoid drinking).

Positive and negative affect schedule (PANAS; Watson et al., 1988)

The PANAS is a 20-item Likert scale that yields scores on positive affect (PA) and negative affect (NA). Results are reported in the Supplementary Materials.

Video questionnaire

This eight-item questionnaire was developed to measure participants’ perception of and engagement with the videos. Participants responded to each item on a 5-point Likert scale, with labels ranging from “strongly disagree” to “strongly agree.” Items are shown in Tables S2A–C.

Visual probe task (for similar tasks see Field et al., 2004; Koster et al., 2004)

This task was programmed in Psychopy v.1.74 (Peirce, 2007) and was administered on a desktop computer with a 15-inch monitor. On each trial, a small white fixation cross was presented in the center of the screen for 500 ms. Immediately after offset, a pair of pictures (each 65 mm high \times 80 mm wide) was presented on the left and right of the screen, 130 mm apart, for either 50 or 500 ms. Immediately after the screen was cleared, the visual probe (a small white arrow that pointed up or down) was presented on either the left or right side of the screen, in the position that had been occupied by one of the pictures. The probe remained on the screen until participants made a response by pressing a key labeled “up” or “down” on the computer keyboard.

Participants were instructed to rest the index fingers of their left and right hands on the “up” and “down” keys, to fixate on the fixation cross at the beginning of each trial, and to rapidly categorize the visual probe as soon as it appeared. The latency and accuracy of responses were recorded. There was an initial practice block of 10 trials in which four pairs of affectively neutral pictures, taken from the International Affective Picture System (IAPS; Lang et al., 2008) were presented. The main block of trials then followed, and this comprised two different types of trials: alcohol-neutral trials, and neutral-neutral trials. For alcohol-neutral trials, a set of seven alcohol-related pictures were each paired with neutral pictures that depicted items of stationery. We used a subset of picture pairs that had been used in earlier studies (Field et al., 2011; Barkby et al., 2012) and the pictures in each pair were matched on perceptual characteristics including brightness and complexity. On neutral-neutral trials, we used four pairs of affectively neutral pictures from the IAPS, as described above. During the main block of trials, there were 112 alcohol-neutral trials and 64 neutral-neutral trials. Each picture pair was presented 16 times, and picture location (left or right), stimulus onset asynchrony (SOA; 50 or 500 ms), probe position (left or right) and probe type (up or down arrow) were counterbalanced for all picture pairs. Trials were presented in a random order.

Video stimuli

We created three different videos in order to manipulate participants’ inclinations to drink alcohol or to refrain from drinking. Videos were created in Windows Movie Maker (version 2.6) and were presented in Windows Media Player (version 7) player in full-screen mode on the computer. Participants wore headphones while watching the videos, all of which were 3 min and 45 s in duration. All video files are available from the Corresponding Author on request.

The *alcohol-positive* video was intended to evoke motivational inclinations to approach alcohol. It comprised still images

depicting people having fun while drinking alcohol, together with some text slides that provided information about the positive consequences of drinking and was accompanied by an upbeat soundtrack. The *alcohol-negative* video was intended to evoke motivational inclinations to avoid alcohol. It comprised still images depicting the negative consequences of drinking, including scenes of alcohol-related violence and vomiting, and other slides depicting graphic government advertisements that warned of the consequences of drink-driving and alcohol-related organ damage and was accompanied by a downbeat soundtrack. The *neutral* video comprised still photos of office equipment and furniture, and was accompanied by non-descript jazz music. All images were obtained using a Google Images search.

Procedure

Participants were randomly allocated to experimental condition. They were tested in a laboratory in the Department of Psychological Sciences at the University of Liverpool. After providing informed consent participants completed the timeline follow back drinking diary, AUDIT, AAAQ, and PANAS (time 1). Then, participants put on the headphones and watched one of the videos (depending on experimental condition), before completing the Video Questionnaire and the AAAQ and PANAS again (time 2). Finally, participants completed the visual probe task. After completing the study, participants were debriefed and offered either course credit or a £5 Shopping Voucher to compensate them for their time.

Results

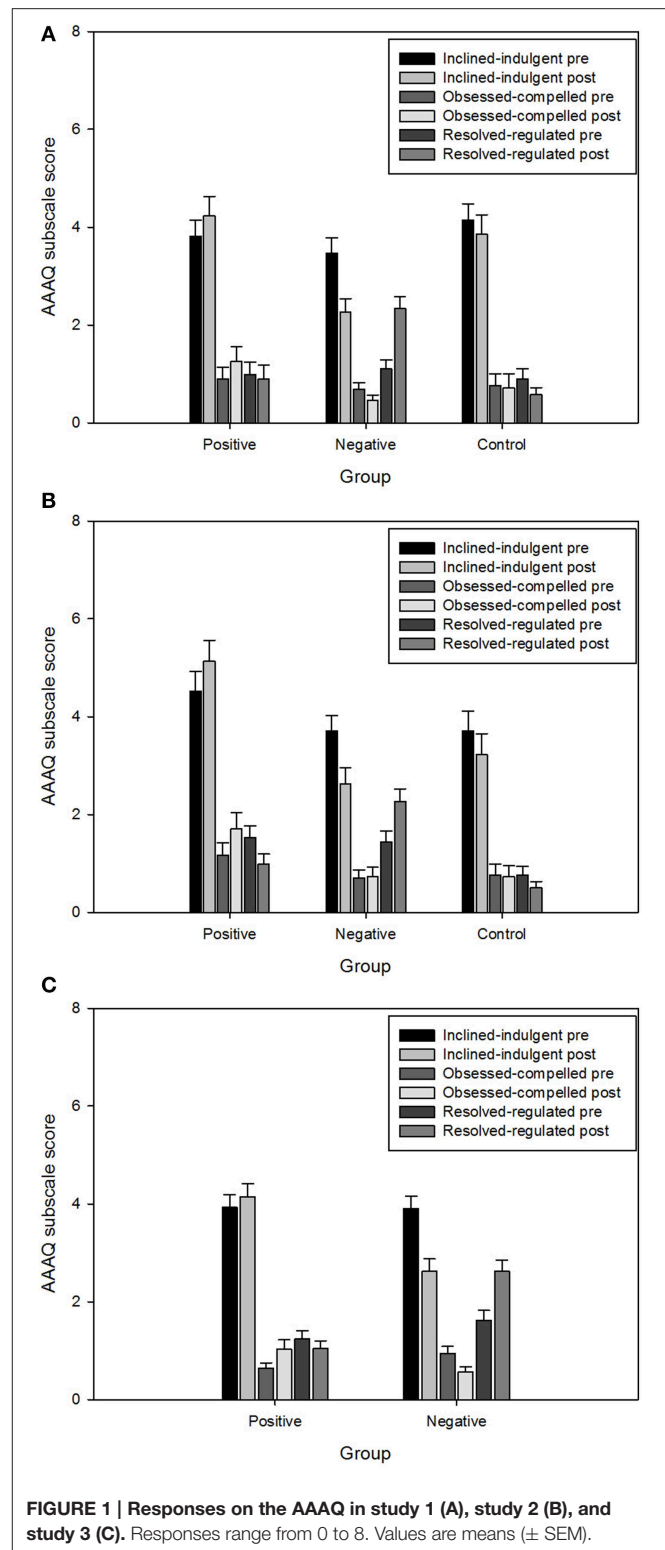
Group Characteristics

Participants reported consuming 20.55 ($SD = 11.53$) units of alcohol per week, and the mean score on the AUDIT was 12.18 ($SD = 5.28$). There were no between-group differences in weekly alcohol consumption or AUDIT scores (Kruskal–Wallis tests $ps > 0.09$), although there was a trend for participants in the alcohol-positive group to be older than participants in the other two groups (Kruskal–Wallis $p = 0.05$). There were no group differences in gender ratio ($\chi^2 = 0.45, p > 0.1$).

Effects of Video Manipulation on AAAQ Ratings (Figure 1A)

AAAQ ratings were analyzed using a mixed design ANOVA, with within-subject factors of sub-scale (3: inclined-indulgent, obsessed-compelled, resolved-regulated), time (2: before video, after video), and group (3: alcohol-positive, alcohol-negative, control). The sub-scale \times time \times group interaction was statistically significant [$F_{(4, 174)} = 27.05, p < 0.001$]. Subsequent *post-hoc* ANOVAs confirmed that the time \times group interaction was significant for all three sub-scales [inclined-indulgent $F_{(2, 87)} = 25.29, p < 0.001$; obsessed-compelled $F_{(2, 87)} = 5.72, p < 0.01$; resolved-regulated $F_{(2, 87)} = 28.32, p < 0.001$].

There were no group differences on any of the AAAQ sub-scales before participants watched the video [inclined-indulgent $F_{(2, 89)} = 1.12, p > 0.1$; obsessed-compelled $F_{(2, 89)} = 0.26, p > 0.1$; resolved-regulated $F_{(2, 89)} = 0.22, p > 0.1$]. As predicted, groups differed on all three sub-scales after watching the video, [inclined-indulgent $F_{(2, 89)} = 9.13, p < 0.001$; resolved-regulated



$F_{(2, 89)} = 17.57, p < 0.01$], although this fell short of significance for the obsessed-compelled sub-scale [$F_{(2, 89)} = 2.93, p = 0.059$]. *Post-hoc* LSD contrasts confirmed that scores on both the inclined-indulgent and obsessed-compelled sub-scales were

higher in the alcohol-positive group compared to the alcohol-negative group ($p < 0.01$), but this pattern was reversed for the resolved-regulated subscale ($p < 0.01$). The direct test of our hypotheses requires contrasts between these groups and the control group. These contrasts revealed that alcohol-positive and control groups did not differ on any subscale ($p > 0.1$). However, scores on the resolved-regulated subscale were higher, and scores on the inclined-indulgent subscale lower, in the alcohol-negative compared to the control group ($p < 0.01$). Alcohol-negative and control groups did not differ on the obsessed-compelled subscale ($p > 0.1$).

Paired-samples *t*-tests revealed that among participants in the alcohol-positive group, scores on the inclined-indulgent and obsessed-compelled sub-scales increased after watching the video [$t_{(28)} = 2.92, p < 0.01$ and $t_{(28)} = 2.29, p < 0.05$], whereas scores on the resolved-regulated sub-scale did not change [$t_{(28)} = 0.61, p > 0.1$]. A different pattern was seen in the alcohol-negative group: inclined-indulgent and obsessed-compelled scores decreased [$t_{(31)} = 6.41, p < 0.001$ and $t_{(31)} = 2.24, p < 0.05$], whereas scores on the resolved-regulated sub-scale increased [$t_{(31)} = 6.34, p < 0.001$]. In the control group, scores on both the inclined-indulgent and resolved-regulated sub-scales decreased after watching the video, although the former failed to reach significance [$t_{(28)} = 1.95, p = 0.06$ and $t_{(28)} = 2.66, p < 0.05$]; scores on the obsessed-compelled sub-scale did not change [$t_{(29)} = 0.30, p > 0.1$].

We also re-ran the omnibus Three-Way ANOVA on AAAQ scores but added PANAS positive and PANAS negative affect after the video as covariates. The three way interaction sub-scale \times time \times group remained statistically significant [$F_{(4, 170)} = 17.25, p < 0.001$]. Therefore, statistically controlling for positive and negative mood at the time did not modify the influence of the videos on the AAAQ.

Visual Probe Task (Table 1)

Data were analyzed in accordance with previous studies (e.g., Field et al., 2004). Firstly, trials with errors were discarded, and then outlying reaction times were removed if they were faster than 200 ms, slower than 2000 ms, and then if they were more than three standard deviations above the individual mean. All data from three participants were excluded as they had an

outlying high rate ($>28\%$) of missing data due to errors and outliers. For the remainder of the sample, on average 7% of trials were missing due to errors and a further 1% due to outliers, and these values did not differ between groups ($p > 0.1$).

Mean reaction times for different trial types and SOAs were analyzed using a $3 \times 2 \times 3$ mixed design ANOVA, with within-subject factors of trial type (3: congruent alcohol trials, incongruent alcohol trials, neutral-neutral trials) and SOA (2: 50, 500 ms), and a between-subjects factor of group. The predicted trial type \times SOA \times group interaction was not statistically significant [$F_{(4, 170)} = 1.16, p > 0.1$]. There was a significant main effect of SOA [$F_{(1, 85)} = 390.10, p < 0.001$], indicating faster reaction times on 500 ms trials compared to 50 ms trials. Importantly, the non-significant main effect of trial type [$F_{(2, 84)} = 1.09, p > 0.1$], and trial type \times SOA interaction [$F_{(2, 84)} = 2.12, p > 0.1$] demonstrate that there was no reliable attentional bias for alcohol cues overall, at either SOA.

Discussion

Overall, results from this study did not support the independence of approach and avoidance orientations for alcohol in either controlled or automatic processes. Data from the AAAQ could be interpreted as independence of self-reported approach and avoidance inclinations for alcohol after watching a video depicting the positive consequences of alcohol consumption, because participants who watched this video reported an increase in self-reported approach inclinations (inclined-indulgent and obsessed-compelled subscales) but no corresponding reduction in avoidance inclinations (the resolved-regulated subscale). However, a video that depicted the negative consequences of alcohol consumption prompted an increase in self-reported avoidance inclinations (the resolved-regulated subscale) in parallel with a decrease in self-reported approach inclinations (the inclined-obsessed and resolved-regulated subscales). Comparisons between these groups and a control group revealed that approach and avoidance inclinations were similar in the control group and the group that had watched the video depicting the positive consequences of alcohol consumption, whereas approach inclinations were lower, and avoidance inclinations higher, in the group that had watched the video depicting the negative consequences of alcohol consumption, compared to the control group.

The visual probe task revealed no evidence of attentional bias or attentional avoidance of alcohol cues in any group (or in the sample as a whole), therefore our hypotheses regarding the influence of the videos on attentional bias can be rejected. One interpretation is that attentional bias is insensitive to experimental manipulations of the motivation to drink or to avoid alcohol, although the absence of attentional bias in the control group argues against this interpretation. In the next study we repeated the general methodology of study 1 so that we were again able to investigate the effects of the different videos on self-reported approach and avoidance of alcohol. Given the null results from the visual probe task, we omitted this task and replaced it with a measure of automatic approach and avoidance tendencies for alcohol cues.

TABLE 1 | Mean reaction times (in milliseconds) from the different trials of the visual probe task in study 1.

	Alcohol-positive	Alcohol-negative	Control
50 ms TRIALS			
Congruent alcohol	732.20 \pm 59.29	725.82 \pm 70.42	710.73 \pm 67.41
Incongruent alcohol	734.37 \pm 58.31	726.36 \pm 73.72	699.75 \pm 67.06
Neutral-neutral	729.02 \pm 53.73	729.39 \pm 73.26	699.35 \pm 65.39
500 ms TRIALS			
Congruent alcohol	669.01 \pm 59.49	670.54 \pm 62.12	654.20 \pm 68.86
Incongruent alcohol	680.24 \pm 43.37	670.88 \pm 62.30	664.84 \pm 60.43
Neutral-neutral	675.89 \pm 48.14	671.86 \pm 67.49	648.36 \pm 73.90

Values are means \pm SD.

Study 2

In the alcohol version of the stimulus-response compatibility (SRC) task (Field et al., 2008), a manikin is presented on a computer screen either above or below an alcohol-related or neutral picture. Participants must move the manikin toward or away from the pictures as quickly as possible. On some blocks of the task, participants must make the manikin move toward alcohol pictures and away from neutral pictures, whereas these instructions are reversed in other blocks of the task. Automatic approach tendencies for alcohol cues are inferred if participants are faster to respond on blocks of the task when alcohol pictures require the “approach” movement in comparison to blocks when alcohol pictures require the “avoidance” movement. By contrast, if participants are faster on the “avoid alcohol” blocks compared to the “approach alcohol” blocks, this would suggest that alcohol cues evoke automatic avoidance tendencies. Heavy drinkers who are not seeking treatment display automatic approach tendencies for alcohol cues (Field et al., 2008, 2011; Christiansen et al., 2012; Kersbergen et al., 2015), whereas alcohol-dependent patients may show the opposite pattern, i.e., they are faster to avoid rather than approach alcohol-related pictures (Spruyt et al., 2013; but see Barkby et al., 2012).

Findings obtained from the standard version of the SRC task must be interpreted cautiously because this task yields an index of automatic approach that is *relative* to automatic avoidance, therefore an apparent bias in automatic approach could be attributed to either strong automatic approach, weak automatic avoidance, or a combination of the two. Among alcohol-dependent patients, if alcohol cues simultaneously evoke strong automatic approach at the same time as strong automatic avoidance, this may explain why they display either no reliable bias on the task (Barkby et al., 2012) or a bias to faster avoidance (Spruyt et al., 2013) depending on the strength of their automatic tendencies to avoid alcohol at the time of testing. In the present study, we overcame this limitation by modifying the task so that it is able to distinguish automatic alcohol approach and avoidance tendencies from each other. This modified version of the task includes neutral movements (to the side) in addition to the standard approach and avoidance movements, and is split into four blocks instead of two (see Baker et al., 2014).

The present study was identical to study 1 with the important difference that participants completed a modified SRC task instead of the visual probe task that was used in study 1. We hypothesized that we would replicate the effects of the videos on the AAAQ that were observed in study 1. Regarding the indices of approach and avoidance tendencies from the SRC task, we hypothesized that, relative to the control group, participants in the alcohol-positive group would show stronger automatic alcohol approach tendencies but these groups would not differ in automatic avoidance tendencies. By contrast, relative to the control group we anticipated stronger automatic avoidance tendencies in the alcohol-negative group, but these two groups would not differ in automatic approach tendencies.

Methods

Participants

Ninety participants (56 Female, mean age 24.56, $SD = 5.34$) were recruited from the local community and students and staff at the University of Liverpool via online and poster advertising. Inclusion and exclusion criteria were identical to those described for study 1. Participants who had taken part in studies 1 or 3 were ineligible to participate. Participants provided informed consent before taking part in the study, which was approved by the University of Liverpool Research Ethics Committee.

Materials

The modified stimulus-response compatibility task

The modified stimulus-response compatibility Task (Baker et al., 2014) is used to measure automatic approach and avoidance responses evoked by alcohol-related cues. Participants are instructed to rapidly categorize alcohol-related and stationery-related (control) pictures by moving a manikin either toward or away from the pictures, or to the left (neutral movement), as quickly as possible by pressing one of three specific keys on the keyboard, which were labeled with arrows pointing up, down, and left. The task was programmed in Inquisit software (Millisecond Software, 2006) and presented on a laptop computer with a 13 inch screen.

The format of the task, trial structure, and perceptual characteristics of the pictorial stimuli were identical to those used in previous studies (Field et al., 2011; Barkby et al., 2012). Fourteen colored pictures (a subset of the picture set used in Barkby et al., 2012) were used in the task: seven pictures of alcoholic drinks and close-ups of individuals holding or consuming those drinks, and seven control pictures of stationery items and close-ups of models interacting with those items.

There were four sub-blocks of the task, which differed according to task instructions. In the “approach alcohol” block, participants were required to move the manikin toward alcohol pictures, and to the left for stationery pictures. In the “avoid alcohol” block, participants moved away from alcohol pictures and to the left for stationery pictures. In the “approach control” block, participants moved toward stationery pictures and to the left for alcohol pictures. Finally, in the “avoid control” block, participants moved away from stationery pictures and to the left for alcohol pictures. Note that in the case of approach and avoidance movements, the position of the manikin was crucial: if the manikin was above the picture, an “approach” response required participants to press the “down” key, and an “avoidance” response required participants to press the “up” key. This was reversed if the manikin was below the picture. Participants were instructed to respond quickly and accurately on each trial. If they pressed the correct key, the manikin moved up, down or to the left in an animation lasting 500 ms. If they pressed the wrong key, error feedback was provided in the form of a large red cross presented in the center of the screen for 500 ms. There was an inter-trial interval of 500 ms.

Each sub-block of the task comprised four practice trials, in which two alcohol pictures and two control pictures were presented, once with the manikin above each picture type and

once with the manikin below. If participants did not understand the task, this practice block was repeated. There then followed 28 “critical” trials, in which each of the 14 pictures was presented twice: once with the manikin above the picture and once with the manikin below. Trials were presented in a new random order for each participant. Participants completed the sub-blocks in a counterbalanced order. Responses and reaction times (in milliseconds) to initiate the manikin movement were recorded on each trial.

Procedure

Participants were randomly allocated to experimental conditions. They were tested in a laboratory in the Department of Psychological Sciences or in quiet public places in which alcohol was not available (e.g., cafes and libraries). After providing informed consent, participants completed the timeline followback drinking diary, AUDIT, AAAQ, and PANAS (time 1). Then, participants put on the headphones and watched one of the videos, before completing the Video Questionnaire and the AAAQ and PANAS again (time 2). Finally, participants completed the SRC task. After completing the study, participants were debriefed and offered either course credit or a £5 Shopping Voucher to compensate them for their time.

Results

Group Characteristics

Participants reported consuming 30.21 ($SD = 23.53$) units of alcohol per week, and the mean score on the AUDIT was 12.85 ($SD = 5.34$). There were no between-group differences in age, weekly alcohol consumption, or AUDIT scores (all Kruskal–Wallis tests $p > 0.1$). There were no group differences in gender ratio ($\chi^2 = 0.66, p > 0.1$).

Effects of Video Manipulation on AAAQ Ratings (Figure 1B)

AAAQ ratings were analyzed using a mixed design ANOVA, with within-subject factors of sub-scale (3: inclined-indulgent, obsessed-compelled, resolved-regulated), time (2: before video, after video), and group (3: alcohol-positive, alcohol-negative, control). The sub-scale \times time \times group interaction was statistically significant [$F_{(4, 174)} = 19.16, p < 0.001$]. Subsequent *post-hoc* ANOVAs confirmed that the time \times group interaction was significant for all three sub-scales [inclined-indulgent $F_{(2, 87)} = 19.67, p < 0.001$; obsessed-compelled $F_{(2, 87)} = 5.28, p < 0.01$; resolved-regulated $F_{(2, 87)} = 13.80, p < 0.001$].

Groups did not differ on the inclined-indulgent [$F_{(2, 89)} = 1.56, p > 0.1$] or obsessed-compelled [$F_{(2, 89)} = 1.26, p > 0.1$] sub-scales before watching the video. However, there was a group difference in the resolved-regulated sub-scale before the video [$F_{(2, 89)} = 3.90, p < 0.05$], and *post-hoc* LSD contrasts revealed that scores were lower in the control group compared to both the alcohol-positive and alcohol-negative groups ($p < 0.05$), who did not differ from each other ($ps > 0.1$). As predicted, groups differed on all three sub-scales after watching the video [inclined-indulgent $F_{(2, 89)} = 10.78, p < 0.001$; obsessed-compelled $F_{(2, 89)} = 4.85, p = 0.01$; resolved-regulated

$F_{(2, 89)} = 20.78, p < 0.001$]. *Post-hoc* LSD contrasts revealed that scores on both inclined-indulgent and obsessed-compelled sub-scales were higher in the alcohol-positive group compared to both alcohol-negative and control groups ($p < 0.01$), who did not differ from each other ($p > 0.1$). On the other hand, scores on the resolved-regulated sub-scale were higher in the alcohol-negative group compared to both alcohol-positive and neutral groups ($ps < 0.01$), who did not differ from each other ($p > 0.08$).

Paired-samples *t*-tests revealed that among participants in the alcohol positive group, scores on the inclined-indulgent and obsessed-compelled sub-scales increased after watching the video [$t_{(29)} = 2.74, p = 0.01$ and $t_{(29)} = 2.84, p < 0.01$], whereas scores on the resolved-regulated sub-scale decreased [$t_{(29)} = 2.90, p < 0.01$]. The reverse pattern was seen in the alcohol negative group: the decrease in inclined-indulgent ratings and the increase in resolved-regulated ratings after watching the video were statistically significant [$t_{(29)} = 6.70, p < 0.001$ and $t_{(29)} = 3.15, p < 0.01$], although there was no significant change in scores on the obsessed-compelled sub-scale [$t_{(29)} = 0.32, p > 0.1$]. In the control group, scores on both the inclined-indulgent and resolved-regulated sub-scales decreased after watching the video [$t_{(29)} = 2.56, p < 0.05$ and $t_{(29)} = 2.63, p < 0.05$], but scores on the obsessed-compelled sub-scale did not change [$t_{(29)} = 0.27, p > 0.1$].

We also re-ran the omnibus Three-Way ANOVA on AAAQ scores but added PANAS positive and PANAS negative affect after the video as covariates. The three way interaction sub-scale \times time \times group remained statistically significant [$F_{(4, 170)} = 9.55, p < 0.001$]. Therefore, statistically controlling for positive and negative mood at the time did not influence the influence of the videos on the AAAQ.

SRC Task (Table 2)

Data were analyzed in accordance with previous studies (e.g., Field et al., 2011). Firstly, trials with errors were discarded, and then outlying reaction times were removed if they were faster than 200 ms, slower than 2000 ms, and then if they were more than three standard deviations above the individual mean. All data from three participants were excluded as they had an outlying high rate (>40%) of missing data due to errors and outliers. For the remainder of the sample, on average 5% of trials were missing due to errors and a further 9% due to outliers, and these values did not differ between groups ($ps > 0.1$).

TABLE 2 | Mean reaction times (in milliseconds) from the different blocks of the SRC task in study 2.

	Alcohol-positive	Alcohol-negative	Control
Approach alcohol	826.51 \pm 165.05	855.10 \pm 186.92	871.38 \pm 208.88
Approach stationery	855.99 \pm 198.96	878.74 \pm 186.11	917.37 \pm 195.66
Avoid alcohol	835.32 \pm 133.79	893.49 \pm 207.91	895.22 \pm 208.31
Avoid stationery	878.81 \pm 174.65	918.22 \pm 202.01	902.41 \pm 211.75

Values are means \pm SD.

Mean reaction times in the different blocks of the task were then analyzed using a mixed design $2 \times 2 \times 3$ ANOVA, with within-subject factors of movement type (2: approach, avoidance) and picture type (2: alcohol, stationery; this refers to the type of picture that the approach or avoidance movement had to be directed toward or away from, with the sideways movement required for the other type of picture), and a between-subjects factor of group. The hypothesized three way interaction was not statistically significant [$F_{(2, 84)} = 0.80, p > 0.1$]. There were, however, significant main effects of picture type [$F_{(1, 84)} = 4.75, p < 0.05$; participants were faster to respond on blocks when the approach or avoidance movement had to be made in response to alcohol pictures rather than stationery pictures], and movement type [$F_{(1, 84)} = 4.82, p < 0.05$; participants were faster on “approach” blocks than “avoid” blocks of the task). The picture type \times movement type interaction was not statistically significant [$F_{(1, 84)} = 0.19, p > 0.1$]. Overall, these results show that participants were faster to make approach rather than avoidance movements, and they were faster to make both approach and avoidance movements in response to alcohol pictures in comparison to stationery pictures. However, the video manipulation had no effect on performance on the task.

Discussion

Consistent with results from study 1, results from this study did not provide clear support for the independence of self-reported approach and avoidance inclinations for alcohol following an experimental manipulation of those inclinations. Between-group contrasts demonstrated that, relative to the control group, scores on the inclined-indulgent and obsessed-compelled subscales of the AAAQ were elevated in participants who had watched a video depicting the positive consequences of alcohol consumption, but this video did not influence scores on the resolved-regulated subscale. The complete opposite pattern was seen in participants who had watched a video depicting the negative consequences of alcohol consumption: increased scores on the resolved-regulated subscale, but there was no difference between this group and the control group in scores on the inclined-indulgent and obsessed-compelled subscales. These contrasts support predictions made by the ambivalence model of craving (Breiner et al., 1999), because they suggest that it is possible to experimentally manipulate subjective approach and avoidance inclinations for alcohol independently of each other. Unfortunately, this conclusion must be heavily caveated given the presence of group differences on the resolved-regulated subscale at baseline, and because within-subject contrasts suggest that increases in approach inclinations were accompanied by decreases in avoidance inclinations in the alcohol positive group, and vice versa for the alcohol negative group.

The data from the SRC task also did not support our hypotheses: there was no evidence that participants were faster to approach or slower to avoid alcohol cues compared to stationery (control) cues, and the experimental manipulation did not influence the task. Therefore, even though our experimental manipulation had a clear influence on self-reported approach and

avoidance inclinations for alcohol, there were no parallel changes in automatic approach or avoidance tendencies evoked by alcohol cues. In our third and final study, we again investigated the influence of alcohol-positive and alcohol-negative videos on self-reported (AAAQ) and automatic (modified SRC task) indices of approach and avoidance motivational orientations for alcohol, but we combined this with an experimental manipulation of thought suppression in an attempt to maximize the influence of the video manipulation on automatic measures of approach and avoidance.

Study 3

People who are attempting to reduce their alcohol consumption often attempt to suppress unwanted thoughts, such as intrusive cravings, in order to achieve their goal (Moss et al., 2015). Unfortunately, thought suppression has unwelcome consequences because it increases, rather than decreases the frequency of intrusive thoughts that are the target of suppression, but only when competing demands are placed on cognitive resources (Wenzlaff and Wegner, 2000). Indeed, attempting to suppress thoughts about alcohol paradoxically increases the accessibility of alcohol-related cognitions as evidenced by increased attentional bias for alcohol words (Klein, 2007) and accessibility of alcohol-related semantic associations (Palfai et al., 1997).

We hypothesized that if participants were primed to think about the positive or negative consequences of alcohol consumption and were then instructed to suppress those thoughts, this should provoke an increase in the accessibility of those thoughts that would manifest itself as a bias in automatic approach or avoidance tendencies in response to alcohol cues. Specifically, participants who viewed a video depicting the positive consequences of alcohol and then attempted to suppress thoughts about alcohol should have stronger automatic approach tendencies for alcohol cues compared to participants who watched the same video but did not attempt to suppress their thoughts. We expected comparable moderating effects of thought suppression on automatic avoidance tendencies in participants who watched a video depicting the negative consequences of alcohol.

Methods

Participants

One hundred participants (53 Female, mean age 27.87, $SD = 6.97$) were recruited from the local community and students and staff at the University of Liverpool via online and poster advertising. Inclusion and exclusion criteria were identical to those described for studies 1 and 2. Participants who had taken part in studies 1 or 2 were ineligible to participate. All participants provided informed consent before taking part in the study, which was approved by the University of Liverpool Research Ethics Committee.

Procedure

Participants were randomly allocated to one of four experimental conditions: (1) alcohol-positive video combined with thought

suppression, (2) alcohol-positive video combined with control manipulation, (3) alcohol-negative video combined with thought suppression, or (4) alcohol-negative video combined with control manipulation. Note that none of the participants in this study watched the neutral video that we used in studies 1 and 2. Participants were tested in a laboratory in the Department of Psychological Sciences or in quiet public places in which alcohol was not served (e.g., cafes, libraries). After providing informed consent, participants completed the timeline follow back drinking diary, AUDIT, AAAQ, and PANAS (time 1). Then, participants put on the headphones and watched one of the videos, before completing the Video Questionnaire and the AAAQ and PANAS again (time 2).

Participants in the thought suppression groups were then instructed to think about anything, but to make every effort to suppress thoughts of alcohol; the importance of the latter was emphasized. Participants in the control groups were instructed to think about anything that came to mind, including alcohol. All participants were then given a 5 min to think freely and to write notes about what they were thinking about on a piece of paper. They were also instructed to place a mark in the right hand margin of the paper each time they thought about alcohol; these marks were subsequently counted up and cross-checked with the content of participants' notes.

Immediately after this 5-min period, participants were asked to respond to two questions by placing a mark on 100 ms visual analog scales (VAS). The questions, which were the same for both groups, were: "To what extent did you think about alcohol?" (anchors "I did not think about alcohol at all" and "I thought about alcohol a lot") and "To what extent did you succeed in complying with the instructions" (anchors "totally unsuccessful" and "totally successful").

Participants then completed the SRC task. The thought suppression or control instructions were re-iterated to participants before each block of the task. In order to increase demands on working memory, participants were given a seven-digit number at the beginning of each sub-block of the task. They were given 50 s to memorize the number and then instructed to hold it in memory, as they would be asked to recall it at the end of each sub-block of the task. Recall of this number was recorded at the end of each sub-block and the process was repeated with a different number at the beginning of each sub-block (see Bryant et al., 2011). After completing all blocks of the SRC task, participants again completed the two 100 ms VAS to indicate the extent to which they had thought about alcohol, and had complied with instructions, whilst they were doing the task. Finally, participants were debriefed and offered either course credit or a £5 Shopping Voucher to compensate them for their time.

Results

Group Characteristics

Participants reported consuming 22.53 ($SD = 15.63$) units of alcohol per week, and the mean score on the AUDIT was 10.56 ($SD = 4.52$). There was no group difference in gender ratio ($\chi^2 = 4.62$, $p > 0.1$), although there were group differences in both weekly alcohol consumption and

AUDIT scores (Kruskal-Wallis tests, $ps < 0.05$). Participants in both thought suppression groups had higher weekly alcohol consumption and higher scores on the AUDIT compared to participants in both control groups. Therefore, we repeated all primary analyses (detailed below) with the addition of weekly alcohol consumption and AUDIT scores as covariates.

Effects of Video Manipulation on AAAQ Ratings (Figure 1C)

AAAQ ratings were analyzed using a mixed design ANOVA, with within-subject factors of sub-scale (3: inclined-indulgent, obsessed-compelled, resolved-regulated), time (2: before video, after video), and between-subject factors of video group (2: alcohol-positive, alcohol-negative), and thought suppression group (2: thought suppression, control). The sub-scale \times time \times video group interaction was statistically significant [$F_{(2, 95)} = 41.60$, $p < 0.001$] but the four way interaction sub-scale \times time \times video group \times thought suppression group was not [$F_{(2, 95)} = 1.17$, $p > 0.1$]. The three way interaction remained significant after adding AUDIT scores and weekly alcohol consumption as covariates [$F_{(2, 93)} = 38.40$, $p < 0.001$]. Subsequent *post-hoc* ANOVAs confirmed that the time \times video group interaction was significant for all three sub-scales [inclined-indulgent $F_{(1, 98)} = 40.51$, $p < 0.001$; obsessed-compelled $F_{(1, 98)} = 23.84$, $p < 0.001$; resolved-regulated $F_{(1, 98)} = 34.35$, $p < 0.001$].

Groups did not differ on any of the sub-scales before watching the video [inclined-indulgent $t_{(98)} = 0.08$, $p > 0.1$; obsessed-compelled $t_{(98)} = 1.62$, $p > 0.1$; resolved-regulated $t_{(98)} = 1.46$, $p > 0.1$]. As predicted, groups differed on all three sub-scales after watching the video [inclined-indulgent $t_{(98)} = 4.06$, $p < 0.001$, higher in the alcohol-positive video group; obsessed-compelled $t_{(98)} = 2.15$, $p < 0.05$, also higher in the alcohol-positive video group; resolved-regulated $t_{(98)} = 5.68$, $p < 0.001$, higher in the alcohol-negative video group].

Paired-samples *t*-tests revealed that in the alcohol-positive video group, scores on the inclined-indulgent and obsessed-compelled sub-scales increased after watching the video [$t_{(49)} = 1.96$, $p < 0.05$ and $t_{(49)} = 3.12$, $p < 0.01$], whereas scores on the resolved-regulated sub-scale decreased [$t_{(49)} = 2.20$, $p < 0.05$]. The reverse pattern was seen in the alcohol-negative video group: scores on the inclined-indulgent and obsessed-compelled sub-scales decreased after watching the video [$t_{(49)} = 6.16$, $p < 0.001$ and $t_{(49)} = 3.95$, $p < 0.001$] whereas scores on the resolved-regulated sub-scale increased [$t_{(49)} = 5.45$, $p < 0.001$].

We also re-ran the omnibus Four-Way ANOVA on AAAQ scores but added PANAS positive and PANAS negative affect after the video as covariates. The three way interaction sub-scale \times time \times group remained statistically significant [$F_{(2, 93)} = 30.79$, $p < 0.001$]. Therefore, statistically controlling for positive and negative mood at the time did not influence the influence of the videos on the AAAQ.

SRC Task (Table 3)

All data were missing from two participants due to an experimenter error. Data were analyzed as described for study 2. All data from two additional participants were excluded as they had an outlying high rate ($>35\%$) of missing data

TABLE 3 | Mean reaction times (in milliseconds) from the different blocks of the SRC task in study 3.

	Suppress positive	Suppress negative	Control positive	Control negative
Approach alcohol	905.47 ± 256.62	815.36 ± 242.46	943.52 ± 261.85	871.95 ± 244.25
Approach stationery	1018.24 ± 308.49	877.71 ± 236.56	960.42 ± 234.93	928.08 ± 289.12
Avoid alcohol	934.29 ± 219.92	842.58 ± 238.03	954.69 ± 207.22	882.97 ± 223.72
Avoid stationery	936.15 ± 244.13	871.00 ± 234.06	941.56 ± 207.04	915.88 ± 249.54

Values are means ± SD.

TABLE 4 | The effect of thought suppression on self-reported alcohol-related thoughts and task instructions.

	Suppress positive	Suppress negative	Control positive	Control positive
IMMEDIATELY AFTER THOUGHT SUPPRESSION AND THOUGHT LISTING EXERCISE:				
"To what extent did you think about alcohol"?	30.92 ± 27.23	37.96 ± 31.43	43.60 ± 32.34	43.60 ± 31.08
"To what extent did you succeed in complying with instructions"?	67.48 ± 26.71	59.76 ± 32.48	74.60 ± 22.63	71.36 ± 25.57
Number of alcohol-related thoughts recorded	3.92 ± 3.92	4.68 ± 3.81	4.32 ± 2.82	4.76 ± 3.50
IMMEDIATELY AFTER COMPLETING SRC TASK:				
"To what extent did you think about alcohol"?	27.64 ± 26.23	18.16 ± 20.45	36.72 ± 32.61	42.32 ± 31.51
"To what extent did you succeed in complying with instructions"?	75.36 ± 15.16	75.76 ± 25.08	78.72 ± 21.35	74.88 ± 21.24

Visual analog scales are 100 ms VAS. Values are means ± SD.

due to errors and outliers. For the remaining participants, on average 4% of trials were missing due to errors and a further 7% due to outliers; these values did not differ between groups ($ps > 0.1$).

Mean reaction times in the different blocks of the task were then analyzed using a mixed design $2 \times 2 \times 2 \times 2$ ANOVA, with within-subject factors of movement type (2: approach, avoidance) and picture type (2: alcohol, stationery; this refers to the type of picture that the approach or avoidance movement had to be directed toward or away from, with the sideways movement required for the other type of picture), and between-subject factors of video group (2: alcohol-positive, alcohol-negative), and thought suppression group (2: thought suppression, control). The hypothesized four way interaction was not statistically significant [$F_{(1, 92)} = 0.99, p > 0.1$]. There was, however, a significant main effect of picture type [$F_{(1, 92)} = 11.22, p < 0.01$] which was subsumed under a significant picture type \times movement type interaction [$F_{(1, 92)} = 7.85, p < 0.01$]. Paired-samples t -tests revealed that participants were significantly faster to approach alcohol rather than control pictures [$t_{(95)} = 4.45, p < 0.001$], but reaction times to avoid alcohol and control pictures did not differ [$t_{(95)} = 0.97, p > 0.1$].

There were no other significant main effects or interactions ($F_s < 1.68, ps > 0.1$), and results were unaffected when the analysis was repeated with AUDIT scores and weekly alcohol consumption added as covariates. Overall, these results show that participants were faster to make approach movements to alcohol pictures than control pictures, but there was no difference in the speed of avoidance movements. Most importantly, neither the video manipulation or the thought suppression manipulation, or the interaction between the two, had any effect on performance on the task.

Thought Suppression and Working Memory Load Manipulation Checks

Responses on the visual analog scales, and the number of alcohol-related thoughts that participants recorded, are shown in **Table 4**. Each VAS was analyzed using a separate 2×2 ANOVA, with between-subject factors of video group (2: alcohol-positive, alcohol-negative), and thought suppression group (2: thought suppression, control). There were no main effects or interactions for the "To what extent did you succeed in complying with the instructions"? VAS at either time ($F_s < 2.99, ps > 0.08$). There were no main effects or interactions for the "To what extent did you think about alcohol" question immediately after the thought suppression and thought listing exercise ($F_s < 2.24, ps > 0.1$), which suggests that the thought suppression manipulation was not effective. However, the main effect of thought suppression group was statistically significant for this question immediately after participants had completed the SRC task [$F_{(1, 99)} = 8.74, p < 0.001$], as participants in the thought suppression group reported significantly fewer alcohol-related thoughts while completing the SRC task than participants in the control group. There were no other main effects or interactions ($F_s < 1.80, ps > 0.1$). Finally, there were no significant main effects or interactions for the number of alcohol-related thoughts that participants recorded during the thought suppression and thought listing exercise ($F_s < 0.72, ps > 0.1$).

Overall, these results indicate that the thought suppression manipulation was not successful, because there were no differences in perceived suppression success or the number of alcohol-related thoughts recorded between thought suppression and control groups during the thought listing exercise. However, when asked immediately after completing the SRC task, participants in the thought suppression group reported that

they thought about alcohol significantly *less* than participants in the control group. Furthermore, the lack of significant thought suppression \times video group interactions for these measures suggests that the alcohol positive and alcohol negative videos did not have differential effects on the success of attempted thought suppression.

Finally, 97% of participants successfully recalled the 7-digit number at the end of each sub-block of the SRC task, which demonstrates that compliance with the manipulation of working memory load was very high.

Discussion

Consistent with the findings from studies 1 and 2, results from this study demonstrated that participants who watched a video depicting the positive consequences of alcohol reported an increase in approach inclinations for alcohol that was accompanied by a reduction in avoidance inclinations; the converse pattern was seen among participants who watched a video depicting the negative consequences of alcohol consumption.

The primary novel feature of this study was the incorporation of a thought suppression manipulation in an attempt to magnify the influence of the videos on automatic approach and avoidance responses evoked by alcohol-related cues. Contrary to our hypotheses, we observed no evidence that the thought suppression manipulation led to an increase in alcohol-related thoughts; by contrast, participants' self-reports indicated that they were able to suppress alcohol when instructed to do so. We observed that participants were faster to approach alcohol rather than control pictures, thereby replicating previous findings using a related task (Field et al., 2008, 2011; Christiansen et al., 2012; Kersbergen et al., 2015). However, this pattern of results was unaffected by the videos, the thought suppression manipulation, or the interaction between the two.

Combined Analysis

Results from all three studies demonstrated that self-reported inclinations to approach and avoid alcohol were not independent of each other: increases in approach inclinations were accompanied by parallel decreases in avoidance inclinations, and vice versa. This interpretation could be bolstered by investigating the strength of the associations between *changes* in subjective approach and avoidance inclinations for alcohol after exposure to videos depicting the positive and negative consequences of alcohol consumption.

To this end, we combined the AAAQ data from all three studies, but disregarded data from the control groups (participants who watched the neutral video) in studies 1 and 2. Given that the thought suppression manipulation had no influence on the AAAQ in study 3, we collapsed these data across thought suppression groups. This combined analysis, with a total sample size of 221, confirmed that the sub-scale \times time \times group interaction was highly statistically significant, and that the time \times group interactions were highly significant for all three of the AAAQ subscales (all $ps < 0.001$). Furthermore, paired samples t -tests confirmed that in the combined alcohol-positive

group, scores on the inclined-indulgent and obsessed-compelled subscales increased, whereas scores on the resolved-regulated subscale decreased, after watching the video. The reverse pattern was seen in the combined alcohol-negative group. Data are shown in Table S4.

We then computed change scores to capture the change in each AAAQ subscale after participants watched the videos. By correlating these change scores with each other we were able to investigate the strength of the association between changes in self-reported approach and avoidance inclinations: as approach inclinations increase, do avoidance inclinations decrease by a similar magnitude (and vice versa)? Overall, intercorrelations between these change scores were statistically significant but small. After watching a video depicting the positive consequences of alcohol consumption, the magnitude of the increase in scores on the inclined-indulgent and obsessed-compelled subscales was associated with the magnitude of the decrease in scores on the resolved-regulated subscale, although the size of the correlation co-efficients ($r = -0.20$) suggests only 4% shared variance. After watching a video depicting the negative consequences of alcohol consumption, the magnitude of the increase in scores on the resolved-regulated subscale was weakly associated with the reduction in scores on the inclined-indulgent subscale ($r = -0.36$; 13% shared variance), but was unrelated to the reduction in scores on the obsessed-compelled subscale.

Finally, we are grateful to the reviewer who suggested the following additional analyses, which are detailed in Table S5. We performed hierarchical linear regression analyses to investigate if the change score for self-reported approach inclinations would be predicted by the video manipulation, even after entering the change score for self-reported avoidance inclinations in the first step of the model. With the score on the inclined-indulgent subscale as the dependent variable, the score on the resolved-regulated subscale accounted for 22% of variance [$F_{(1, 279)} = 77.29, p < 0.01$], but addition of experimental group as a predictor in the subsequent step of the regression accounted for an additional 8% of variance [$F\Delta_{(1, 277)} = 29.73, p < 0.01$]. Similarly, with the score on the obsessed-compelled subscale as the dependent variable, the score on the resolved-regulated subscale accounted for 7% of variance [$F_{(1, 279)} = 22.20, p < 0.01$], but addition of experimental group as a predictor accounted for an additional 6% of variance [$F\Delta_{(1, 277)} = 18.42, p < 0.01$].

However, we did not observe the same pattern when scores on the resolved-regulated subscale were entered as the dependent variable, and scores on the inclined-indulgent and obsessed-compelled subscales were entered as independent variables (in separate analyses). In the first case, after accounting for the 22% of variance attributable to the inclined-indulgent subscale [$F_{(1, 279)} = 77.29, p < 0.01$], the addition of experimental group as a predictor did not account for additional variance in scores on the resolved-regulated subscale [$<1\%$ of additional variance; $F\Delta_{(1, 277)} = 1.07, p = 0.30$]. Similarly, after accounting for the 7% of variance attributable to the obsessed-compelled subscale [$F_{(1, 279)} = 22.20, p < 0.01$], the addition of experimental group as a predictor did not account for additional variance in scores

on the resolved-regulated subscale [$<1\%$ of additional variance; $F\Delta(1, 277) = 0.20, p = 0.66$].

These analyses reveal that self-reported approach and avoidance inclinations for alcohol can operate at least partly independently of each other, but the effect is not symmetrical. Changes in self-reported avoidance inclinations after watching videos depicting the positive or negative consequences of alcohol consumption were completely accounted for by changes in self-reported approach inclinations. However, the changes in self-reported approach inclinations that were evoked by these videos were at least partly independent of changes in self-reported avoidance inclinations.

General Discussion

A number of consistent findings emerged from the three studies reported here. When participants viewed short videos that depicted the positive or negative consequences of alcohol consumption, their self-reported approach and avoidance inclinations for alcohol tended to change in parallel: as approach inclinations increased, avoidance inclinations decreased, and vice versa. However, between-group contrasts suggested some degree of independence of approach and avoidance inclinations, although these findings were not consistent across studies. In addition, although *changes* in approach and avoidance inclinations tended to be inversely correlated (as one increased, the other decreased), these correlations were small and there was some evidence that changes in approach inclinations were partly independent of changes in avoidance inclinations. Finally, we found no evidence in support of our predictions that alcohol-related implicit cognitions would be influenced by experimental manipulations of motivational orientations for alcohol, regardless of whether we measured attentional biases or automatic approach / avoidance tendencies, or whether the experimental manipulation was combined with a thought suppression exercise.

One of the primary aims of these studies was to expand on findings from previous studies that used the Approach and Avoidance of Alcohol Questionnaire (AAAQ) in order to test predictions made by the ambivalence model of craving (Breiner et al., 1999). Specifically, if subjective approach and avoidance inclinations for alcohol are independent of each other, it should be possible to dissociate them by exposing participants to experimental manipulations that are designed to increase one but not the other. Overall, our findings demonstrated that approach and avoidance inclinations tend to change in parallel because as one increased, the other tended to decrease. This casts doubt on the independence of these constructs as predicted by the ambivalence model. Specifically, within-subject contrasts revealed that, after watching a video depicting the positive consequences of alcohol consumption, approach inclinations increased and avoidance inclinations decreased, whereas the reverse pattern was seen in participants who watched a video depicting the negative consequences of alcohol consumption. Although there were some minor inconsistencies between studies, the results of a combined analysis of data from all studies confirmed that approach and avoidance inclinations

tended to fluctuate alongside each other. Furthermore, the combined analysis revealed that the magnitude of changes in approach and avoidance inclinations over time were reliably negatively correlated with each other: the magnitude of the increase in the strength of approach inclinations was associated with the magnitude of the corresponding decrease in avoidance inclinations, and vice versa. Finally, the regression analysis confirmed that variation in the change in avoidance inclinations after watching the videos was completely accounted for by the change in approach inclinations.

However, other analyses suggested that self-reported approach and avoidance inclinations could be characterized as at least partly independent of each other. Although the combined analysis confirmed inverse correlations between the magnitude of changes in approach and avoidance inclinations over time, these relationships were weak (with, at most, 13% shared variance), and regression analyses demonstrated that changes in approach inclinations were at least partly independent of changes in avoidance inclinations. More importantly, between-subject contrasts with a control group suggested that approach inclinations increased without a corresponding change in avoidance inclinations, and vice versa. Unfortunately, we have limited confidence in these group differences because they were only seen in study 2; in study 1, group differences in approach inclinations were accompanied by group differences in avoidance inclinations.

Our findings are consistent with some previous observations that subjective approach and avoidance inclinations for alcohol tend to change in parallel after exposure to appetitive alcohol cues (Curtin et al., 2005), although one study reported a dissociation between approach and avoidance inclinations after cue exposure (Jones et al., 2013). Importantly, the studies reported here are the very first to investigate the influence of an experimental manipulation that was intended to activate motivational orientations to avoid drinking; findings from participants in the alcohol-negative groups in all studies clearly suggest that this manipulation led to the predicted increase in self-reported avoidance inclinations that was accompanied by a decrease in approach inclinations. However, it is important to clarify that previous studies demonstrated that approach and avoidance inclinations have independent predictive validity for individual differences in drinking behavior and prospective drinking behavior (Curtin et al., 2005; Schlauch et al., 2012, 2013a,b,c; Klein and Anker, 2013). The three studies reported here cannot speak to the predictive validity of these constructs.

In contrast to the robust effects on self-reported approach and avoidance motivational orientations (assessed with the AAAQ), the videos depicting the positive and negative consequences of alcohol consumption had no effect on measures of motivational orientations operating within automatic processes, that is attentional biases (study 1) and approach / avoidance tendencies (studies 2 and 3). To our knowledge, these are the first studies that experimentally manipulated the motivation to avoid drinking and our findings suggest that these automatic processing biases are impervious to motivational orientations to avoid drinking. This interpretation is consistent with findings from an earlier study (Baker et al., 2014) in which we attempted

to prime motivational orientations by presenting alcohol-positive and alcohol-negative primes below the threshold of conscious awareness; this manipulation also failed to influence attentional biases and approach / avoidance tendencies. However, on the basis of previous demonstrations of robust, albeit weak associations between subjective craving (typically assessed with self-report instruments that capture only “approach” inclinations) and attentional bias (Field et al., 2009), and demonstrations that experimental manipulations of craving such as negative mood induction, exposure to alcohol cues, and acute alcohol intoxication all lead to increases in attentional bias (see Field and Cox, 2008), we anticipated elevated attentional biases in the “attend positive” vs. the control groups. This pattern of results was not seen. Perhaps most importantly, with one exception (the bias to more rapidly approach alcohol rather than control images in study 3), we found no evidence of attentional or approach or avoidance biases in any of the studies, and no evidence that individual differences in alcohol consumption, hazardous drinking or scores on the AAAQ were associated with these implicit processing biases in any of the studies (see Supplementary Materials). These findings cast doubt on the validity and sensitivity of the tasks that were used in the current studies, all of which were slightly modified versions of tasks that are more commonly used in the literature. Future investigations of this research question should attempt to develop measures of approach and avoidance inclinations operating in automatic processes that have acceptable construct validity and sensitivity for this purpose.

The studies reported here have other weaknesses in addition to the questionable construct validity of the implicit measures. All participants consumed alcohol in excess of UK government guidelines and therefore their alcohol consumption was placing their health at risk. However, we did not attempt to recruit participants who were concerned about or attempting to limit their alcohol consumption, and we did not measure participants’ motivation to change using a validated self-report measure, so it is possible that our participants were relatively insensitive to our experimental manipulations that were designed to exaggerate their ambivalence about alcohol consumption. Future studies could investigate this issue by recruiting heavy drinking participants who are currently motivated to reduce their alcohol consumption (and are actively attempting to do so), because motivational orientations in these participants might be expected

to be more sensitive to the experimental manipulations that were used in the present study. A further limitation is that we did not record participants’ occupational or socioeconomic status so we are unable to fully characterize participants who took part in these studies. In addition, it is possible that the videos had robust effects on self-report but not computerized measures because participants always completed the former before the latter; this could be investigated by counterbalancing the order in which assessments are administered in future studies. Our study also had strengths, including measurements of participants’ subjective mood after they had watched the videos, which enabled us to rule out changes in mood as a contributor to the influence of those videos on self-reported and automatic motivational orientations.

In conclusion, findings from the three studies reported here question the degree of independence of self-reported approach and avoidance inclinations for alcohol, which tended to co-vary in response to experimental manipulations of inclinations to drink or inclinations to avoid alcohol. However, results from a combined analysis of data from all studies suggest that changes in inclinations to drink may be at least partially independent of changes in inclinations to avoid alcohol. Our findings also suggest that measures of alcohol-related motivational orientations that operate in automatic processes are impervious to these experimental manipulations, although the modified tasks that we used here have questionable construct validity and sensitivity which suggests that these findings should be interpreted with caution.

Funding

Funded by a research grant from the Wellcome Trust, reference 086247/Z/08/Z, awarded to MF and JD.

Acknowledgments

We thank Abigail Kirkland, Georgina Schwarz, and Michael Tapp for their help with data collection.

Supplementary Material

The Supplementary Material for this article can be found online at: <http://journal.frontiersin.org/article/10.3389/fpsyg.2015.01465>

References

- Baker, S., Dickson, J. M., and Field, M. (2014). Implicit priming of conflicting motivational orientations in heavy drinkers. *BMC Psychol.* 2:28. doi: 10.1186/s40359-014-0028-1
- Barkby, H., Dickson, J. M., Roper, L., and Field, M. (2012). To approach or avoid alcohol? Automatic and self-reported motivational tendencies in alcohol dependence. *Alcohol. Clin. Exp. Res.* 36, 361–368. doi: 10.1111/j.1530-0277.2011.01620.x
- Breiner, M. J., Stritzke, W. G. K., and Lang, A. R. (1999). Approaching avoidance: a step essential to the understanding of craving. *Alcohol. Res. Health* 23, 197–206.
- Bryant, R. A., Wyzenbeek, M., and Weinstein, J. (2011). Dream rebound of suppressed emotional thoughts: the influence of cognitive load. *Conscious. Cogn.* 20, 515–522. doi: 10.1016/j.concog.2010.11.004
- Christiansen, P., Cole, J. C., Goudie, A. J., and Field, M. (2012). Components of behavioural impulsivity and automatic cue approach predict unique variance in hazardous drinking. *Psychopharmacology* 219, 501–510. doi: 10.1007/s00213-011-2396-z
- Christiansen, P., Schoenmakers, T. M., and Field, M. (2015). Less than meets the eye: reappraising the clinical relevance of attentional bias in addiction. *Addict. Behav.* 44, 43–50. doi: 10.1016/j.addbeh.2014.10.005
- Curtin, J. J., Barnett, N. P., Colby, S. M., Rohsenow, D. J., and Monti, P. M. (2005). Cue reactivity in adolescents: measurement of separate approach and avoidance reactions. *J. Stud. Alcohol* 66, 332–343. doi: 10.15288/jsa.2005.66.332

- Dickson, J. M., Gately, C., and Field, M. (2013). Alcohol dependent patients have weak negative rather than strong positive implicit alcohol associations. *Psychopharmacology* 228, 603–610. doi: 10.1007/s00213-013-3066-0
- Field, M., Caren, R., Fernie, G., and De Houwer, J. (2011). Alcohol approach tendencies in heavy drinkers: comparison of effects in a relevant stimulus-response compatibility task and an approach/avoidance Simon Task. *Psychol. Addict. Behav.* 25, 697–701. doi: 10.1037/a0023285
- Field, M., and Cox, W. M. (2008). Attentional bias in addictive behaviors: a review of its development, causes, and consequences. *Drug Alcohol Depend.* 97, 1–20. doi: 10.1016/j.drugalcdep.2008.03.030
- Field, M., Kiernan, A., Eastwood, B., and Child, R. (2008). Rapid approach responses to alcohol cues in heavy drinkers. *J. Behav. Ther. Exp. Psychiatry* 39, 209–218. doi: 10.1016/j.jbtep.2007.06.001
- Field, M., Mogg, K., and Bradley, B. P. (2005). Craving and cognitive biases for alcohol cues in social drinkers. *Alcohol Alcohol.* 40, 504–510. doi: 10.1093/alcac/agh213
- Field, M., Mogg, K., Mann, B., Bennett, G. A., and Bradley, B. P. (2013). Attentional biases in abstinent alcoholics and their association with craving. *Psychol. Addict. Behav.* 27, 71–80. doi: 10.1037/a0029626
- Field, M., Mogg, K., Zetteler, J., and Bradley, B. P. (2004). Attentional biases for alcohol cues in heavy and light social drinkers: the roles of initial orienting and maintained attention. *Psychopharmacology* 176, 88–93. doi: 10.1007/s00213-004-1855-1
- Field, M., Munafò, M. R., and Franken, I. H. A. (2009). A Meta-Analytic investigation of the relationship between attentional bias and subjective craving in substance abuse. *Psychol. Bull.* 135, 589–607. doi: 10.1037/a0015843
- Hettema, J., Steele, J., and Miller, W. R. (2005). Motivational interviewing. *Annu. Rev. Clin. Psychol.* 1, 91–111. doi: 10.1146/annurev.clinpsy.1.102803.143833
- Jones, A., Rose, A. K., Cole, J. C., and Field, M. (2013). Effects of alcohol cues on craving and *ad libitum* alcohol consumption in social drinkers: the role of disinhibition. *J. Exp. Psychopathol.* 4, 239–240. doi: 10.5127/jep.031912
- Kersbergen, I., Woud, M. L., and Field, M. (2015). The validity of different measures of automatic alcohol action tendencies. *Psychol. Addict. Behav.* 29, 225–230. doi: 10.1037/adb0000009
- Klein, A. A., and Anker, J. J. (2013). A psychometric evaluation of the approach and avoidance of alcohol questionnaire among alcohol-dependent patients attending residential treatment. *J. Psychopathol. Behav. Assess.* 35, 205–214. doi: 10.1007/s10862-012-9322-5
- Klein, A. A., Stasiewicz, P. R., Koutsky, J. R., Bradizza, C. M., and Coffey, S. F. (2007). A psychometric evaluation of the approach and avoidance of alcohol questionnaire (AAAQ) in alcohol dependent outpatients. *J. Psychopathol. Behav. Assess.* 29, 231–240. doi: 10.1007/s10862-007-9044-2
- Klein, A. A. (2007). Suppression-induced hyperaccessibility of thoughts in abstinent alcoholics: A preliminary investigation. *Behav. Res. Ther.* 45, 169–177. doi: 10.1016/j.brat.2005.12.012
- Koster, E. H. W., Crombez, G., Verschuere, B., and De Houwer, J. (2004). Selective attention to threat in the dot probe paradigm: differentiating vigilance and difficulty to disengage. *Behav. Res. Ther.* 42, 1183–1192. doi: 10.1016/j.brat.2003.08.001
- Lang, P. J., Bradley, M. M., and Cuthbert, B. N. (2008). *International Affective Picture System (IAPS). Affective Ratings of Pictures and Instruction Manual*. Technical Report A-8. Gainesville, FL: University of Florida.
- Lee, S., Cho, S., and Lee, J. H. (2014). Approach-avoidance pattern of visual attention in hazardous drinkers with ambivalence. *Addict. Behav.* 39, 669–676. doi: 10.1016/j.addbeh.2013.12.001
- McEvoy, P. M., Stritzke, W. G. K., French, D. J., Lang, A. R., and Ketterman, R. L. (2004). Comparison of three models of alcohol craving in young adults: a cross-validation. *Addiction* 99, 482–497. doi: 10.1111/j.1360-0443.2004.00714.x
- Millisecond Software. (2006). *Inquisit (version 2)*. Seattle, WA: Millisecond Software.
- Moss, A. C., Erskine, J. A. K., Albery, I. P., Allen, J. R., and Georgiou, G. J. (2015). To suppress, or not to suppress? That is repression: controlling intrusive thoughts in addictive behaviour. *Addict. Behav.* 44, 65–70. doi: 10.1016/j.addbeh.2015.01.029
- Noël, X., Colmant, M., Van Der Linden, M., Bechara, A., Bullens, Q., Hanak, C., et al. (2006). Time course of attention for alcohol cues in abstinent alcoholic patients: the role of initial orienting. *Alcohol. Clin. Exp. Res.* 30, 1871–1877. doi: 10.1111/j.1530-0277.2006.00224.x
- Palfai, T. P., Monti, P. M., Colby, S. M., and Rohsenow, D. J. (1997). Effects of suppressing the urge to drink on the accessibility of alcohol outcome expectancies. *Behav. Res. Ther.* 35, 59–65. doi: 10.1016/S0005-7967(96)00079-4
- Peirce, J. W. (2007). PsychoPy - Psychophysics software in Python. *J. Neurosci. Methods* 162, 8–13. doi: 10.1016/j.jneumeth.2006.11.017
- Roefs, A., Quaedackers, L., Werrij, M. Q., Wolters, G., Havermans, R., Nederkoorn, C., et al. (2006). The environment influences whether high-fat foods are associated with palatable or with unhealthy. *Behav. Res. Ther.* 44, 715–736. doi: 10.1016/j.brat.2005.05.007
- Saunders, J. B., Aasland, O. G., Babor, T. F., de la Fuente, J. R., and Grant, M. (1993). Development of the alcohol use disorders identification test (AUDIT): WHO collaborative project on early detection of persons with harmful alcohol consumption II. *Addiction* 88, 791–804. doi: 10.1111/j.1360-0443.1993.tb02093.x
- Schlauch, R. C., Breiner, M. J., Stasiewicz, P. R., Christensen, R. L., and Lang, A. R. (2013a). Women inmate substance abusers' reactivity to visual alcohol, cigarette, marijuana, and crack-cocaine cues: approach and avoidance as separate dimensions of reactivity. *J. Psychopathol. Behav. Assess.* 35, 45–56. doi: 10.1007/s10862-012-9313-6
- Schlauch, R. C., Gwynn-Shapiro, D., Stasiewicz, P. R., Molnar, D. S., and Lang, A. R. (2013b). Affect and craving: Positive and negative affect are differentially associated with approach and avoidance inclinations. *Addict. Behav.* 38, 1970–1979. doi: 10.1016/j.addbeh.2012.12.003
- Schlauch, R. C., Levitt, A., Bradizza, C. M., Stasiewicz, P. R., Lucke, J. F., Maisto, S. A., et al. (2013c). Alcohol craving in patients diagnosed with a severe mental illness and alcohol use disorder: Bidirectional relationships between approach and avoidance inclinations and drinking. *J. Consult. Clin. Psychol.* 81, 1087–1099. doi: 10.1037/a0033914
- Schlauch, R. C., Stasiewicz, P. R., Bradizza, C. M., Coffey, S. F., Gulliver, S. B., and Gudleski, G. D. (2012). Relationship between approach and avoidance inclinations to use alcohol and treatment outcomes. *Addict. Behav.* 37, 824–830. doi: 10.1016/j.addbeh.2012.03.010
- Sharbanee, J. M., Stritzke, W. G. K., Wiers, R. W., and Macleod, C. (2013a). Alcohol-related biases in selective attention and action tendency make distinct contributions to dysregulated drinking behaviour. *Addiction* 108, 1758–1766. doi: 10.1111/add.12256
- Sharbanee, J. M., Stritzke, W. G. K., Wiers, R. W., Young, P., Rinck, M., and Macleod, C. (2013b). The interaction of approach-alcohol action tendencies, working memory capacity, and current task goals predicts the inability to regulate drinking behavior. *Psychol. Addict. Behav.* 27, 649–661. doi: 10.1037/a0029982
- Sobell, L. C., and Sobell, M. B. (1992). "Timeline Follow-back: a technique for assessing self-reported ethanol consumption," in *Measuring Alcohol Consumption: Psychosocial and Biological Methods*, eds J. Allen and R. Z. Litten (Totowa, NJ: Humana Press), 41–72. doi: 10.1007/978-1-4612-0357-5_3
- Spruyt, A., De Houwer, J., Tibboel, H., Verschuere, B., Crombez, G., Verbanck, P., et al. (2013). On the predictive validity of automatically activated approach/avoidance tendencies in abstaining alcohol-dependent patients. *Drug Alcohol Depend.* 127, 81–86. doi: 10.1016/j.drugalcdep.2012.06.019
- Stacy, A. W., and Wiers, R. W. (2010). Implicit cognition and addiction: a tool for explaining paradoxical behavior. *Annu. Rev. Clin. Psychol.* 6, 551–575. doi: 10.1146/annurev.clinpsy.121208.131444
- Stormark, K. M., Field, N. P., Hugdahl, K., and Horowitz, M. (1997). Selective processing of visual alcohol cues in abstinent alcoholics: an approach-avoidance conflict? *Addict. Behav.* 22, 509–519. doi: 10.1016/S0306-4603(96)00051-2
- Townshend, J. M., and Duka, T. (2001). Attentional bias associated with alcohol cues: differences between heavy and occasional social drinkers. *Psychopharmacology* 157, 67–74. doi: 10.1007/s002130100764
- Townshend, J. M., and Duka, T. (2007). Avoidance of alcohol-related stimuli in alcohol-dependent inpatients. *Alcohol. Clin. Exp. Res.* 31, 1349–1357. doi: 10.1111/j.1530-0277.2007.00429.x
- Vollstädt-Klein, S., Loeber, S., Von Der Goltz, C., Mann, K., and Kiefer, F. (2009). Avoidance of alcohol-related stimuli increases during the early stage of abstinence in alcohol-dependent patients. *Alcohol Alcohol.* 44, 458–463. doi: 10.1093/alcac/agg056

- Watson, D., Clark, L. A., and Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: the PANAS scales. *J. Pers. Soc. Psychol.* 54, 1063–1070. doi: 10.1037/0022-3514.54.6.1063
- Wenzlaff, R. M., and Wegner, D. M. (2000). Thought suppression. *Annu. Rev. Psychol.* 51, 59–91. doi: 10.1146/annurev.psych.51.1.59
- Wiers, R. W., Rinck, M., Dictus, M., and van den Wildenberg, E. (2009). Relatively strong automatic appetitive action-tendencies in male carriers of the OPRM1 G-allele. *Genes Brain Behav.* 8, 101–106. doi: 10.1111/j.1601-183X.2008.00454.x

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2015 Di Lemma, Dickson, Jedras, Roefs and Field. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Embracing comorbidity: a way toward understanding the role of motivational and control processes in cannabis use disorders

Janna Cousijn *

Departments of Developmental and Experimental Psychology, Utrecht University, Utrecht, Netherlands

Keywords: cannabis use disorders, depression, anxiety, comorbidity, motivation, control

Background

Although the general public perceives cannabis as one of the less harmful illicit drugs, the past decades saw a surge in treatment demands for CUDs (UNODC, 2014). Cannabis nowadays is the primary illicit drug of concern in drug treatment services across North America, Oceania and Africa (UNODC, 2014). The low perceived harms of cannabis use are reflected in the small number of studies investigating the neurocognitive processes underlying CUDs [e.g., only 3 published functional Magnetic Resonance Imaging (fMRI) studies in individuals with a diagnosed CUD compared to controls, contrasting more than 1000 studies in individuals with an Alcohol Use Disorder]. Most studies on the mechanisms underlying cannabis abuse, including my own, investigated heterogeneous groups of chronic or heavy cannabis users with various levels of cannabis use related problems, not groups with diagnosed CUDs.

Even though a substantial part of regular cannabis users will not experience any clear negative social and health consequences of cannabis, this does not imply that CUDs are less severe than other Substance Use Disorders (SUDs). The mental health issues associated with CUDs are substantial and often include comorbid psychiatric disorders including depression and anxiety (Stinson et al., 2006). Moreover, CUDs are difficult to treat and long-term abstinence is achieved by fewer than 20% (Danovitch and Gorelick, 2012). This urgently calls for a better understanding of CUDs. It is therefore time to reach out to those coping with CUDs by studying the mechanisms underneath. The goal of this opinion article is twofold: First, I want to address the strong need for neurocognitive studies in CUDs. Second, I propose that studying neurocognitive commonalities and differences between CUDs and comorbid disorders like depression and anxiety has great potential to unravel the mechanisms underlying CUDs and to eventually reveal new treatment targets.

Motivational and Control Processes in Cannabis Use Disorders

Strong motivations towards drug use (e.g., craving, automatic tendencies to attend to and approach the drug), paired with an insufficient capacity to keep these under control are thought to play a prominent role in SUDs (Goldstein and Volkow, 2002; Robinson and Berridge, 2003; Dawe and Loxton, 2004; Everitt and Robbins, 2005; Wiers et al., 2007; Verdejo-Garcia and Bechara, 2009). Recent behavioral studies suggest that this is also the case in CUDs: confrontation with cannabis or related objects and contexts (i.e., cues) can trigger craving (e.g., Gray et al., 2011; Lundahl and Johanson, 2011), capture their attention (attentional bias; e.g., Cousijn et al., 2013b; Asmaro et al., 2014), and activate approach tendencies (approach bias; e.g., Field et al., 2006; Cousijn et al., 2011). In addition, cognitive control-related functions like planning, organizing, problem

OPEN ACCESS

Edited by:

Frank Ryan,
Imperial College London, UK

Reviewed by:

Anne Marije Kaag,
University of Amsterdam, Netherlands

*Correspondence:

Janna Cousijn,
j.cousijn@gmail.com

Specialty section:

This article was submitted to
Psychology for Clinical Settings,
a section of the journal
Frontiers in Psychology

Received: 06 February 2015

Accepted: 08 May 2015

Published: 27 May 2015

Citation:

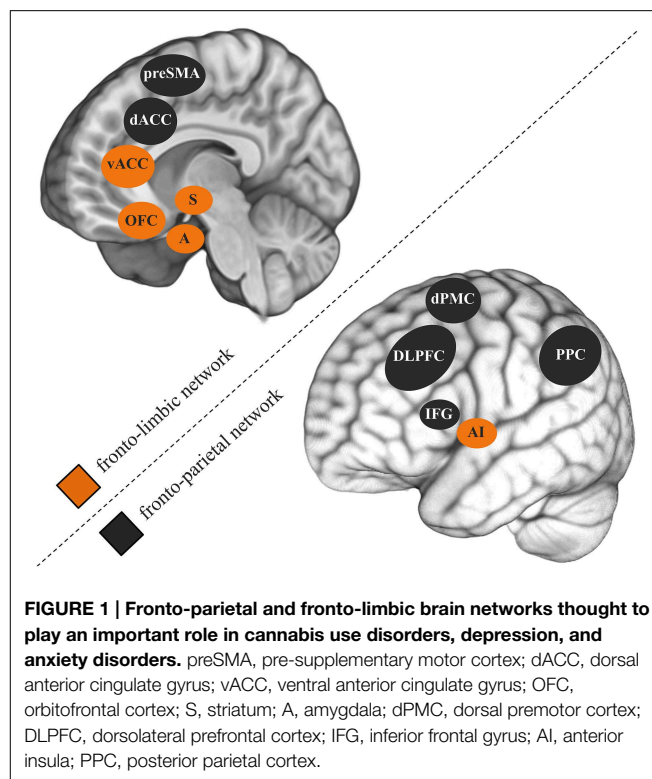
Cousijn J (2015) Embracing
comorbidity: a way toward
understanding the role of motivational
and control processes in cannabis use
disorders. *Front. Psychol.* 6:677.
doi: 10.3389/fpsyg.2015.00677

solving, decision-making, and working-memory appear to be impaired in individuals with a CUD (Fernandez-Serrano et al., 2011). Chronic cannabis exposure may (temporarily) impair cognitive control, but cognitive control deficits may also be a risk factor for the onset of cannabis use and escalation into CUDs (Cousijn et al., 2014).

Embracing Comorbidity as a Tool

While the comorbidity between CUDs and other psychiatric disorders is widely accepted, neurocognitive studies mostly study disorders in isolation. Comorbid symptoms are often even controlled for by excluding such participants. A 3-year longitudinal epidemiological study specifically investigated the role of mental health factors in non-dependent versus dependent heavy cannabis use (Van Der Pol et al., 2013). Although externalizing psychiatric disorders like ADHD and conduct disorder were common to non-dependent and dependent users, internalizing psychiatric disorders such as mood and anxiety disorders were uniquely associated with dependence. Combined, the 283 almost daily cannabis users that participated in my previous studies revealed a correlation of $r = 0.50$ between cannabis use-related problems and depression symptoms (e.g., Cousijn et al., 2012; Beraha et al., 2013; Cousijn et al., 2013a,b). Similarly as in SUDs, neurocognitive models of depression (Weir et al., 2012) and anxiety disorders (Bruhl et al., 2014) stress the importance of dyscontrol over motivational processes and abnormal functioning of the underlying brain systems in the emergence of these disorders. Fronto-parietal and fronto-limbic brain networks are thought to play a key role in this (Figure 1; Seeley et al., 2007). The fronto-parietal network is thereby the main substrate for relatively *cold* executive control (e.g., working memory, attention, inhibition). The fronto-limbic network is primarily involved in emotion regulation, salience attribution and the integration of motivational information (e.g., reward, emotions) into decision processes.

The overlap in neurocognitive mechanisms underlying SUDs with depression and anxiety appears evident. Little is known, however, about why certain symptoms cluster together and what differentiates disorders. From a clinical perspective, the vague boundaries between psychiatric disorders, the heterogeneity in psychiatric problems within patient groups and the poor treatment response in a substantial number of patients also underline the need to look beyond dichotomous disorder classifications (Casey et al., 2013). The new edition of the diagnostic statistical manual (DSM-V) introduced stages of disorder severity but still relies on self-reports (American Psychiatric Association, 2013). In the quest to identify more objective biomarkers for psychiatric problems, the US National Institute of Mental Health recently called for a transdiagnostic dimensional approach in the study of psychiatric disorders, in which the neurobiology underlying symptom dimensions is central, not the disorder classification itself (Casey et al., 2013). Embracing comorbid psychiatric problems, rather than factoring them out in neurocognitive studies, is an important step in this and I believe that such an approach has great potential to advance



our knowledge of psychiatric disorders, including CUDs. An additional advantage of such an approach is that participants with comorbid problems are more representative of individuals with (sub-threshold) psychiatric problems in the general population and of patients in treatment.

Studying the common and unique neurocognitive mechanisms underlying psychiatric disorders may help us to identify new biomarker that could advance prevention and treatment. In the case of CUDs, we can only speculate about the neurocognitive mechanisms underlying CUDs, let alone understand why depression and anxiety disorders are associated with CUDs. Cognitive control deficits and malfunctioning of the underlying fronto-parietal brain networks may be shared between all three disorders, posing a general risk factor for the development of CUDs, depression and anxiety disorders (Koob and Volkow, 2010; Weir et al., 2012; Cousijn et al., 2013b; Bruhl et al., 2014; Peterson et al., 2014). In contrast, motivational processes within specific emotional and rewarding contexts may differentiate disorders. Although abnormal approach-avoidance behavior is common to all three disorders, depression and anxiety are associated with overactive avoidance of certain social and emotional situations (Trew, 2011; Caouette and Guyer, 2014), whereas CUDs may be associated with overactive approach of cannabis cues (Cousijn et al., 2011). Moreover, SUDs including CUDs and depression are both characterized by low positive affect (anhedonia) and abnormal reward responsiveness within various fronto-limbic brain areas (Koob and Volkow, 2010; Hatzigiakoumis et al., 2011; Elman et al., 2013; Morgan et al., 2013; Telzer et al., 2014). Interestingly, a recent PET study among 14 heavy cannabis users showed a link between

anhedonia and reduced dopamine transmission in the striatum (Bloomfield et al., 2014). Unlike, CUDs, depression and anxiety further show abnormal processing of social emotional stimuli in the amygdala (Burghy et al., 2012; Weir et al., 2012; Caouette and Guyer, 2014). Further, amygdala connectivity with the ventral medial prefrontal cortex may differentiate between anxiety and depression by uniquely contributing to certain symptoms (McClure et al., 2007; Beesdo et al., 2009; Burghy et al., 2012).

Genetics are also known to play an important role in the risk for CUDs, depression and anxiety disorders. Motivational and control processes are influenced by genetic factors, including genes involved in drug metabolism and neurotransmission (Sweitzer et al., 2012). Motivational and control processes may thereby, at least partly, mediate the genetic vulnerability to all three disorders. For example the D2 dopamine receptor gene (DRD2) Taq1 A polymorphism affects dopamine binding in the striatum and is consistently associated with SUDs, depression and anxiety disorders (gene-disorder association indices retrieved from Gene Prospector; Yu et al., 2008). The A1 allele of the DRD2 Taq1 A polymorphism has been linked to reduced dopamine D2 receptor availability in the striatum, which could in turn reduce general reward responsiveness (Belcher et al., 2014). Another polymorphism consistently associated with all three disorders is COMTval158met (Yu et al., 2008). The COMT gene encodes an enzyme that is involved in the inactivation of catecholamine neurotransmitters like dopamine, epinephrine, and norepinephrine. The COMTval158met polymorphism has been linked to altered dopamine signaling in the prefrontal cortex, thereby influencing cognitive control (Bruder et al., 2005). Important to note CUDs, depression and anxiety disorders are polygenetic. Single genes are often only weakly associated with the risk for certain psychiatric disorders. To investigate genetic factors underlying polygenetic disorders large-scale multicenter genome-wide studies are needed. To allow DNA data contribution of small studies to large-scale consortia, DNA data collection should be facilitated for new studies, even though the primary objectives do not necessarily comprise genetics. Moreover, epigenetics should be considered, that is the processes involved in long-term changes in gene expression that are heritable to daughter cells. Interestingly, a recent study in mice showed that a single epigenetic mechanisms (histone methylation of *fosb*) can influenced gene expression in the nucleus accumbens

and induce depression and addiction like behavior (Heller et al., 2014).

A Critical Note

Although knowledge of the common and unique neurobiology underlying comorbid disorders could identify biomarkers, researchers and clinicians should carefully evaluate and compare the clinical value of such measures for the individual patient. Our group-based findings may not necessarily translate to the individual. Also, neuroimaging techniques are expensive compared to questionnaires and neuropsychological tests. It is therefore important to explicitly test if certain neural indices explain unique variance on top of simpler (and cheaper) methods.

Conclusions

The worldwide high treatment demands for CUDs, but the significant lack of studies investigating it warrant new studies that investigate neurocognitive functions in cannabis users with a clinically diagnosed CUDs. Uncovering the common and unique neurocognitive mechanisms and associated (epi)genetics underlying CUDs and highly comorbid disorders like depression and anxiety can provide valuable knowledge for improving current state-of-the-art treatments and for developing new neuroscience based interventions, such as neurocognitive training (e.g., approach-action retraining; Wiers et al., 2011), neuromodulation (e.g., stimulating brain areas involved in control; Ressler and Mayberg, 2007; Berlim et al., 2013; Da Silva et al., 2013) and pharmacotherapy (e.g., medication that enhances emotion regulation; Ressler and Mayberg, 2007; Sofuoglu, 2010; Mohler, 2012; Farb and Ratner, 2014). I reiterate that it is vital to study motivational processes and cognitive control in ecologically valid groups of individuals, that is, by including those coping with comorbid psychiatric problems.

Acknowledgments

JC is supported by The Consortium Individual Development (CID). CID is funded through the Gravitation program of the Dutch Ministry of Education, Culture, and Science and the Netherlands Organization for Scientific Research (NWO grant number 024.001.003 awarded to Chantal Kemner, Utrecht University, The Netherlands).

References

- American Psychiatric Association. (2013). *Diagnostic and Statistical Manual of Mental Disorders, 5th Edn.* (DSM-5). Washington, DC: American Psychiatric Publishing.
- Asmaro, D., Carolan, P. L., and Liotti, M. (2014). Electrophysiological evidence of early attentional bias to drug-related pictures in chronic cannabis users. *Addict. Behav.* 39, 114–121. doi: 10.1016/j.addbeh.2013.09.012
- Beesdo, K., Lau, J. Y., Guyer, A. E., McClure-Tone, E. B., Monk, C. S., Nelson, E. E., et al. (2009). Common and distinct amygdala-function perturbations in depressed vs anxious adolescents. *Arch. Gen. Psychiatry* 66, 275–285. doi: 10.1001/archgenpsychiatry.2008.545
- Belcher, A. M., Volkow, N. D., Moeller, F. G., and Ferre, S. (2014). Personality traits and vulnerability or resilience to substance use disorders. *Trends Cogn. Sci.* 18, 211–217. doi: 10.1016/j.tics.2014.01.010
- Beraha, E. M., Cousijn, J., Hermanides, E., Goudriaan, A. E., and Wiers, R. W. (2013). Implicit associations and explicit expectancies toward cannabis in heavy cannabis users and controls. *Front. Psychiatry* 4:59. doi: 10.3389/fpsy.2013.00059
- Berlim, M. T., Van Den Eynde, F., and Daskalakis, Z. J. (2013). Clinical utility of transcranial direct current stimulation (tDCS) for treating major depression: a systematic review and meta-analysis of randomized, double-blind and sham-controlled trials. *J. Psych. Res.* 47, 1–7. doi: 10.1016/j.jpsychires.2012.09.025

- Bloomfield, M. A., Morgan, C. J., Kapur, S., Curran, H. V., and Howes, O. D. (2014). The link between dopamine function and apathy in cannabis users: an [18F]-DOPA PET imaging study. *Psychopharmacology (Berl)*. 231, 2251–2259. doi: 10.1007/s00213-014-3523-4
- Bruder, G. E., Keilp, J. G., Xu, H., Shikhman, M., Schori, E., Gorman, J. M., et al. (2005). Catechol-O-methyltransferase (COMT) genotypes and working memory: associations with differing cognitive operations. *Biol. Psychiatry* 58, 901–907. doi: 10.1016/j.biopsych.2005.05.010
- Bruhl, A. B., Delsignore, A., Komossa, K., and Weidt, S. (2014). Neuroimaging in social anxiety disorder-A meta-analytic review resulting in a new neurofunctional model. *Neurosci. Biobehav. Rev.* 47c, 260–280. doi: 10.1016/j.neubiorev.2014.08.003
- Burghy, C. A., Stodola, D. E., Ruttle, P. L., Molloy, E. K., Armstrong, J. M., Oler, J. A., et al. (2012). Developmental pathways to amygdala-prefrontal function and internalizing symptoms in adolescence. *Nat. Neurosci.* 15, 1736–1741. doi: 10.1038/nn.3257
- Caouette, J. D., and Guyer, A. E. (2014). Gaining insight into adolescent vulnerability for social anxiety from developmental cognitive neuroscience. *Dev. Cogn. Neurosci.* 8, 65–76. doi: 10.1016/j.dcn.2013.10.003
- Casey, B. J., Craddock, N., Cuthbert, B. N., Hyman, S. E., Lee, F. S., and Ressler, K. J. (2013). DSM-5 and RDoC: progress in psychiatry research? *Nat. Rev. Neurosci.* 14, 810–814. doi: 10.1038/nrn3621
- Cousijn, J., Goudriaan, A. E., Ridderinkhof, K. R., Van Den Brink, W., Veltman, D. J., and Wiers, R. W. (2012). Approach-bias predicts development of cannabis problem severity in heavy cannabis users: results from a prospective fMRI study. *PLoS ONE* 7:e42394. doi: 10.1371/journal.pone.0042394
- Cousijn, J., Goudriaan, A. E., and Wiers, R. W. (2011). Reaching out towards cannabis: approach-bias in heavy cannabis users predicts changes in cannabis use. *Addiction* 106, 1667–1674. doi: 10.1111/j.1360-0443.2011.03475.x
- Cousijn, J., Snoek, R. W., and Wiers, R. W. (2013a). Cannabis intoxication inhibits avoidance action tendencies: a field study in the Amsterdam coffee shops. *Psychopharmacology (Berl)*. 229, 167–176. doi: 10.1007/s00213-013-3097-6
- Cousijn, J., Watson, P., Koenders, L., Vingerhoets, W. A., Goudriaan, A. E., and Wiers, R. W. (2013b). Cannabis dependence, cognitive control and attentional bias for cannabis words. *Addict. Behav.* 38, 2825–2832. doi: 10.1016/j.addbeh.2013.08.011
- Cousijn, J., Wiers, R. W., Ridderinkhof, K. R., Van Den Brink, W., Veltman, D. J., and Goudriaan, A. E. (2014). Effect of baseline cannabis use and working-memory network function on changes in cannabis use in heavy cannabis users: a prospective fMRI study. *Hum. Brain Mapp.* 35, 2470–2482. doi: 10.1002/hbm.22342
- Danovitch, I., and Gorelick, D. A. (2012). State of the art treatments for cannabis dependence. *Psychiatr. Clin. North Am.* 35, 309–326. doi: 10.1016/j.psc.2012.03.003
- Da Silva, M. C., Conti, C. L., Klauss, J., Alves, L. G., Do Nascimento Cavalcante, H. M., Fregni, F., et al. (2013). Behavioral effects of transcranial Direct Current Stimulation (tDCS) induced dorsolateral prefrontal cortex plasticity in alcohol dependence. *J. Physiol. Paris* 107, 493–502. doi: 10.1016/j.jphysparis.2013.07.003
- Dawe, S., and Loxton, N. J. (2004). The role of impulsivity in the development of substance use and eating disorders. *Neurosci. Biobehav. Rev.* 28, 343–351. doi: 10.1016/j.neubiorev.2004.03.007
- Elman, I., Borsook, D., and Volkow, N. D. (2013). Pain and suicidality: insights from reward and addiction neuroscience. *Prog. Neurobiol.* 109, 1–27. doi: 10.1016/j.pneurobio.2013.06.003
- Everitt, B. J., and Robbins, T. W. (2005). Neural systems of reinforcement for drug addiction: from actions to habits to compulsion. *Nat. Neurosci.* 8, 1481–1489. doi: 10.1038/nn1579
- Farb, D. H., and Ratner, M. H. (2014). Targeting the modulation of neural circuitry for the treatment of anxiety disorders. *Pharmacol. Rev.* 66, 1002–1032. doi: 10.1124/pr.114.009126
- Fernandez-Serrano, M. J., Perez-Garcia, M., and Verdejo-Garcia, A. (2011). What are the specific vs. generalized effects of drugs of abuse on neuropsychological performance? *Neurosci. Biobehav. Rev.* 35, 377–406. doi: 10.1016/j.neubiorev.2010.04.008
- Field, M., Eastwood, B., Bradley, B. P., and Mogg, K. (2006). Selective processing of cannabis cues in regular cannabis users. *Drug Alcohol Depend.* 85, 75–82. doi: 10.1016/j.drugalcdep.2006.03.018
- Goldstein, R. Z., and Volkow, N. D. (2002). Drug addiction and its underlying neurobiological basis: neuroimaging evidence for the involvement of the frontal cortex. *Am. J. Psychiatry* 159, 1642–1652. doi: 10.1176/appi.ajp.159.10.1642
- Gray, K. M., Larowe, S. D., Watson, N. L., and Carpenter, M. J. (2011). Reactivity to in vivo marijuana cues among cannabis-dependent adolescents. *Addict. Behav.* 36, 140–143. doi: 10.1016/j.addbeh.2010.08.021
- Hatzigiakoumis, D. S., Martinotti, G., Giannantonio, M. D., and Janiri, L. (2011). Anhedonia and substance dependence: clinical correlates and treatment options. *Front. Psychiatry* 2:10. doi: 10.3389/fpsy.2011.00010
- Heller, E. A., Cates, H. M., Pena, C. J., Sun, H., Shao, N., Feng, J., et al. (2014). Locus-specific epigenetic remodeling controls addiction- and depression-related behaviors. *Nat. Neurosci.* 17, 1720–1727. doi: 10.1038/nn.3871
- Koob, G. F., and Volkow, N. D. (2010). Neurocircuitry of addiction. *Neuropsychopharmacology* 35, 217–238. doi: 10.1038/npp.2009.110
- Lundahl, L. H., and Johanson, C. E. (2011). Cue-induced craving for marijuana in cannabis-dependent adults. *Exp. Clin. Psychopharmacol.* 19, 224–230. doi: 10.1037/a0023030
- McClure, E. B., Adler, A., Monk, C. S., Cameron, J., Smith, S., Nelson, E. E., et al. (2007). fMRI predictors of treatment outcome in pediatric anxiety disorders. *Psychopharmacology (Berl)*. 191, 97–105. doi: 10.1007/s00213-006-0542-9
- Mohler, H. (2012). The GABA system in anxiety and depression and its therapeutic potential. *Neuropharmacology* 62, 42–53. doi: 10.1016/j.neuropharm.2011.08.040
- Morgan, J. K., Olino, T. M., Mcmakin, D. L., Ryan, N. D., and Forbes, E. E. (2013). Neural response to reward as a predictor of increases in depressive symptoms in adolescence. *Neurobiol. Dis.* 52, 66–74. doi: 10.1016/j.nbd.2012.03.039
- Peterson, B. S., Wang, Z., Horga, G., Warner, V., Rutherford, B., Klahr, K. W., et al. (2014). Discriminating risk and resilience endophenotypes from lifetime illness effects in familial major depressive disorder. *JAMA Psychiatry* 71, 136–148. doi: 10.1001/jamapsychiatry.2013.4048
- Ressler, K. J., and Mayberg, H. S. (2007). Targeting abnormal neural circuits in mood and anxiety disorders: from the laboratory to the clinic. *Nat. Neurosci.* 10, 1116–1124. doi: 10.1038/nn1944
- Robinson, T. E., and Berridge, K. C. (2003). Addiction. *Annu. Rev. Psychol.* 54, 25–53. doi: 10.1146/annurev.psych.54.101601.145237
- Seeley, W. W., Menon, V., Schatzberg, A. F., Keller, J., Glover, G. H., Kenna, H., et al. (2007). Dissociable intrinsic connectivity networks for salience processing and executive control. *J. Neurosci.* 27, 2349–2356. doi: 10.1523/JNEUROSCI.5587-06.2007
- Sofuoglu, M. (2010). Cognitive enhancement as a pharmacotherapy target for stimulant addiction. *Addiction* 105, 38–48. doi: 10.1111/j.1360-0443.2009.02791.x
- Stinson, F. S., Ruan, W. J., Pickering, R., and Grant, B. F. (2006). Cannabis use disorders in the USA: prevalence, correlates and co-morbidity. *Psychol. Med.* 36, 1447–1460. doi: 10.1017/S0033291706008361
- Sweitzer, M. M., Donny, E. C., and Hariri, A. R. (2012). Imaging genetics and the neurobiological basis of individual differences in vulnerability to addiction. *Drug Alcohol Depend.* 123(Suppl. 1), S59–S71. doi: 10.1016/j.drugalcdep.2012.01.017
- Telzer, E. H., Fuligni, A. J., Lieberman, M. D., and Galvan, A. (2014). Neural sensitivity to eudaimonic and hedonic rewards differentially predict adolescent depressive symptoms over time. *Proc. Natl. Acad. Sci. U.S.A.* 111, 6600–6605. doi: 10.1073/pnas.1323014111
- Trew, J. L. (2011). Exploring the roles of approach and avoidance in depression: an integrative model. *Clin. Psychol. Rev.* 31, 1156–1168. doi: 10.1016/j.cpr.2011.07.007
- UNODC. (2014). *World Drug Report 2014*. Vienna: United Nations Office on Drugs and Crime, United Nations publication.
- Van Der Pol, P., Liebrechts, N., De Graaf, R., Ten Have, M., Korf, D. J., Van Den Brink, W., et al. (2013). Mental health differences between frequent cannabis users with and without dependence and the general population. *Addiction* 108, 1459–1469. doi: 10.1111/add.12196
- Verdejo-Garcia, A., and Bechara, A. (2009). A somatic marker theory of addiction. *Neuropharmacology* 56(Suppl. 1), 48–62. doi: 10.1016/j.neuropharm.2008.07.035

- Weir, J. M., Zakama, A., and Rao, U. (2012). Developmental risk I: depression and the developing brain. *Child Adolesc. Psychiatr. Clin. N. Am.* 21, 237–259. doi: 10.1016/j.chc.2012.01.004
- Wiers, R. W., Bartholow, B. D., Van Den Wildenberg, E., Thush, C., Engels, R. C., Sher, K. J., et al. (2007). Automatic and controlled processes and the development of addictive behaviors in adolescents: a review and a model. *Pharmacol. Biochem. Behav.* 86, 263–283. doi: 10.1016/j.pbb.2006.09.021
- Wiers, R. W., Eberl, C., Rinck, M., Becker, E. S., and Lindenmeyer, J. (2011). Retraining automatic action tendencies changes alcoholic patients' approach bias for alcohol and improves treatment outcome. *Psychol. Sci.* 22, 490–497. doi: 10.1177/0956797611400615
- Yu, W., Wulf, A., Liu, T., Khoury, M. J., and Gwinn, M. (2008). Gene Prospector: an evidence gateway for evaluating potential susceptibility genes

and interacting risk factors for human diseases. *BMC Bioinformatics* 9:528. doi: 10.1186/1471-2105-9-528

Conflict of Interest Statement: The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2015 Cousijn. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



The computational psychiatry of reward: broken brains or misguided minds?

M. Moutoussis^{1*}, G. W. Story^{1,2} and R. J. Dolan^{1,3}

¹ Wellcome Trust Centre for Neuroimaging, University College London, London, UK, ² Centre for Health Policy, Institute of Global Health Innovation, Imperial College, London, UK, ³ Max Planck UCL Centre for Computational Psychiatry and Ageing Research, University College London, London, UK

OPEN ACCESS

Edited by:

Nikolina Skandali,
University of Cambridge, UK

Reviewed by:

Michelle Dow Keawphalouk,
Harvard University – Massachusetts
Institute of Technology, USA
Stefan Borgwardt,
University of Basel, Switzerland

*Correspondence:

M. Moutoussis,
Wellcome Trust Centre
for Neuroimaging, University College
London, 12 Queen Square,
London WC1N 3BG, UK
m.moutoussis@ucl.ac.uk

Specialty section:

This article was submitted to
Psychology for Clinical Settings,
a section of the journal
Frontiers in Psychology

Received: 26 June 2015

Accepted: 09 September 2015

Published: 29 September 2015

Citation:

Moutoussis M, Story GW
and Dolan RJ (2015)
The computational psychiatry
of reward: broken brains or misguided
minds? *Front. Psychol.* 6:1445.
doi: 10.3389/fpsyg.2015.01445

Research into the biological basis of emotional and motivational disorders is in danger of riding roughshod over a patient-centered psychiatry and falling into the dualist errors of the past, i.e., by treating mind and brain as conceptually distinct. We argue that a psychiatry informed by computational neuroscience, computational psychiatry, can obviate this danger. Through a focus on the reasoning processes by which humans attempt to maximize reward (and minimize punishment), and how such reasoning is expressed neurally, computational psychiatry can render obsolete the polarity between biological and psychosocial conceptions of illness. Here, the term ‘psychological’ comes to refer to information processing performed by biological agents, seen in light of underlying goals. We reflect on the implications of this perspective for a definition of mental disorder, including what is entailed in asserting that a particular disorder is ‘biological’ or ‘psychological’ in origin. We propose that a computational approach assists in understanding the topography of mental disorder, while cautioning that the point at which eccentric reasoning constitutes disorder often remains a matter of cultural judgment.

Keywords: computational psychiatry, dualism, optimality, psychiatric nosology, Bayesian inference

*I'm gonna, I'm gonna lose my baby/So I always keep a bottle near
[The psychiatrist] said, "I just think you're depressed."/This, me, yeah, baby, and the rest.
A. Winehouse (2007), musician who died of alcohol intoxication in 2011*

Introduction

The idea that reward processing is important in emotional and motivational psychiatric disorders comes from a view of the mind-as-decision-maker. This idea has been developed within the nascent field of computational psychiatry, the clinical offshoot of computational neurobiology. Within this framework, ‘psychiatric disorder’ entails a breakdown in the brain’s inability to optimize decisions. Thus, to the extent that good decisions set up the individual to optimally obtain reward, ‘psychiatric disorder’ entails a suboptimal seeking of reward within an environment. As an approach computational psychiatry promises much by way of future diagnostic and therapeutic applications (Huys et al., 2011; Montague et al., 2012).

We are of course mindful that psychiatry has seen many promising directions that have delivered much less than hoped. In this article we argue that computational psychiatry has already

made major contributions in resolving important conceptual divides in mental health. These have been expressed in varying ways but are located around biological/psychological – diagnostic/whole-person polarities (Boyle and Johnstone, 2014; Hayes and Bell, 2014). This has led to a situation where biological research is accused of shocking oversimplification of the mind, and psychosocial research accused of an equally shocking neglect of the brain ('mindlessness vs. brainlessness'). Intimately related is the question of when psychiatric intervention is justified to address mental symptoms¹. Here medical professionals may inappropriately diagnose and prosecute biological interventions (Szasz, 1960), while psychological therapists can be just as disempowering (Dolnick, 1998; Romito, 2008). These splits, like old religious conflicts in Europe, concern resource or power struggles among 'denominations' as much as they concern disagreements of substance (Bentall, 2009). It is important to note that a resolution of the latter, to which our present work contributes, may only make slow inroads into the former.

An unhealthy mind is one disposed to make bad decisions and there is no end of examples in psychiatry. Decisions are not just the sine qua non of overt actions, such as a decision of a patient with depression to stay all day in bed, or drink a lethal quantity of vodka and die. We are also 'deciding' when we believe a proposition such as 'my wife has been replaced by a double,' or believe our senses when they inform that 'I look fat' – as in the body image distortion seen in anorexia – right through to a conclusion that 'the voice is real' in psychosis. Good decisions on the other hand entail those (among others) that lead to a healthy life, maintain safety and successful reproduction. Computational psychiatry goes further, postulating that *healthy organisms take optimal decisions, given their resources*. 'Good decisions' cannot but be those that successfully obtain the 'best reward,' those which are good in life and for life. We can call this the Leibnitz² principle – the best possible world of decision-making is with us. Within this framework, 'psychiatric disorder' entails an inability to optimize decisions. Thus, to the extent that good decisions set up an individual to optimally obtain reward, 'psychiatric disorder' entails a suboptimal reaping of possible reward.

Common sense tells us that motivational and emotional disorders are central to most psychiatric disorders such as drug dependence, clinical depression or schizophrenia. For example, craving is a key motivational disturbance and DSM5 rightly includes it in the 11 criteria of Substance Use Disorder (APA, 2013). In disorders of mood, the symptom of anhedonia is a core criterion of clinical depression, marking it as a motivational and emotional disorder. It is also likely that the fears expressed within persecutory delusions are signs of deeply disordered emotional processing, whereby a diseased brain has recruited basic motivational and emotional mechanisms, originally meant

to warn and protect the individual against dire threat, in a completely unwarranted fashion. It may also seem obvious that psychopaths can be construed as individuals inadequately motivated by the pain of others.

At the same time the study of reward in Psychiatry necessitates a widening of the scope of classic computational neurobiology to take seriously the *subjective experience* of motivational and emotional symptoms. Psychiatry is first and foremost a branch of medicine, not of engineering. Psychiatrists recommend biological, psychological and social interventions first and foremost in order to alleviate the suffering of a patient, and those around the patient. Unlike other disciplines, changing people's behavior is not the final goal but a part – usually a very important part – of restoring health. Conversely, understanding *behavior* motivated by reward and loss is important for psychiatric research. If we were concerned with physical trauma or viral illnesses, a thorough understanding of the body's mechanisms of immunity and tissue repair would be important, while supporting and correcting such processes would constitute practicing medicine. On the one hand, health research strives to understand both the physiology (the healthy function) and the pathophysiology (function-in-illness) of an underlying biological substrate. On the other, the clinician helps people who suffer as best as possible, while neither over- or under- applying their craft, as condensed in the dictum 'only the expert surgeon knows when *not* to operate.' As there is much suffering which medical interventions do not help, much maladaptive behavior and loss-related suffering is within the frame of scientific interest but outside the clinical scope of psychiatry.

Computational psychiatry focuses on those brain-based mechanisms which strive to optimize reward within the environment. We claim that this indivisible coupling of brain-function-environment has already transcended the troublesome polarities of biological vs. psychological, diseased brain vs. maladjusted mind. Furthermore, once an optimizing of function is understood in relation to an individual patient's needs, the approach also transcends polarities of normative vs. libertarian and reductionist vs. anti-scientific psychiatry. If the study of reward-related decision-making is our new analytic tool, then the goal of this article is to clarify how much emotional and motivational disorders might yield to its explanatory power. Here we also consider foreseeable pitfalls in addition to how this new way of seeing disorder transcends the old polarities that still haunt psychiatry.

Methods: Review of the Normative Account

One opportunity that a working hypothesis of optimal reward-seeking, given one's resources, affords is that of normativity. Behavior (be it choice between A and B or free, creative expression) is no longer judged in comparison to a reference sample, a 'healthy control' group with all the limitations this entails. Instead behavior is compared to demonstrably optimal solutions in face-valid but solvable tasks.

¹ A most curious term: a symptom without a mental dimension is not a symptom but a sign.

² Leibnitz claimed that we live in the best world that could possibly (logically, self-consistently) exist. He was famously satirized by Voltaire in 'Candide' [en. wikipedia.org/wiki/Candide], exemplifying the normativity-pathology dialectic that is highly relevant to us.

The Bayesian Approach

In an *uncertain* world, each piece of *information* is used to *update* the person's *beliefs* about the *reality underlying appearances* according to this person's rule-book of *how reality gives rise to appearance*. This is what's called 'Bayesian inference' see **Table 1** for an illustrative toy example.

This toy example does not include decisions about which action to take as yet, only decisions about which state the world is in (here, a self-worth state). Neither motivation nor reward, the central topics of this work are, as yet, explicit.

It is still necessary to write down the stages of information processing leading to normative decisions, and therefore classify where the process may break down in psychiatric disorders. Agents must

- (1) have an adequate repertoire of classes (dimensions or categories) that can describe the environment in which they take decisions. 'Can describe' here means that beliefs about contexts and states, including prior beliefs, are expressed in terms of this repertoire. Does the set {worthless, worthy} form an adequate repertoire?
- (2) have an adequate generative model of what states within the environment can give rise to the observations they make (likelihood of states, including the intentions of other people). Was the rule in the toy example accurate?
- (3) be able to invert the generative model so as to determine what the state of the world and the self is likely to be at any given moment. Assuming that the world-view of the toy example was correct, was the update belief warranted or unwarranted?

We now consider decisions about actions, rather than passive beliefs. If, as computational neuroscience claims, brains seek the best possible decisions then they should have values that they seek to optimize, values which are meaningful even if not explicitly represented.

Utility as Consistent Probability Representation

If the value of different outcomes that can be obtained via different behavioral strategies in a given context is well defined

for an agent, we can call these values the 'utilities' of the different outcomes and map them to the probability of an agent adopting the corresponding strategy. Rewards are outcomes that reinforce human behavior or are reported as appetitive, desirable, hedonic, pleasant by healthy humans. Confronted with known choices A, B, and C an agent will ascribe 'utilities' $u(A)$, $u(B)$, $u(C)$ such that they can choose by applying a well-defined choice probability. For choice A, this would be $\pi(A; u(A), u(B), u(C))$. Here we operationalize the motivational value of an outcome as the relative (but otherwise consistent) probability with which it is chosen. Let's call this 'consistent probability representation' on the part of the agent. The fact that this can be well-defined is a hypothesis with extremely productive consequences, which we presently describe. It is consistent probability representation that makes it possible to construct a full Bayesian Decision-making psychiatry (BDP), (Montague et al., 2012; Huys et al., 2014). In addition to 1–3 above, agents need -

- (4) an adaptive utility function, as just described; and
- (5) a generative model that includes an accurate prediction of which outcomes will follow which decision, given a state of the world and the self (model of control), so that they can choose actions that will impact upon the state of the world to produce outcomes with maximum utility.

We immediately note that consistent probability representation firmly maps utility to probability – which is, in Bayesian terms, just another kind of belief (Friston et al., 2013); where the model of control is just part of the generative model. Thus (4) and (5) are not additions to the Bayesian schema but are special cases of its elements.

'Mental disorder' can be said to exist when this decision-making apparatus itself is impaired, rather than reflecting any issues with its inputs. Note, however, that the decision-making apparatus by virtue of its Bayesian nature accumulates experience. Every updated belief contains the weight of its priors and forms the prior of the next update. Hence an 'impairment' may consist in the development of a decision-making apparatus poorly adapted for the circumstance in question so that there is no firm distinction between maladapted and diseased decision-making apparatus. At the same time there is no guarantee that brain development will not encode posteriors into irreversible structure. As an example, the accent with which one speaks is part of the posteriors about the world encoded in childhood. It is very difficult to learn to pronounce a foreign language like a native in adulthood. Hence with respect to an environment where speaking this new language without a foreign accent is optimal, child development has in this broad definition "damaged" the brain.

We can now illustrate this scheme by locating motivational and emotional problems to distinct parts of this apparatus:

- (1) Development may not have equipped the patient with an adequate repertoire of classes. A traumatic life which has set the prior probability that others will dislike me as equal to one can be seen to correspond to the extreme of Beck's notion of a core belief, which says 'I am worthless.' In the opening

TABLE 1 | A toy example of Bayesian reasoning.

Commonsense term	Example	Bayesian term
Belief before considering information (before thinking, not before an event!)	"I am either 'worthless' or 'worthy' – both are equally likely"	<i>Prior belief</i>
Salient information	"My paper was rejected"	Data (or observation)
Rule – how the world works	The worthless get rejected, the worthy are appreciated.	<i>Generative model</i> of the world, which provides the <i>Likelihood</i> that an underlying state of the world will produce a datum.
Updated belief	"I am worthless"	<i>Posterior belief</i>

quote, “I’m gonna lose my baby” might be an exemplar of such a (prior) certainty.

- (2) The generative model of self and world may be inadequate, leading to wrong estimations of likelihood. For example, one may not have the requisite knowledge that there are common causes of palpitations and shortness of breath other than serious illness, setting the ground for panic anxiety. In the Winehouse quote the dismissal of the psychiatrists’ opinion (‘you’re depressed’) as unlikely to lead to good care may in fact be a sign of such poverty-of-generative-model.
- (3) The person may simply be cognitively impaired, so that they can’t work backward from observations to the underlying reality (technically, ‘model inversion’).

As above, focusing at reward contingent on actions yields two further potential problem areas:

- (4) They may believe that no decisions are available to them that are associated with dependably good outcomes (e.g., learnt helplessness, OCD).
- (5) They may attach too much utility to certain components of a decision (e.g., relief of negative affect by ‘keeping a bottle near’) rather than to others. In the drinking example, if the value of the awful state that will ensue once the drug wears off is discounted, the sufferer heads for a vicious cycle.

However, not all is perfect with utility-maximizing Bayesian schema: human beings appear to violate systematically³ the hypothesis of ‘consistent probability representation.’ We now briefly consider an alternative proposal of how people may represent their preferences, which offers a potential solution to these violations.

Bussemeyer: Preference for Reward as a Mixture State

Suppose that decision-making is probed first according to one of three options, for example by asking how much one prefers A over (B or C). We then probe how much B is preferred over C. It turns out that the choice probabilities⁴ (and implied utilities) experimentally measured are not consistent with performing the same experiment in the alternative possible orders. This violation of consistent probability representation – here an order effect – is one of several apparent inconsistencies in probabilistic reasoning that people display. Various explanations have been put forward, ranging from an erroneous bias to invocation of specialized context effects.

But what if before *making* the choice between A, B and C a person does not in fact encode all decision probabilities $\pi(A; u(A), u(B), u(C))$ etc. ? It may be that their psychological and neurobiological state is better described as the subject being in two or more minds, in a so-called “mixture state” $s = (a(A), a(B), a(C))$ where a are the amplitudes of the mixture. In this scenario the process of *making the choice* is implemented

³Though not grossly.

⁴Consistent choice utilities map to consistent choice probabilities. In a Bayesian world we would require such consistency of utility. In reality the decision-making processes people use mean that this isn’t so. This is called ‘violation of revealed preference theory.’

as a reduction or projection of the mixture state. The rigorous formalism used to describe the dynamics of mixture states, and what happens at the point of reduction, was first described in quantum physics and recently introduced into decision making by Bussemeyer et al. (2009). The reduction or projection process naturally produces order effects: essentially, enquiring about A automatically affects B and C, and so on. The framework is referred to as ‘quantum probability’ (QP) – an unfortunate term in the current context as no physical quanta are involved at all.

The consequences of this framework have not been worked out nearly as fully as, for example, a Bayesian framework. We suggest that the experimental evidence supporting it casts some doubt on the fundamental idea that people represent preference probabilities corresponding to a well defined utility function. There are, however, many cases in which the two frameworks concur (Bussemeyer and Bruza, 2012) and so in this instance we proceed with the better-worked-out Bayesian framework, mindful that there are good reasons why its assumptions might provide poor approximations of psychology.

Marr: Process Models in the Brain

So far we have considered reward and emotion at the level of information processing. Computational psychiatry, however, is not just behavioral economics or behaviorist psychology. Following Marr (1982), we seek evidence that specific, normative, computations we hypothesize are instantiated in neural wetware. This in turn raises the thorny issue that a specific computation – in the sense of a specific normative solution – can be achieved with different problem-solving techniques. It is the signatures of these algorithms that we look for in the neural substrate, and the complete account – from stimulus to neural response, to neural computation to its representation in experimental data – is the ‘process model.’ The best-established process models relevant to the computational psychiatry of reward are arguably those that posit the basal ganglia as representing reward-based learning prediction errors (Seymour et al., 2004) and of the ventral and medial prefrontal cortex representing the values of different actions available to the subject (Rushworth et al., 2011).

Modeling Motivation and Emotion

There is one issue in motivational and emotional research which has been relatively overlooked within the framework of reward processing, and if unaddressed might reinforce dualist splits. The working definition of motivation within the Bayesian framework appears to claim something trivial, namely defining the motivational power of an outcome as the frequency with which it is chosen. We can choose to call this ‘motivation,’ and this is fine if we were talking about math or physics, where there is no danger of confusing a rigorously defined quantity, say the charge of a quark, with a property of the mind. However, here we are *also* talking about motivation as experienced by patients, so we need to be clear about what sort of claim we are making about the semantic referents to which the term ‘motivation’ belongs. More specifically, are we claiming that the choice-frequency definition of ‘motivation’ is to be taken for granted, while the phenomenal experience of ‘motivation’ is a subject for a future, maybe more optional, clarification or research? This

would constitute a linguistic coup d'état! The hard problem of consciousness need not concern us here: we only need to avoid dualism and – like good Bayesians – optimally combine both linguistic and decision-behavioral evidence.

People place great importance in the distinction between 'I can't' and 'I don't care.' 'He doesn't care about me' is a much more serious accusation than 'he can't understand me.' Yet our measurement of motivation as the currency between observable outcomes and decision probability often makes this distinction quite difficult. Suppose button A gives me a piece of jellied eel four times out of 10, and button B six times out of 10. If I prefer them equally, is it that I am very good at working out frequencies but I don't care about jellied eel (no motivation), or that I'm very keen on eel but I am incapable of working out frequencies (no ability)? Similarly, if task performance depends on some other psychiatric variable (say on anxiety) we could easily confuse performance at the left side of the Yerkes–Dodson curve (arousal and motivation too low) with performance on the right (high motivation, but arousal detrimentally high). It is not, of course, impossible to distinguish between the 'I can't' and 'I don't care' but ideally both phenomenological and behavioral enquiry are needed. It is interesting to note that the individual's 'I can't' may be the genetic pool's 'I haven't learnt to appreciate.'

Models traditionally address the issue of motivation-per-outcome by fitting a single parameter (often called 'temperature') for each agent. More recently models have parametrized two different aspects of how motivating reward are, even before considering the phenomenological level. The first relates to how often a choice would be made if the reward emanating from it were immediately obtained with great certainty. Even an obviously preferable outcome ('do you want £5 or £0?') may not be chosen 100% of the time due, for example, to lapses in attention/misunderstanding. The second aspect has to do with how motivation to make a decision changes as the outcomes of these decisions are, with time, more reliably inferred. This can be seen as an 'motivational exchange rate' or 'decision temperature' pertaining to a unit change of outcome away from the point of indifference. This pair of concepts is codified as 'lapse rate and inverse temperature' in the classic RL temperature (Guitart-Masip et al., 2012) and 'goal priors and action precision' in an active inference framework (Friston et al., 2013). Note these are not just different names for the same variables and although they refer to related concepts they have subtly different computational roles.

Although we have a working definition of motivation, we have less of a handle on the term emotion. Our *implied* definition of emotion: a positive or negative utility attaches a value upon the outcomes with which it is associated, and thus upon the states and decisions that lead to them, corresponding to more positive or negative emotional states respectively. Emotion contains *as inseparable parts of each unitary phenomenological state* not only valence and magnitude but rich information about context, intention etc. The desire for sex and the desire for knowledge are not just differently tagged emotions, they are different emotions.

At the moment the way that researchers relate computational variables relate to emotions (if at all) is haphazard; yet tentative progress is being made. In one path breaking study, Rutledge

et al. (2014) related changes in subjective well-being to several aspects of a participants' reward – such as their cumulative reward ('wealth'), immediate reward and most importantly *immediate reward compared to expectations* – their reward prediction error (RPE). Here changes in subjective wellbeing, 'how happy do you feel at the moment,' were best predicted by RPEs. In a bold formulation, Joffily and Coricelli (2013) posited that the phenomenology of several emotions, not just the single dimension of higher vs. lower wellbeing, is intimately linked to *both the temporal dynamics and the certainty* of the beliefs about how one's state evolves relative to one's goals or desires. Thus not only does a person feel 'positive' as their beliefs shift toward a desired state (as a positive RPE would entail); but this positive emotion has the color of happiness if the current belief is certain but the color of hope if the corresponding belief is uncertain.

This experimental and theoretical progress attests to the feasibility of unifying the 'client' (subjective) and the 'decision-maker' (objective) perspectives on emotion. The links between the dynamics of reward and the dynamics of emotion show great promise and need a lot of experimental testing, but the first steps of clinical importance have been taken.

Results

Approaching motivational and emotional disorders through the lens of (computational) reward processing furnishes a number of important results with respect to two of the polarities that have plagued psychiatry but has not made as much progress with respect to a third.

Biological vs. Psychosocial

Computational psychiatry simultaneously addresses the computational level of what the problem is, the algorithmic level of how it can be operationalized in terms of information processing, and the implementation level in terms of the neural substrate. More practical considerations, such as the behavioral economics of interpersonal exchanges (Camerer, 2003), has obliged scientists to integrate social psychology and neuroscience with basic, or impersonal, reward processing. Let us consider two findings: first, that subjective wellbeing follows RPEs (Rutledge et al., 2014) as above. Second, during interpersonal exchange people may encode both ordinary RPEs (e.g., I'm pleasantly surprised with what she gave me) but also person-representation prediction errors [She will be pleasantly surprised about me, as I'm about to reciprocate generously (Xiang et al., 2012)]. If ordinary RPEs drive some aspects of emotion, it would be strange indeed if person-representation RPEs were unrelated to the strong emotions we experience in an interpersonal sphere: for example, their fragility in emotionally unstable personality or their presumed dearth in psychopathy.

On the other hand, of course, we are far from elucidating the actual way in which social emotions and non-social emotions are represented in their neurological substrates and, inversely, how social and non-social emotional processing changes this substrate, be it through trauma (Chen and Etkin, 2013), learning in psychosis (Murray et al., 2008) or subtle plasticity (Garvert

et al., 2015). Signatures of biased reward processing have been found in several disorders but they are far from explaining these disorders either in the sense of explaining symptoms in the here and now or in the sense of predicting the course of the disorder much better than traditional methods (Whelan et al., 2014).

Disease vs. Maladjustment

Learning about reward takes place at different levels of information processing. Let us consider the example of psychosis. The early, and celebrated, aberrant salience hypothesis of psychotic disorders (Kapur et al., 2005) postulated a disease level wherein dopamine discharges might be epileptic-like, unrelated to information processing, leading to the establishment of psychotic associations (both beliefs and choices) at the phenomenological and behavioral levels. Such an account separates the diseased brain reporting aberrant increased salience; and the healthy brain downstream that tries to make sense of this abnormal salience. However no epileptiform activity has been demonstrated. Increased aberrant salience has been demonstrated in association with schizotypy in healthy individuals and in medicated patients with delusions (Roiser et al., 2009); however, it does not seem to be prominent in prepsychotic and early psychotic states, where no changes in aberrant salience have been found so far (Smieskova et al., 2015). At the same time there is evidence that exaggerated dopamine reactivity to stress is associated with psychotic experiences in predisposed individuals (Hernaus et al., 2015).

Therefore the evidence points toward disease being an overall brain-state, the result of adjustment to psychobiological challenges performed by the individual's neural phenotype. Computationally, this is inference about salient stimuli at the developmental timescale; while genetically it is likely to be based on 'intermediate phenotypes,' e.g., of atypical connectivity (Cao et al., 2016). We can see that this framework renders the dualistic view of 'disease' and 'maladjustment' obsolete. The canonical teaching of an illness being explained in terms of predisposing, precipitating and perpetuating factors fits much more comfortably with the dynamical view of computational psychiatry, wherein dopamine reactivity or the interplay of prior and posterior beliefs are meaningful (if suboptimal) at different but intimately linked Marrian levels. The computational models of Ruppín and coworkers (Horn and Ruppín, 1994) illustrate a beautiful early example of such thinking. They suggested that the brain performed compensatory adjustments to long-range dysconnectivity in order to preserve the ability to activate appropriate perceptions in response to stimuli. However, these compensatory adjustments result in a propensity for percepts that bear small correlation to stimuli (i.e., hallucinations) to arise. Neurobiological and computational research has greatly refined these insights. We close this brief foray into psychosis research by point out a promising theme relevant to the role of reward, the focus of this issue. From the early theories of dopamine-dependent signal-to-noise (Cohen and Servan-Schreiber, 1992; Servan-Schreiber et al., 1996) to the influential analysis of the role of precision at sensory vs. cognitive levels (Adams et al., 2013) to the findings of exaggerated dopamine reactivity to stress (Hernaus et al., 2015), psychosis has been about aberrations and

compensatory changes in synaptic gain. The original aberrant salience theory of psychosis has opened new horizons regarding the role of reward- and threat- anticipation in psychosis; yet it may be the increasingly sophisticated understanding of synaptic gain, especially in its guise as precision calculated in cortical NMDA fields (Adams et al., 2013), that helps us go beyond the oversimplified aspects of salience theory.

At a theoretical level some biological factors are so dominant that to call them 'predisposing factors' is misleading (e.g., Down's syndrome causing Alzheimer's disease). These can be thought of as maladjustments at another level of the hierarchy – where an evolving reproductive apparatus has not learnt to avoid trisomies. Such maladjustments may be chance events or indeed the result of optimizing compromises between priorities.

At the same time the normative view of reward processing contains an ambiguity that needs acknowledgment and resolution. This is that for *any* input-output behavioral pattern a cost structure can be found for which this pattern of behavior is optimal (Daunizeau et al., 2010). For *any* behavior we can simply say that the person in question emits it because it genuinely optimizes their happiness. This is analogous to the psychological assertion that a patient 'refuses to change because it would be too painful for them,' or that an addict or pedophile simply finds indulging too rewarding to trade it against an alternative. Given a conception of what is valuable, e.g., making the most money, we can offer to explain how people attempt to optimize their behavior, and which parts of the process may go wrong. In current practice most research that investigates abnormalities of reward processing takes as a starting point an assumption that there are rewards out there, which have a normative relationship with the individual's behavior and that people *should* value and *should* seek. When a rat or human are hungry, two lumps of sugar are more rewarding than one, and we can measure how much harder subjects are willing to work for the chance to get them. We have a normative yardstick: our subject should work just hard enough to maximize the utility of (sugar + effort). Motivational disorder is then defined as a statistically significant deviation from this norm. However, in the real world it is hard to know what people *should* care about and computational, biological and psychosocial research agendas could do well to take seriously what we don't know.

Computational psychiatry does not do as well, as yet, when dealing with the complexity of human emotion. The problem is acute not because we should address emotion in its huge complexity, but because we have so far dealt with it by a simplification into positive vs. negative emotion, albeit tagged according to experimental tasks in question. If human emotion relevant to psychopathology contains multiple facets as inseparable parts of a phenomenological states, and if these rich states have computational relevance, then current studies are likely to be very remote from actual clinical relevance. Paradiso and Rudrauf (2012) put it eloquently: "... the fear experienced by a mountain climber in potential danger has levels of social complexity unlikely to be reached in mice. In addition to fearing his own end, the mountain climber anticipating a possible death is equally likely also to be scared of losing his spouse and children, leaving them fatherless and exposed to

dangers, of the financial consequences of his death on them, of the emotional effects on his parents, and so on. He may simultaneously experience shame (another social emotion) and danger (perhaps toward his self) for having neglected what he thinks were routine safety measures. A human facing the possibility of ceasing to exist has emotions that encompass the inescapable social nature and interconnectedness of our species and multiple levels of self-representation and projection.” We don’t really know which complex emotional constellations found in psychiatric disorders are most relevant, especially for decision-making that can be considered pathological. At the moment we haven’t developed a good way of addressing this most important question scientifically either.

Discussion

A computational psychiatry of emotional disorders has begun to put on the table key issues that have plagued psychiatry. It provides a framework for bridging biological-psychological-social divides and offers novel perspectives on the question of emotional-motivational ‘diseases’ versus ‘problems.’ This is rendered possible by formulating disorders of motivation and emotion within a normative probabilistic framework which offers sophisticated and neurobiologically plausible accounts of how reward motivate decisions. Many challenges remain. Phenomenology is only tentatively connected to computation; much-promising theoretical concepts have not been put to experimental test, while their normative basis is not understood. For example, we have no rigorous normative account of what utility structures correspond to mental health. A key example is how reward *should* be discounted in the face of time (inter-temporal discounting), valence (complex discounting of negative future events, including dread) or social distance (social discounting). Therefore the statistical connections that have been found between temporal discounting and addictive disorders lack a true normative basis.

Let us now consider a libertarian (or Szaszian) critique of reward processing as a basis for psychiatric research. Szasz protested against a medicalization of deviant behavior, believing that so called psychiatric disorders lack an adequate biological basis. Hence ‘medicalizing’ unjustifiably transgressed peoples’ autonomy (Szasz, 1960). Deciding *a priori* what reward people should value more (as manifested in their choices) or what reward they should care about (as manifested in their phenomenology) is just as much ‘playing god’ once we move beyond trivial choices: in many cases psychiatrically relevant situations are complex enough to negate a dream of finding a normative standard against which to measure motivational disorder. Reward processing should maximize long-term outcomes and so in research practice we use paradigms that have well-defined ends or may be thought of as going on ‘for ever’ (as for example near the beginning of a task with hundreds of trials). Yet what sorts of long-term outcomes are involved in the long-term reward processing important for psychiatry? Individual reproductive fitness? We have no clear idea, and the temptation is to

import convenient social norms, rendering our framework only pseudo-normative. Even in the simple example of working for lumps of sugar, mentioned above, there will usually be *some* evaluation of effort and sugar⁵ that renders behavior optimal. This evaluation may be normative with respect to the person’s history, not the task. In any psychiatrically relevant situation considerations rapidly multiply. For example, what if our hungry human is overweight? And what if the reproductive fitness associated with slimness (attracting mates) is socially constructed?

Thomas Szasz and the libertarian tradition (to which the authors belong) argue that rather than impose norms on people – say about which reward would maximize their life expectancy, their reproductive success – we should respect the priorities they have and, by definition, accept a person’s autonomy to seek their own reward by deploying their own motivational structures. So is there no such thing as a motivational or emotional disorder and in fact everyone is just doing the best they can? Szasz would claim that the dream of aberrant reward processing pinning down what’s essential about motivational and emotional disorders is no more solid than the ‘chemical imbalance’ theory of depression or of Freud’s ‘unconscious motivation’ theory of mental illness. To be more specific if the Reward processing domain of the otherwise promising ‘Research Domain Criteria’ framework (Casey et al., 2013) is applied too simplistically we may end up with exactly the same mistakes as in previous biological or psychoanalytic normative straightjackets.

If a Szaszian position simply accepts peoples’ choices for what they are, its extreme opposite would be a 1984 world where people have been taught through social, psychological and biological interventions, not only *what to decide* but actually *what to desire*. While we recoil from the Szaszian extreme as it is dismissive of the importance of psychiatric suffering, psychiatrists cannot dictate what patients should care about – even about their symptoms. The so-called recovery movement can already teach computational neuroscientists that the rewards that patients really care about are not so much to do with their symptoms as with their life goals and values. In that case perhaps the priorities for researching archetypal motivational disorders like depression are not about ‘what motivational disturbance underpins depression’ but ‘what decision structures of the depressed can help them fulfill their values’ (Hayes et al., 1999). Here we have dialectic, because the scientific baby should not be thrown away with the essentialist bath water. The clinician could bring to the patient a biopsychosocial assessment of ‘wrong priors,’ ‘wrong models,’ or ‘wrong utilities.’ They would then decide *in dialog with a patient*, with a diagnosis of say the successor of ‘Depressive Episode,’ now defined in computational terms, what key needs must be targeted and optimized. When it comes to the severe mental illnesses, formulations that go beyond ‘Schizophrenia,’ or indeed ‘Abnormal Salience syndrome’ (Van Os, 2009) will help clinicians and patients consider emotional and motivational dispositions both as threats and as

⁵i.e., the person’s goal or preference priors: Friston et al. (2013).

instruments toward recovery. Of course this account assumes patients with some capacity to consider the issues in question, which may itself be severely compromised – for example in acute psychosis.

Why bring in the concept of need when considering reward and emotional disorder? Because biologically reward is not an end in itself, but a good surrogate toward longer-term biological goals. The stability properties of a self-perpetuating system, like a species in an ecosystem, can be conceptualized in terms of having the ‘purpose’ or ‘goal’ to keep perpetuating the system (e.g., the species). One has to be careful philosophically to avoid false teleological justifications, but in the first instance this is small print. We assert that there are physiological homeostatic needs, reproductive/sexual needs, and more complex ones such as needs for social contact. Furthermore, people are motivated by reward that extend beyond their own lives. They will often, in fact, sacrifice their life for much less than ‘two brothers or four cousins,’ as mathematical evolutionary biologists have put it (Maynard Smith, 1993). Each of these needs entails goals, desires and reward; all are relevant to psychiatry; but probably few can be the target of fruitful intervention for each particular patient.

References

- Adams, R. A., Stephan, K. E., Brown, H. R., Frith, C. D., and Friston, K. J. (2013). The computational anatomy of psychosis. *Front. Psychiatry* 4:47. doi: 10.3389/fpsyt.2013.00047
- APA (2013). *DSM-5 Diagnostic and Statistical Manual of Mental Disorders*. Arlington, TX: American Psychiatric Association.
- Bentall, R. (2009). *Doctoring the Mind: Is Our Current Treatment of Mental Illness Really any Good?* New York, NY: NYU Press.
- Boyle, M., and Johnstone, L. (2014). Alternatives to psychiatric diagnosis. *Lancet Psychiatry* 1, 409–411. doi: 10.1016/S2215-0366(14)70359-1
- Busmeyer, J., and Bruza, P. (2012). *Quantum Models of Cognition and Decision*. Cambridge: Cambridge University Press. doi: 10.1017/CBO9780511997716
- Busmeyer, J., Wang, Z., and Lambert-Mogiliansky, A. (2009). Empirical comparison of Markov and quantum models of decision making. *J. Math. Psychol.* 53, 423–433. doi: 10.1016/j.jmp.2009.03.002
- Camerer, C. F. (2003). *Behavioral Game Theory: Experiments in Strategic Interaction*. Princeton, NJ: Princeton University Press.
- Cao, H., Dixon, L., Meyer-Lindenberg, A., and Tost, H. (2016). Functional connectivity measures as schizophrenia intermediate phenotypes: advances, limitations, and future directions. *Curr. Opin. Neurobiol.* 36, 7–14.
- Casey, B. J., Craddock, N., Cuthbert, B. N., Hyman, S. E., Lee, F. S., and Ressler, K. (2013). DSM-5 and RDoC: progress in psychiatry research? *Nat. Rev.* 4, 810–814. doi: 10.1038/nrn3621
- Chen, A., and Etkin, A. (2013). Hippocampal network connectivity and activation differentiates post-traumatic stress disorder from generalized anxiety disorder. *Neuropsychopharmacology* 38, 1889–1898. doi: 10.1038/npp.2013.122
- Cohen, J. D., and Servan-Schreiber, D. (1992). Context, cortex, and dopamine: a connectionist approach to behavior and biology in schizophrenia. *Psychol. Rev.* 99, 45–77. doi: 10.1037/0033-295X.99.1.45
- Daunizeau, J., den Ouden, H. E. M., Pessiglione, M., Kiebel, S. J., Friston, K. J., and Stephan, K. E. (2010). Observing the observer (II): deciding when to decide. *PLoS ONE* 5:e15555. doi: 10.1371/journal.pone.0015555
- Dolnick, E. (1998). *Madness on the Couch: Blaming the Victim in the Heyday of Psychoanalysis*. New York, NY: Simon and Schuster.
- Friston, K. J., Schwartenbeck, P., FitzGerald, T., Moutoussis, M., Behrens, T., and Dolan, R. J. (2013). The anatomy of choice: active inference and agency. *Front. Hum. Neurosci.* 7:598. doi: 10.3389/fnhum.2013.00598
- Garvert, M. M., Moutoussis, M., Kurth-Nelson, Z., Behrens, T., and Dolan, R. J. (2015). Learning-induced plasticity in medial prefrontal cortex predicts preference malleability. *Neuron* 85, 418–428. doi: 10.1016/j.neuron.2014.12.033
- Guitart-Masip, M., Huys, Q., Fuentemilla, L., Dayan, P., Duzel, E., and Dolan, R. (2012). Go and no-go learning in reward and punishment: interactions between affect and effect. *NeuroImage* 62, 154–166. doi: 10.1016/j.neuroimage.2012.04.024
- Hayes, J., and Bell, V. (2014). Diagnosis: one useful method among many. *Lancet Psychiatry* 1, 412–413. doi: 10.1016/S2215-0366(14)70399-2
- Hayes, S. C., Strosahl, K. D., and Wilson, K. G. (1999). *Acceptance and Commitment Therapy: An Experiential Approach to Behavior Change*. New York, NY: Guilford.
- Hernaus, D., Collip, D., Lataster, J., Viechtbauer, W., Myin, E., Ceccarini, J., et al. (2015). Psychotic reactivity to daily life stress and the dopamine system: a study combining experience sampling and [¹⁸F]fallypride positron emission tomography. *J. Abnorm. Psychol.* 124, 27–37. doi: 10.1037/abn0000010
- Horn, D., and Ruppert, E. (1994). Synaptic compensation in attractor neural networks: modeling neuropathological findings in schizophrenia. *Neural Computation* (in press).
- Huys, Q., Guitart-Masip, M., and Dayan, P. (2014). Decision theoretic psychiatry. *Clin. Psychol. Sci.* 3, 374–377.
- Huys, Q., Moutoussis, M., and Williams, J. (2011). Are computational models of any use to psychiatry? *Neural Netw.* 24, 544–551. doi: 10.1016/j.neunet.2011.03.001
- Joffily, M., and Coricelli, G. (2013). Emotional valence and the free-energy principle. *PLoS Comput. Biol.* 9:e1003094. doi: 10.1371/journal.pcbi.1003094
- Kapur, S., Mizrahi, R., and Li, M. (2005). From dopamine to salience to psychosis—linking biology, pharmacology and phenomenology of psychosis. *Schizophr. Res.* 79, 59–68. doi: 10.1016/j.schres.2005.01.003
- Marr, D. (1982). *Vision*. New York, NY: Freeman.
- Maynard Smith, J. (1993). *The Theory of Evolution*. Cambridge: Cambridge University Press.
- Montague, P. R., Dolan, R. J., Friston, K. J., and Dayan, P. (2012). Computational psychiatry. *TICS* 16, 72–80. doi: 10.1016/j.tics.2011.11.018

Conclusion: Computational Psychiatry must be Profoundly Biopsychosocial

In the best possible world scientists will take seriously the question of what needs really matter for patients, what reward form the best surrogates or milestones toward the fulfillment of such needs and will do so in open collaboration with relevant stakeholders. At first sight the rigorous, biologically based discipline of computational psychiatry seems far from patients’ expressed needs, yet the fact that it puts reward and motivation at the center of understanding psychiatric disorder gives it a privileged vantage point toward serving patients.

Acknowledgments

RD is supported by a Wellcome Trust Senior Investigator Award (ref 098362/Z/12/Z). The current work is funded by a Strategic Award by the Wellcome Trust (ref 095844/7/11/Z). MM is also supported by the Biomedical Research Council.

- Murray, G., Corlett, P., Clark, L., Pessiglione, M., Blackwell, A., Honey, G., et al. (2008). Substantia nigra/ventral tegmental reward prediction error disruption in psychosis. *Mol. Psychiatry* 13, 267–276. doi: 10.1038/sj.mp.4002058
- Paradiso, S., and Rudrauf, D. (2012). Struggle for life, struggle for love and recognition: the neglected self in social cognitive neuroscience. *Dialogues Clin. Neurosci.* 14, 65–75.
- Roiser, J. P., Stephan, K. E., den Ouden, H. E. M., Barnes, T. R. E., Friston, K. J., and Joyce, E. M. (2009). Do patients with schizophrenia exhibit aberrant salience? *Psychol. Med.* 39, 199–209. doi: 10.1017/S0033291708003863
- Romito, P. (2008). *A Deafening Silence: Hidden Violence Against Women and Children*. Bristol: Policy Press.
- Rushworth, M., Noonan, M., Boorman, E., Walton, M., and Behrens, T. (2011). Frontal cortex and reward-guided learning and decision-making. *Neuron* 70, 1054–1069. doi: 10.1016/j.neuron.2011.05.014
- Rutledge, R., Skandali, N., Dayan, P., and Dolan, R. (2014). A computational and neural model of momentary subjective well-being. *Proc. Natl. Acad. Sci. U.S.A.* 111, 12252–12257. doi: 10.1073/pnas.1407535111
- Servan-Schreiber, D., Cohen, J. D., and Steingard, S. (1996). Schizophrenic deficits in the processing of context. A test of a theoretical model. *Arch. Gen. Psychiatry* 53, 1105–1112. doi: 10.1001/archpsyc.1996.01830120037008
- Seymour, B. O., Doherty, J. P., Dayan, P., Koltzenburg, M., Jones, A. K., Dolan, R. J., et al. (2004). Temporal difference models describe higher-order learning in humans. *Nature* 429, 664–667. doi: 10.1038/nature02581
- Smieskova, R., Roiser, J. P., Chaddock, C. A., Schmidt, A., Harrisberger, F., Bendfeldt, K., et al. (2015). Modulation of motivational salience processing during the early stages of psychosis. *Schizophr. Res.* 166, 17–23. doi: 10.1016/j.schres.2015.04.036
- Szasz, T. (1960). The myth of mental illness. *Am. Psychol.* 15, 113–118. doi: 10.1037/h0046535
- Van Os, J. (2009). 'Salience syndrome' replaces 'schizophrenia' in DSM-V and ICD-11: psychiatry's evidence-based entry into the 21st century? *Acta Psychiatrica Scandinavica* 120, 363–372. doi: 10.1111/j.1600-0447.2009.01456.x
- Whelan, R., Watts, R., Orr, C., Althoff, R., Artiges, E., Banaschewski, T., et al. (2014). Neuropsychosocial profiles of current and future adolescent alcohol misusers. *Nature* 512, 185–189. doi: 10.1038/nature13402
- Xiang, T., Ray, D., Lohrenz, T., Dayan, P., and Montague, P. R. (2012). Computational phenotyping of two-person interactions reveals differential neural response to depth-of-thought. *PLoS Comput. Biol.* 8:e1002841. doi: 10.1371/journal.pcbi.1002841

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2015 Moutoussis, Story and Dolan. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



A Computational Analysis of Aberrant Delay Discounting in Psychiatric Disorders

Giles W. Story^{1,2,3*}, Michael Moutoussis^{1,2} and Raymond J. Dolan^{1,2}

¹ Max Planck University College London Centre for Computational Psychiatry and Ageing Research, University College London, London, UK, ² Wellcome Trust Centre for Neuroimaging, University College London, London, UK, ³ Centre for Health Policy, Imperial College London, Institute of Global Health Innovation, St. Mary's Hospital, London, UK

OPEN ACCESS

Edited by:

Gianluca Castelnuovo,
Università Cattolica del Sacro Cuore,
Italy

Reviewed by:

Michelle Dow Keawphalouk,
Harvard and Massachusetts Institute
of Technology, USA
Warren K. Bickel,
Virginia Polytechnic Institute and State
University, USA

*Correspondence:

Giles W. Story
g.story@ucl.ac.uk

Specialty section:

This article was submitted to
Psychology for Clinical Settings,
a section of the journal
Frontiers in Psychology

Received: 29 June 2015

Accepted: 04 December 2015

Published: 13 January 2016

Citation:

Story GW, Moutoussis M and
Dolan RJ (2016) A Computational
Analysis of Aberrant Delay Discounting
in Psychiatric Disorders.
Front. Psychol. 6:1948.
doi: 10.3389/fpsyg.2015.01948

Impatience for reward is a facet of many psychiatric disorders. We draw attention to a growing literature finding greater discounting of delayed reward, an important aspect of impatience, across a range of psychiatric disorders. We propose these findings are best understood by considering the goals and motivation for discounting future reward. We characterize these as arising from either the opportunity costs of waiting or the uncertainty associated with delayed reward. We link specific instances of higher discounting in psychiatric disorder to heightened subjective estimates of either of these factors. We propose these costs are learned and represented based either on a flexible cognitive model of the world, an accumulation of previous experience, or through evolutionary specification. Any of these can be considered suboptimal for the individual if the resulting behavior results in impairments in personal and social functioning and/or in distress. By considering the neurochemical and neuroanatomical implementation of these processes, we illustrate how this approach can in principle unite social, psychological and biological conceptions of impulsive choice.

Keywords: discounting, time preference, psychiatric, computational psychiatry, mental illness, biopsychosocial

INTRODUCTION

Vitae summa brevis spem nos vetat incohare longam

Life's short span forbids our embracing far-reaching hopes - Horace, Odes (23BC)

Humans and animals often accept a smaller reward immediately, rather than wait to receive a larger reward in the future (Ainslie, 1974; Thaler, 1981; Thaler and Shefrin, 1981; Fishburn and Rubinstein, 1982; Frederick et al., 2002; McClure et al., 2007; Kalenscher and Pennartz, 2008; Pine et al., 2009). In economic terms, this behavior indicates that the subjective value of reward decreases as it is delayed, a process referred to as temporal discounting (for reviews see Frederick et al., 2002; Kalenscher and Pennartz, 2008). As we will discuss, biological agents have good reason to discount delayed rewards, since these might either fail to materialize or arrive too late to satisfy the organism's current needs. Indeed, as pointed out by the Roman poet Horace in the quotation above, the ultimate motive for discounting is that the agent will die before deferred rewards are realized.

In humans, temporal discounting can be measured by examining choices between quantities of money at varying delays (Mazur, 1987; Kirby and Maraković, 1995; Myerson et al., 2001; Green and Myerson, 2004). The most commonly used method elicits choices between a larger, delayed amount

of money, (e.g., “\$100 in 6 months”), and a series of immediate amounts of decreasing magnitude (e.g., “\$80 today”). By observing at each delay the magnitude of smaller-sooner reward at which the participant switches to preferring the later reward, the decrease in value of the later reward can be plotted as a function of delay. A non-parametric estimate of discounting can be derived by taking the area beneath this indifference curve (Myerson et al., 2001). Alternatively, the shape of the curve can be fitted with a discount *function*.

Samuelson (1937), and later Strotz (1957), showed that a decision-maker who discounts future benefits according to an exponentially decreasing function (and behaves *as if* to maximize the sum of exponentially discounted reward) allocates resources across time in a self-consistent manner. Under the classical model, the effect of delay, d , is described by an (exponential) discount function, here denoted by $\Delta(d)$, such that:

$$\Delta(d) = e^{-kd} \quad (1)$$

Where k is an exponential discount rate, such that higher values of k lead to a steeper decrease in reward value with delay. The effect of reward magnitude, here signified by r , is independently described by an instantaneous utility function, $u(r)$, such that the subjective utility of a stream of future rewards is then given by:

$$U(r_t, r_{t+1}, r_{t+2} \dots r_{T-1}, r_T) = \sum_t^T u(r_t) \Delta(\tau - t) \quad (2)$$

As reviewed by Frederick et al. (2002), the above account was not intended as a veridical psychological model of choice over time. In keeping with this, many experimental studies have shown that a discounting function is better approximated via a hyperbolic than an exponential function (e.g., Green et al., 1994; Kirby and Herrnstein, 1995; Kirby and Maraković, 1995; Myerson and Green, 1995; Laibson, 1997; van der Pol and Cairns, 2002; Rubinstein, 2003), of the form:

$$\Delta(d) = \frac{1}{1 + kd} \quad (3)$$

Here k denotes a hyperbolic discount rate (though for alternative accounts see Read, 2001; Kable and Glimcher, 2010; Read et al., 2012; Luhmann, 2013).

Temporal discounting has received considerable attention in human behavioral neuroscience, not least because many forms of maladaptive behavior are readily characterized as pursuit of immediate gratification at the expense of reaping greater rewards in the future (Critchfield and Kollins, 2001; Bickel et al., 2007, 2014a; Koffarnus et al., 2013; Story et al., 2014). Indeed, lending validity to the discounting construct, steeper discounting is positively associated with behaviors with potentially harmful long-term consequences such as tobacco smoking (Odum et al., 2002; Epstein et al., 2003; Reynolds et al., 2004; Bickel et al., 2008; MacKillop and Kahler, 2009; Fields et al., 2009a,b; Reynolds and Fields, 2012), alcohol use (Van Oers et al., 1999; Mazas et al., 2000; Petry, 2001; Field et al., 2007; Reynolds et al., 2007; Rossow, 2008; MacKillop and Kahler, 2009; Moore and Cusens, 2010), illicit

drug misuse (Kirby et al., 1999; Petry and Casarella, 1999; Kollins, 2003; Petry, 2003; Kirby and Petry, 2004; Washio et al., 2011; Stanger et al., 2012), credit card debt (Meier and Sprenger, 2012) and risky sexual or drug-taking practices (Odum et al., 2000; Dierst-Davies et al., 2011). Also, many authors have explored how discounting relates to demographic variables, finding that measured discounting decreases across the lifespan (Green et al., 1996, 1999; Chao et al., 2009; Steinberg et al., 2009), is negatively correlated with income (Green et al., 1996; Eckel et al., 2005; Reimers et al., 2009), and tends to be lower in individuals living in the developed world than in the developing world (Wang et al., 2010). Furthermore, although discounting is sensitive to a gamut of contextual factors (for a review see Koffarnus et al., 2013), the level of discounting has been shown to exhibit high test-retest reliability when measured under similar conditions (Odum, 2011), and the extent of individual discounting for different forms of reward is correlated (Odum, 2011), suggesting that discounting has a substantial trait component.

More recently, researchers have taken an interest in comparing discounting behavior in groups who exhibit symptoms of a given psychiatric disorder and those who do not. These studies have found evidence for steeper discounting amongst patients with symptoms of schizophrenia (Heerey et al., 2007, 2011; Ahn et al., 2011; MacKillop and Tidey, 2011; Wing et al., 2012; Avsar et al., 2013; Weller et al., 2014), depression (Takahashi et al., 2008; Dennhardt and Murphy, 2011; Dombrowski et al., 2012; Imhoff et al., 2014; Pulcu et al., 2014), mania (Mason et al., 2012), attention deficit hyperactivity disorder (ADHD) (Barkley et al., 2001; Tripp and Alsop, 2001; Bitsakou et al., 2009; Paloyelis et al., 2010a,b; Scheres et al., 2010; Scheres and Hamaker, 2010), anxiety disorder (Rounds et al., 2007) and cluster B personality disorder (Dougherty et al., 1999; Moeller et al., 2002; Petry, 2002; Dom et al., 2006a,b; Lawrence et al., 2010; Coffey et al., 2011). This line of enquiry is not without theoretical justification, for example the broader construct of impulsivity, defined as taking action without forethought or regard for consequences (Moeller et al., 2001), of which discounting is an element, is a defining feature of some psychiatric disorders, for example borderline personality disorder (Moeller et al., 2001; DSM V, 2013) and mania (Swann, 2009). Also, psychiatric disorders are strongly associated with poor health choices, including but not limited to cigarette smoking, and drug and alcohol misuse (Robson and Gray, 2007), which have themselves been associated with steeper discounting (Bickel et al., 2012b, 2014a,b; Story et al., 2014). However, in many cases this research, although clearly valuable, appears to have been opportunist.

In this article we attempt to understand increases in discounting seen across a range of psychiatric disorders in light of the *reasons* why people should discount the future in the first place. We propose that the study of intertemporal impulsivity in psychiatric disorders would benefit from fractionating these underlying motives, and that parsing discounting in this manner can assist in drawing out the contributing psychological and biological processes. Our approach follows that of the neuroscientist David Marr (Marr, 1982), who proposed that information processing systems can be understood at three

levels of analysis: a “computational” level, specifying what information processing problem is being solved by the system, an “algorithmic” level, formalizing how the system attempts to solve the problem, and an “implementational” level, denoting how these processes are realized physically.

For the case of discounting, the computational problem is easily defined in economic terms: to optimize the sum of future reward. However, this definition obscures a difficult question as to what constitutes “reward” (Moutoussis et al., 2015). It is convenient here to assume that all biological agents share some fundamental objective function. Rather than attempting to characterize the objective function directly, we assume some consensus on the kinds of outcome that organisms often seek, and that can therefore be considered “rewarding.” We then consider a subset of generic scenarios under which behavior consistent with discounting would indeed optimize the sum of future “reward.” This will give us some insight as to the contexts that agents, who discount future reward in different ways, including humans deemed to have mental disorders, might be adapted to.

We go on to speculate as to the broad classes of algorithms that biological agents might use to optimize reward, and where relevant their possible neural implementation. We argue that the application of this approach to psychiatric disorders, the bedrock of the emerging field of computational psychiatry (Huys et al., 2011; Montague et al., 2012; Friston et al., 2014; Stephan and Mathys, 2014; Wang and Krystal, 2014), can help to bridge a gap between psychological and biological conceptions of mental ill health (for further discussion see Moutoussis et al., 2015).

MARR’S COMPUTATIONAL LEVEL: REASONS TO DISCOUNT FUTURE REWARD

The discount function estimated from the analysis of intertemporal choice paradigms is likely to reflect the influence of factors jointly serving to make impatience potentially advantageous. A key ambiguity in the classical economic model concerns whether these factors should be properly assigned to the time series of future rewards, or to the discount function (Frederick et al., 2002; Frederick and Loewenstein, 2008; Friston et al., 2013; for a review of contextual influences on discounting see Koffarnus et al., 2013). The following discussion illustrates that if they are made fully explicit in the utility function, behavior consistent with temporal discounting emerges.

Opportunity Cost Growth and Missed Investment

For most organisms growth and development are necessary to reach reproductive capacity (Williams, 1957). For humans, development also extends to furthering one’s social status. Growth potential motivates obtaining rewards sooner rather later, since earlier rewards can be invested—effectively loaned out at some rate of interest (see Rachlin, 2006; Kacelnik, 2011). The form of discounting that results depends on whether or not interest can be re-invested. Under the most straightforward

scenario, referred to as simple interest, interest is not reinvested during the term of the loan. Consider a reward with utility r (for simplicity we omit the instantaneous utility function) invested for a period of time, d , to yield a larger payout, R . With simple interest:

$$R = r + krd \quad (4)$$

Solving for r and expressing as a ratio of the payout gives:

$$\frac{r}{R} = \frac{1}{1 + kd} \quad (5)$$

A decision-maker should therefore be indifferent between a larger reward of utility, R , received after a delay, d , and a smaller reward, r , received immediately. Thus, linear growth (simple interest) motivates hyperbolic discounting (see Read, 2004; Rachlin, 2006).

In the above example, after the delay has lapsed the agent ought to reclaim their money and re-invest the entire payout to avoid losing out to a lower rate of interest. Compound interest represents a continual reinvestment of the payout, and generates exponential growth, such that the payout accrued at time d after choosing r is given by:

$$R = re^{gd} \quad (6)$$

Where g reflects the interest rate. Rearranging as before gives:

$$\frac{r}{R} = e^{-gd} \quad (7)$$

Thus, compound interest motivates exponential discounting.

Missed Income

In the natural world, delay often entails inactive waiting, during which other sources of reward cannot be harvested. The cost associated with an inactive delay can be quantified as the reward that is missed out on while waiting (Kacelnik, 2011). Under one such formulation, organisms should consequently choose an action which maximizes a *rate* of reward per unit time, a concept that has arisen in ecological theory independently from the notion of discounting (Stevens and Krebs, 1986). Under this formulation, discounted value is simply inversely proportional to delay (Chung and Herrnstein, 1967). It can be easily shown however that if even “immediate” rewards are associated with some small delay, m , where $m = 1/k$, this is equivalent to hyperbolic discounting (Daw and Touretzky, 2000). Thus, at indifference:

$$\frac{r}{m} = \frac{R}{m + d} \quad (8)$$

Rearranging as previously:

$$\frac{r}{R} = \frac{m}{m + d} = \frac{1}{1 + d/m} = \frac{1}{1 + kd} \quad (9)$$

A corollary of this theory is that the opportunity cost of delaying reward on a particular option depends on the average rate of

reward from all other options (Chung and Herrnstein, 1967; Daw and Touretzky, 2000; Niv et al., 2007).

Inactive waiting leads to interesting results if other options become available only once the delays associated with the current choice have lapsed. Consider for example a lawyer who is paid by the hour for seeing clients at weekdays, but does not work at weekends. Say that he or she has two lunch options, either waiting in a long queue for a tasty lunch at a popular café, or being able to buy an equally calorific but less enjoyable meal straightaway at a sandwich bar. The lawyer might be optimally inclined to choose the sandwich bar on weekdays, so as to facilitate a sooner return to work, but might choose to wait at the café if faced with the same choice on a weekend. Here the intertemporal choice is influenced by other available sources of reward, which are inaccessible during the delay. In ecological terms, if an organism is foraging in a reward-rich area, the opportunity cost of delaying foraging by engaging in other activities is greater than when foraging in a reward-poor area (Niv et al., 2007).

Thus, expressed in terms of the total reward received, and letting the average rate of reward *available after the delay* be signified by ρ , then at indifference:

$$R = r + \rho d \quad (10)$$

Thus:

$$r = R - \rho d \quad (11)$$

This arrangement allows for the possibility that a delayed reward carries negative value, whereby a decision-maker would willing to pay so as to be able to resume seeking rewards at the average rate, rather than to wait for the delayed reward.

Uncertainty Probability and Hazard

Whenever reward (capital) is stored for the future, for example when a person lends money to another person or when an animal stores food, there is some possibility that the capital will be lost (for example if a conspecific raids the food store or the debtor defaults on their loan). If there is some constant probability per unit time, referred to as a hazard rate, that future rewards do not materialize as promised, the expected value of reward (magnitude \times probability) decreases exponentially with delay and gives rise to exponential discounting (Sozou, 1998).

Following the notation above at indifference:

$$r = Re^{-\lambda t} \quad (12)$$

Rearranging:

$$\frac{r}{R} = e^{-\lambda t} \quad (13)$$

Where λ denotes a constant hazard rate.

Thus, the agent choosing whether to store reward should adopt a discount rate appropriate to the estimated hazard rate. For example a creditor ought to demand a rate of interest that is commensurate with the risk of the debtor's chance of default

per unit time. Interestingly, where the appropriate hazard rate is uncertain, decision-makers ought to weight each possible hazard rate by its probability of being the true rate; such a weighted average of exponential rates approximates hyperbolic discounting (Sozou, 1998; Kurth-Nelson and Redish, 2009). As shown by Sozou (1998), hyperbolic discounting results exactly if:

$$\int_0^\infty f(\lambda) e^{-\lambda t} d\lambda = \frac{1}{1 + kt} \quad (14)$$

Where $f(\lambda)$ is a probability density function over hazard rates. The above is satisfied if:

$$f(\lambda) = \frac{1}{k} e^{-\lambda/k} \quad (15)$$

i.e., if there is an exponential prior distribution over hazard rates, where k determines the shape of this distribution. In support of Sozou's theory, Takahashi et al. (2007) find that the subjective probability of receiving delayed reward in standard intertemporal choice tasks indeed decays hyperbolically.

As the quotation at the start of this article encapsulates, death creates a fundamental motive not to defer rewards for too long. In computational terms death can be considered to be an absorbing state, from which no future reward can be harvested. Notably a hazard rate for the event of dying can be seen to depend on the organism's current state, such that a greater physiological deficit is associated with a greater probability of dying per unit time. The fundamental value of reward is then its effect to reduce the hazard rate for dying (before successfully securing one's legacy). This argument suggests that it is optimal for biological agents to discount future reward more steeply when they are currently far from a physiological set point, based simply on an increased probability of their dying before future reward is attained.

Volatility

In summary, environmental hazards create a motive to discount the future, since future rewards might not materialize as promised. In addition, the utility of future rewards might be more uncertain, in the sense of having higher *variance* than immediate rewards (when the variance is known the resulting uncertainty is referred to as risk). Many behavioral economic studies have shown that people tend to be risk averse (Kahneman and Tversky, 1979; Holt and Laury, 2002; Trepel et al., 2005; Andersen et al., 2008; Platt and Huettel, 2008; Jones and Rachlin, 2009), in so far as they will accept a smaller expected payoff over a larger expected payoff with higher variance. If future events tend to evolve with a random component, the uncertainty associated with future events increases with delay (Mathys et al., 2011). To take an example, a decision-maker responding to a discounting questionnaire might have some degree of uncertainty about the subjective utility of a \$20 payout received immediately (if this appears implausible, imagine being paid in a foreign currency, whose worth is uncertain). However, owing to volatility governing future events in their lives (e.g., becoming ill, falling into debt, national economic collapse), uncertainty regarding the utility of the \$20 ought to increase as it is delayed. In combination with risk aversion this motivates delay discounting. In support

of this idea, individual discount rates are correlated with risk aversion (Leigh, 1986; Anderhub et al., 2001; Eckel et al., 2005; Jones and Rachlin, 2009; Dohmen et al., 2010).

Notably, risk aversion can be expressed in terms of *probability* discounting, which is found to be hyperbolic in the odds against receiving a reward. Whilst probability discounting and temporal discounting are often found to be correlated across individuals (e.g., Jones and Rachlin, 2009), they are subject to distinct influences. For example, increasing reward magnitude increases probability discounting (i.e., risk aversion) and decreases temporal discounting (Green and Myerson, 2004). This is often taken as evidence that temporal discounting does not encompass an estimate of the risk associated with future rewards. However, pertinent to discounting is how a person estimates risk to be dependent on delay. Probability discounting offers a measure of risk aversion but does not access this time-dependent representation of risk. In support of this idea Takahashi et al. (2007) find that while probability and temporal discounting are uncorrelated across individuals, temporal discounting does correlate with the rate of decay in the subjective probability of receiving reward after increasing delay. This may help explain why psychiatric disorders are often associated with increased inter-temporal discounting but not necessarily with excessive probability discounting.

MARR'S ALGORITHMIC LEVEL: PROCESSES SUB-SERVING INTERTEMPORAL CHOICE

In the preceding analysis we have outlined some generic scenarios under which behavior consistent with discounting would be optimal. These scenarios illustrate that discounting need not be considered as a unitary process, rather as (implicitly or explicitly) reflecting an expectation of different environmental contingencies. Under reinforcement learning formulations, such contingencies are seen as engendering transitions in a state-space (Sutton and Barto, 1998; Dayan and Balleine, 2002; Dayan and Daw, 2008; Kurth-Nelson and Redish, 2009). That it is, an action is assumed to move the agent from one (discrete) state to another, where each state may be associated with a varying quantity of reward. The state-space is equivalent to the vector of rewards described in the classical economic model (Equation 2), though may also be made contingent on the agent's future behavior, giving rise to a matrix, or "decision-tree." A key question for this account is whether the (discounted) utility of a delayed reward is directly parameterized, which is to say that there is no more inference or learning beyond the state where this utility is considered, or whether the delayed reward is instead considered as part of a cascade of preceding states.

A Parametric Discount Function?

If higher organisms indeed represent a discount function parametrically, they would require a widespread and efficient system for making this information accessible for decision-making. Neuromodulatory systems, with their diffuse connections to many areas of the brain, would be well

placed to achieve this, and several authors have speculated that neuromodulators, such as dopamine and norepinephrine might represent some of the relevant parameters. For example, Niv et al. (2007) have proposed that the average rate of reward is signaled in the mammalian brain by tonic levels of extracellular dopamine in the striatum, suggesting that increased striatal dopamine availability might increase discounting by increasing the implicit opportunity cost of delay. Commensurate with this hypothesis, systemic administration in humans of the dopamine precursor L-Dopa increases discount rates (Pine et al., 2010), although potentially countervailing evidence is that decreasing dopamine transmission in rats by administration of haloperidol (Denk et al., 2005) or flupenthixol (Floresco et al., 2008) has been found to increase discounting, or in other studies to exert no significant effect on discounting (Winstanley et al., 2005).

Similarly, a good deal of decision-making neuroscience seeks to uncover how uncertainty is represented neurally (see Behrens et al., 2007; Wilson et al., 2010; Mathys et al., 2011; Nassar et al., 2012). A recent suggestion is that operating in an unstable environment is associated with tonic release (over a time course of minutes) of norepinephrine (Yu and Dayan, 2003, 2005). The latter would suggest that tonic norepinephrine might signal environmental volatility, and thus influence discounting. Clearly, further psychopharmacological work is needed to fully uncover the role of monoaminergic signaling on discounting behavior. Also, if organisms indeed have a parametric model of discounting in the strictest sense, then this ought to be revealed in the manner in which estimates of discounting are updated in light of changes in the environment, and careful behavioral work is required to probe this possibility.

Discounting as a Revealed Phenomenon

According to a second possibility outlined above, choosing a delayed reward leads to a cascade of states, and may (or may not) lead to the promised reward, which if it occurs, may be delivered in a variety of future states (just in time for Christmas, after I've been killed by a bus, etc.) (see Peters and Büchel, 2010). If an agent uses this cascade of states to evaluate their actions, only the resulting transitions will endow this action with whatever value percolates through from the end states. Here discounting takes place due to learning and/or inference, where the value of the reward gradually evaporates as inference (or learning) propagates through a cascade of states. Given the properties of organisms and their environments, as outlined above, behavior consistent with discounting would simply emerge as the end result of applying these learning processes to situations where there is delay in the receipt of reward. Under this possibility, in terms of the economic model, all relevant information is summarized in an agent's utility function, which then implicitly incorporates the discount function. It appears likely that organisms use parallel mechanisms to calculate the value of the resulting state-space, operating across different timescales of information integration, ranging from updating innate behaviors through evolution, through learning from experience, to inferring future states via deployment of a cognitive map or model of the world.

Reliable valuations may be refined and passed on through genetic inheritance and evolution. For example, the possibility of

death, and its associated opportunity cost, is likely incorporated through evolution, whereby internal states deviating from a homeostatic ideal, such as hunger and thirst, are assigned an innate cost as a proxy (see Keramati and Gutkin, 2011). Thus, discounting for food would be expected to increase when hungry, due to innate negative value associated with prolonging a state of hunger. Furthermore, actions themselves might in some cases be selected from an innately determined repertoire. Through Pavlovian conditioning, a stimulus (termed unconditioned stimulus, US, e.g., food) that elicits an innate response (the unconditioned response, e.g., salivation), can become associated with another stimulus (conditioned stimulus, CS, e.g., a tone), such that the latter subsequently becomes capable of eliciting an appropriate innate response independently (Rescorla and Solomon, 1967; Williams and Williams, 1969; Hershberger, 1986; Pavlov, 2003). Here the conditioning process, whereby CS becomes associated with US, can incorporate the cost of delay to conform to the optimal adaptations of some of the computational processes above. For example, if delivery of food follows a tone, with an intervening delay of 10 s, the “Pavlovian value” of the tone may be temporally discounted by a given proportion per unit time relative to that of the food (Domjan, 2003). Algorithmic accounts of classical conditioning, such as temporal difference learning, thus incorporate an exponential discount factor (O’Doherty et al., 2003; Moutoussis et al., 2008; Dayan, 2009; Kurth-Nelson and Redish, 2009). Exactly how such discounting is represented at a neurobiological process level remains unclear, but the influences outlined must be important. For example, the incremental process of temporal-difference learning, including Rescorla-Wagner learning (Domjan, 2003), means that the strength of the association between CS and US comes to reflect their probabilistic relationship.

Organisms can also learn the value of actions based simply on whether or not they yielded benefits in the past, referred to as instrumental conditioning (Domjan, 2003). In algorithmic terms, this can be most parsimoniously achieved by integrating the history of reinforcement following a given action, without representing an explicit model of the relationship between actions and their outcomes (Watkins and Dayan, 1992; Daw et al., 2005; Seymour et al., 2005; Schultz, 2006; Moutoussis et al., 2008; McDannald et al., 2011). This is referred to as *model-free* reinforcement learning, and corresponds to the “Thorndikian” Law of Effect (Thorndike, 1927), or “habit” learning (Dickinson et al., 1995; Ouellette, 1998; Neal, 2006; Tricomi et al., 2009; Dolan and Dayan, 2013; Orbell and Verplanken, 2014). Instrumental learning would be expected to incorporate discounting, to the extent that the environmental influences described earlier in this article affect the timecourse of reward contingent on a particular action.

Finally, biological agents can be availed of a cognitive map, or model, of the world, detailing the results of different actions and their respective values (Dickinson and Balleine, 1994; Balleine and Dickinson, 1998; Gläscher et al., 2010; Daw et al., 2011; McDannald et al., 2011). The choice of action proceeds by thinking forward through the map (or tree), and considering the consequences of alternative actions (see Seymour and Dolan, 2008). This mode of control is referred to in reinforcement

learning applications as *model-based* (Gläscher et al., 2010; Daw et al., 2011; Wunderlich et al., 2012; Smittenaar et al., 2013; Lucantonio et al., 2014), and corresponds to the definition of goal-directed behavior in animal learning as being rapidly sensitive to changes in the contingency between action and outcome, or to devaluing the outcome (Dickinson and Balleine, 1994; Balleine and Dickinson, 1998). An advantage of the model-based approach lies in its flexibility. For example, this approach is necessary to generate appropriate intertemporal choices in esoteric scenarios, to which a smooth discount function is not well adapted. For example, say a generous experimenter offers me a choice between \$100 today and \$125 4 weeks from today. The knowledge that I will be receiving my monthly pay of \$1000 exactly 4 weeks from today, and that without additional income I am likely to exceed my overdraft limit next week by around \$50, incurring a heavy fine, would likely encourage me to choose the immediate money. If I were to try choose between the immediate and delayed money according to a parametric discount function *alone*, without considering extraneous sources of (dis)utility, I might lose out to the overdraft fine. In summary, through the above innate and instrumental learning processes, given appropriate experience of the cost of delay, an organism can behave in a manner consistent with discounting without directly computing discounted value at all.

(MAL)ADAPTIVE DISCOUNTING IN PSYCHIATRIC DISORDERS

We propose that whether parametric, or revealed through the above valuation processes, discounting nevertheless represents encoding of different environmental contingencies. It is therefore noteworthy, where changes in discounting are observed, for example in psychiatric disorders, to consider such changes in light of the environment to which a given individual might be “tuned to” (see also Del Giudice, 2014). The key point here is that, the decision-maker brings to a laboratory intertemporal choice task their previous experience of delay and may also consider the rewards of the task in the context of other future outcomes they expect to receive. We consider particular instances of this below.

Mania as a State of Increased Opportunity Cost

Might steeper discounting in some pathological states reflect increased estimates of opportunity cost? In support of discounting being sensitive to changes in opportunity cost, discount rates for money have been shown to increase in line with increases in inflation (Ostaszewski et al., 1998). More speculatively, steeper discount rates in childhood and adolescence which decline into adulthood (Green et al., 1999; Chao et al., 2009; Steinberg et al., 2009) might even reflect greater potential for growth in adolescence. We propose that the pathological state of mania is associated with perceived high rates of reward and high growth potential, creating a heightened opportunity cost associated with inaction. Mania is known to be associated with impulsive behavior, such as overspending, rash financial decision-making or drug-taking (Swann, 2009),

and one study (Mason et al., 2012) finds evidence for steeper discounting in an intertemporal choice task with real-time delays in the order of seconds in individuals prone to hypomanic symptoms.

Notably growth potential creates something of a paradox. On the one hand investing reward to achieve growth implies that the decision maker has adopted a long-term view. On the other hand, having something worthwhile to invest in favors choices that obtain rewards sooner rather than later, so that they too can be invested. For example, imagine you are starting a new business venture. Whilst this is necessarily a long-term project, you might sacrifice other potential rewards, such as your health or relationships, in order to invest resources in the business, which can be seen as borrowing predicated on a high level of return from your new business. Manic individuals generate novel, and often unrealistically ambitious, goals, for example, enlisting on education courses, or indeed starting new business ventures (DSM V, 2013). We propose that these goals create high opportunity costs to delaying reward, increasing preference for immediate rewards, so as to enlist resources for goal-pursuit. This offers a putative psychological explanation for why increased impulsivity in mania (Swann, 2009), including steeper discounting (Mason et al., 2012), manifests alongside an apparent increase in goal-directed activity.

The investment in apparently long-term goals in mania seems to occur at the expense of patients correctly “playing out” or “forward modeling” future scenarios themselves. This explains why the same (mal)adaptation is found across several behavioral domains. McClure and colleagues (McClure et al., 2004, 2007) have suggested that the explicit influence of larger-later options on behavior is associated with greater cognitive control, which is reduced in mania in tandem with prefrontal activation (Murphy et al., 1999; Townsend et al., 2010). This reduction in “forward modeling” is in fact consistent—if not necessary—for the suggestion we make here to work. That is, if a person with mania were to consider in detail the path ahead leading to their goals, they would realize that the projection implicit in their growth estimate is unrealistic and they would feel able to afford to be patient. A further interesting possibility, discussed further below, is that such forward modeling itself takes time, and that in the face of high opportunity costs, the depth of such model-based strategies is reduced in favor of more rough-and-ready heuristics, or more Pavlovian or habitual responding (Dezfouli, 2009; Huys et al., 2012). Future investigations of mania might focus on measuring beliefs about growth and opportunity cost directly, and whether such beliefs correlate with changes in discounting. Interestingly, Dezfouli (2009) similarly propose that the abnormally high rewards engendered by drugs of abuse lead to an artificially elevated estimate of the average reward rate in the environment, and that this accounts for increased discounting seen amongst substance abusers (e.g., Kirby et al., 1999; Kollins, 2003; Kirby and Petry, 2004).

Finally, we have shown above how an increase in the rate of reward available from activities *other than those currently on offer* increases impatience to complete the current activity as soon as possible (i.e., increases discounting for rewards obtained from the task in hand). Niv et al. (2007) use the same approach to explain variations in response vigor. In their model they propose

that the agent can choose to reduce latency of its responses, at some energetic cost that is proportional to the latency reduction. Thus, choosing how *quickly* to perform a particular action itself becomes an intertemporal choice. As their model illustrates, greater vigor (shorter response latency) is then optimal where the average reward rate is higher, in order that agents can resume reward seeking as soon as possible. This description accords well with that of mania, where sufferers often describe the need to complete various tasks with great urgency and where the general vigor of behavior is markedly increased. Furthermore, the model of Niv and colleagues incorporates a latency-independent cost associated with switching tasks. As the authors show, at high reward rates latency-dependent costs tend to dwarf the switching cost, leading to greater task switching than at low reward rates. This too is in keeping with behavior exhibited in manic states, where sufferers have difficulty sustaining tasks.

Economic Poverty as a Deficit State

In keeping with the normative notion that deficit states increase a hazard rate for losing out on future reward, discounting indeed tends to be higher in states of monetary or physiological deficit. For example, steeper discounting is observed in individuals with lower incomes (Green et al., 1996; Reimers et al., 2009), an effect which remains after controlling for level of education. Of course, such studies are correlational, making it difficult to conclude that changes in income directly alter discounting. However, an interesting study by Callan et al. (2011) provides indirect support for a more causal role of low income in increasing discounting. The authors found that a manipulation which lead people to believe that their income was lower than their peers brought about an increase in discounting, relative to a group who were lead to believe that their income was similar to that of their peers. The manipulation was interpreted as priming personal notions of deservedness, though this might just as easily be formalized as a shift toward a perceived deficit state. In a conceptually related study Haushofer et al. (2013) performed an experiment in which subjects performed an effort task for monetary reward, after which different groups received either an increase in income from a low starting endowment, or a decrease in income from a high starting endowment. The design thus allowed the effect of (experimental) wealth changes to be dissociated from absolute wealth. Subjects' temporal discount rates were measured before and after the task, with the finding that negative income shocks lead to an increase in discounting, while positive income shocks effected a small decrease in discounting. Starting wealth was found to be unrelated to discounting. Notably, the size of an experimental endowment might not be expected to have an effect on discounting, since the endowment was likely to be small in comparison to subjects' total real-world wealth. The effect of negative income shocks, which might be interpreted as having primed an increased hazard rate for future earnings, suggests that instability in earnings, rather than simply total wealth, is an important determinant of the relationship between socioeconomic status and discounting.

A study in women deprived of food and water (for 4 h after their usual waking time) found that women given a pre-loading meal prior to testing chose an option leading to the delayed, rather than immediate, delivery of juice significantly

and significantly more so than women who had not received a preloading meal (Kirk and Logue, 1997). Also, Wang and Dvorak (2010) measured monetary discounting before and after participants drank either a sugary or a sugar-free drink (both caffeine-free), finding a significant decrease in discounting in the group who drank the sugary drink and a significant increase in the control group. This finding suggests that raising blood glucose decreases discounting, an idea congruent with increased discounting associated with deficit states.

Economic poverty may well underlie *some* of the steeper discounting seen in psychiatric disorders, through an association between mental illness and lower socioeconomic status (e.g., Weich and Lewis, 1998; Lorant et al., 2007) (however in several studies associations remain after controlling for socioeconomic characteristics). Notably, there may be an interdependent relationship between low socioeconomic status, discounting, and mental ill health, whereby impatience for rewards leads to maladaptive choices such as substance misuse, which in turn are associated with worsening finances, further increases in discounting and increased risk of psychiatric disorder (e.g., Fields et al., 2009b; Leitão et al., 2013). A similar idea has been championed by Bickel et al. (2014b), who propose that the environment associated with low socioeconomic status promotes steeper discounting, which in turn engenders unhealthy choices, thus contributing to known socioeconomic gradients in health status (Adler and Rehkopf, 2008). This is supported by evidence that cigarette smoking, obesity, alcohol use and illicit drug use all exhibit negative relationships with socioeconomic status (Conner and Norman, 2005), that these behaviors are associated with poor executive functioning (e.g., Bickel et al., 2012a), and that economic poverty is prospectively associated with poor executive functioning (Lupien et al., 2007; Noble et al., 2007; Evans and Schamberg, 2009). We discuss this interaction between environment and cognition in Section The Cost of Thinking in Economic Poverty, Borderline Personality Disorder and Schizophrenia below.

ADHD as a Deficit State

Interestingly, the effects of deprivation appear to cross modalities of reward. For example, mild opioid deprivation in opioid dependent individuals increases discounting for money as well as heroin (Giordano et al., 2002). Arguably this might be motivated by a desire on the part of subjects to obtain money sooner so as to buy drugs. However, it might equally be attributable to a more global alteration in decision-making associated with physiological deficit states (see also Loewenstein, 1996; Metcalfe and Mischel, 1999). In further support of this idea, exposure to erotic cues increases discounting for money, as well as for candy bars or soda drinks in men (Van den Bergh et al., 2008). Furthermore, the effect of sex cues to increase discounting for food and drink rewards was attenuated by satiation with money, providing evidence for a global physiological signaling mechanism. Niv et al. (2007) propose that this mechanism “global drive” mechanism might involve modulation in tonic dopamine signaling.

In some cases steeper discounting observed in psychiatric disorders might reflect processes associated with normal deficit

states. ADHD is a possible example. ADHD is defined by behavioral symptoms of inattentiveness, over-activity and impulsivity, of long-standing duration and is most commonly diagnosed in school-aged children (DSM V, 2013). Many studies have shown that children with ADHD have a greater tendency than controls to choose immediate over delayed rewards in single choices (e.g., Sonuga-Barke et al., 1992; Schweitzer and Sulzer-Azaroff, 1995; Kuntsi et al., 2001; Bitsakou et al., 2009; for reviews see Luman et al., 2005; Paloyelis et al., 2009) and (relative to controls) are biased toward choosing tasks which yield earlier, rather than delayed, reinforcement (Tripp and Alsop, 2001). Also, on delay of gratification tasks (Mischel et al., 1989) children with hyperactivity exhibit a greater tendency to terminate the delay to obtain a smaller reward, rather than waiting an allotted time for a larger reward (Rapport et al., 1986). Furthermore, several studies now report steeper monetary discounting in children with ADHD (Paloyelis et al., 2009; Scheres et al., 2010; Wilson et al., 2011; Demurie et al., 2012) or in adults with previous ADHD (Hurst et al., 2011).

We hypothesize that the increased discounting rates found in ADHD reflect both the well-known genetic vulnerability for this disorder but also encode the more deprived environments that lead to increased expression of this disorder (Apperley and Mittal, 2013; Russell et al., 2015). In support of this, in one study boys with ADHD symptoms who had been reared in deprived institutions showed increased aversion to delay compared with ADHD controls compared to less deprived patients (Loman, 2012). Thus, seeking of immediate reward in ADHD might reflect underlying mechanisms linking increased discounting with states of internal deprivation. One such mechanism would be that outlined above of higher rates of reward available from alternative tasks. For example, say that children with ADHD have an internal state resembling a deprivation of loving attention; their performance of tasks that do not offer this attention, such as quiet private study, is likely to be more impatient, so as to more quickly return to actions that do command attention from others.

Increased Estimates of Uncertainty and Hazard

Although conventional discounting tasks offer choices between rewards that are promised to be delivered with certainty, decision-makers likely come to the task with a prior belief regarding the level of hazard in the environment, and so tend to implicitly distrust the experimenter's assertion that the future rewards are guaranteed. In support of this, discount rates amongst cigarette smokers have been shown to correlate positively with their belief that the future reward will be delivered (Reynolds et al., 2007). Also, within a standard discounting questionnaire, people discount more steeply when rewards are framed as being received from fictive characters rated as untrustworthy, as opposed to from characters perceived as trustworthy (Michaelson et al., 2013).

In an interesting study, Callan et al. (2009) measured discounting in 56 undergraduate students who first watched an interview with a HIV-positive woman. One group were told that she had acquired HIV through unprotected sex and the other group that she had acquired the virus via an infected

blood transfusion. The latter group exhibited significantly steeper discounting, an effect which was proposed to result from the story of the infected blood transfusion having primed a belief that the world is unjust. A related explanation, independent of feelings of injustice *per se*, would be that the transfusion scenario increased the perceived hazard rate for adverse life events.

Finally, as described previously, the ultimate hazard is that one will die before the future reward occurs. In keeping with this, in a South African population, discounting was found to be higher amongst individuals with the lowest perceived survival probability than amongst those with average survival probability (Chao et al., 2009), and to correlate with the number of bereavements of close family members reported by North Americans (a factor putatively increasing perceived mortality risk) (Pepper and Nettle, 2013). Furthermore, discounting has been shown to increase on conscription into the Israeli army (Lahav et al., 2011), and to be higher in youths living in slums in Rio De Janeiro than in an age matched sample of university students (Ramos et al., 2013).

Populations with psychiatric disorders might well believe that future rewards are less likely to materialize (a higher hazard rate) than do healthy control populations, for quite rational reasons, given their life experiences (Hill et al., 2008). In other words, the past is the best predictor of the future, and this may be why psychiatric disorders associated with hazardous development are characterized by higher discounting rates. Populations with psychiatric illness have experienced an excess of major life events compared with the healthy population (Paykel, 1978), and have excess mortality from physical health conditions compared with the general population (Robson and Gray, 2007). The latter would be expected to be associated with lower perceived survival probability, given correlations between perceived and actual mortality in the general population (Idler and Benyamini, 1997). To our knowledge no previous studies have examined this. This may in turn result in decisions that perpetuate or worsen the disorder. Indeed, Sonuga-Barke has hypothesized that the high discounting rates measured in the laboratory in youths with conduct disorder represent an accurate—and hence adaptive in their native environment—summary of the increased hazards that these youths so commonly have experienced (Barke, 2014). An interesting possibility for future research would be to elicit beliefs of groups with psychiatric disorder about the likelihood that future reward will be forthcoming, and to regress this against their discounting choices. Similarly further research is needed to examine relationships between an individual's experience of significant life events, their confidence in the future, and their level of temporal discounting.

The Cost of Thinking in Economic Poverty, Borderline Personality Disorder and Schizophrenia

It appears that a greater engagement of model-based control, a faculty tightly dependent on working memory, is associated with more future-oriented responses on discounting paradigms. Promoting mental simulations of future outcomes by cueing participants with episodes in their lives corresponding to

the timing of the options decreases measured discount rates (Peters and Büchel, 2010). Higher working memory capacity is associated with both lower discounting (Shamosh et al., 2008), and an increased emphasis on model-based control (Eppinger et al., 2013), while working memory training in substance misusers has been found to decrease their delay discounting (Bickel et al., 2011b).

In keeping with the above, functional neuroimaging studies have found that the dorsolateral prefrontal cortex (dlPFC), an area often implicated in tasks dependent on working memory (Curtis and D'Esposito, 2003), is sensitive to model-based learning signals (Gläscher et al., 2010). This area is also known to be active when choosing delayed rewards on intertemporal choice paradigm (McClure et al., 2004, 2007). Furthermore, disrupting dlPFC function (using either transcranial magnetic stimulation or transcranial direct current stimulation) both decreases the emphasis on model-based control (Smittenaar et al., 2013) and increases temporal discounting (Hecht et al., 2013). The process of mentally simulating future outcomes is also known to be dependent on the hippocampus (Hassabis et al., 2007; Johnson et al., 2007; Schacter et al., 2008; Schacter and Schacter, 2008), and rats with hippocampal lesions have been found to exhibit increased discounting (Mariano et al., 2009). Taken together these results suggest that mental simulation of the future tends to generate more patient intertemporal choices, and that this process is working memory dependent.

A plausible explanation for the above is that mentally simulating the future resolves uncertainty about the utility of larger-later rewards (see Daw et al., 2005). For example, I might be uncertain about how much I am likely to require money in 7 months' time, but if I remember that my partner's birthday is in seven and a half months' time, and I anticipate needing the money to buy him or her an expensive present, I might revise my estimate of the utility of the future money. An interesting possibility is that decision-makers face a trade-off between making the best possible decisions and doing so in a timely manner with the minimum of effort. Model-based simulation of the future is computationally costly, i.e., consumes time and energy. If conditions are sufficiently unpredictable, then attempting to explicitly plan out future possibilities is futile, and may even be disadvantageous (see Daw et al., 2005). Thus, prolonged exposure to an unstable environment during development ought to both discourage the use of model-based strategies and increase discounting via greater uncertainty associated with future rewards. This possibility would conceptually bind together an unstable childhood environment, diminished cognitive ability and steeper discounting of reward, providing a tentative theoretical basis for explaining the association between these factors in several psychiatric disorders. For example, people with borderline personality disorder are likely to have experienced childhood abuse (Lewis and Christopher, 1989; Ogata et al., 1990; Zanarini et al., 1997), exhibit below average cognitive function (Swirsky-Sacchetti et al., 1993) and discount the future more steeply than healthy controls (Lawrence et al., 2010).

A similar interaction might in part underlie associations between low socioeconomic status, steeper discounting and

psychiatric disorder. Bickel et al. (2014a, 2011a) propose a neuropsychological explanation for relationships between low socioeconomic status and unhealthy lifestyle choices, in terms of a dual-systems model of cognition, whereby low socioeconomic status encourages engagement of a more “impulsive” decision-making system, putatively mediated by limbic brain structures, over an “executive” decision-making system, mediated by parts of frontal cortex. The authors point to evidence that several neurocognitive abilities including working memory, declarative memory, and cognitive control exhibit socioeconomic gradients (Noble et al., 2007). This association appears to hold in prospective analyses too. On a developmental timescale, Evans and Schamberg (2009) show that childhood poverty predicts lower working memory in young adulthood, and that high levels of childhood stress mediate this relationship. State-based effects of poverty on cognitive function are also evident, for example Indian sugar-cane farmers exhibit worse cognitive performance before their harvest, when they are poor, than after their harvest, when they are richer, even controlling for levels of stress (Mani et al., 2013). The dual-systems approach is not incompatible with our three-way division of behavioral control. The model-based system for instance appears to depend on executive functions such as working memory, but has the advantage of carrying a specific algorithmic meaning. Also, we envisage the three-controllers as sharing the mutual goal of maximizing reward (Dayan et al., 2006), and suggest that their relative deployment is also subject to a cost-benefit trade-off (Daw et al., 2005; Dezfouli, 2009; Huys et al., 2012). We therefore go as far as to propose that diminished deployment of model-based control in states of deprivation might reflect an evolutionary milieu in which such changes were approximately optimal, for example in response to irreducible future uncertainty.

Deficits in future thinking appear likely to underlie steeper discounting seen in patients diagnosed with schizophrenia compared with healthy controls (Heerey et al., 2007, 2011), in keeping with observations that such patients often exhibit cognitive and executive dysfunction. Furthermore, patients with schizophrenia exhibit atrophy of frontal and temporal brain regions (Madsen et al., 1999; Velakoulis et al., 2001; van Haren et al., 2008), a pattern which would be expected to be accompanied by shortened time perspective, given the role of these structures in imagining future scenarios (Hassabis et al., 2007; Johnson et al., 2007; Schacter et al., 2008; Schacter and Schacter, 2008). Heerey et al. (2011) present evidence to support this view, comparing measures of discounting, cognitive function and “future representation” in 39 patients with schizophrenia and 25 healthy control participants. Patients discounted more steeply than controls, and when asked to list events which they thought might happen to them in their lives, on average reported future life-events that were nearer in time. This shortened future perspective correlated with lower working memory scores in both patients and controls, to the extent that controlling for working memory abolished the effect of schizophrenia status on discounting. These results suggest that discounting deficits in schizophrenia are attributable to an impaired ability to imagine the future, a faculty that is limited by working memory capacity.

FUTURE DIRECTIONS

The above account leaves considerable room for future research. The foregoing discussion has largely focused on appetitive processes evoked in the appraisal of future *rewards*. A complementary, but distinct, set of principles might apply to how humans evaluate future punishment. For example, as a complement to the theory that tonic dopamine signals the average reward rate, it has been proposed that tonic serotonin signals the long run average punishment rate, and thus controls the vigor of avoidance behavior (Dayan, 2012a,b, see also Crockett et al., 2012). This idea might hold relevance for increased discounting in depression, which is associated with both marked avoidance (Ferster, 1973) and possible serotonergic abnormalities (e.g., Mann et al., 2000). Although a normative account of the role of serotonin in depression remains elusive, it is interesting that decreasing serotonin availability (achieved by tryptophan depletion) in healthy subjects acts to increase discounting (Tanaka et al., 2007; Schweighofer et al., 2008), commensurate with increased discounting seen in depression (Takahashi et al., 2008; Dennhardt and Murphy, 2011; Dombrowski et al., 2011, 2012; Imhoff et al., 2014; Pulcu et al., 2014) (For further discussion of temporal preferences for punishment see Berns et al., 2006; Story et al., 2013, 2015).

A further area for future research concerns the effect of stress on discounting (e.g., Diller et al., 2011; Kimura et al., 2013). A recent meta-analysis (Fields et al., 2014) of 16 studies examining the relationships between delay discounting or delay of gratification and subjective or physiological measures of stress and found that stress was associated with steeper discounting, with a large aggregate effect size (Hedge's $g = 0.59$). Seemingly contradicting these findings, low baseline cortisol levels have been associated with *increased* delay discounting (Takahashi, 2004), and similarly predict higher discounting at 6 month follow up (Takahashi et al., 2009). A possible explanation would be that baseline stress and responsivity to stress manipulations exert distinct influences on discounting. In part supporting this idea, Lempert et al. (2012) found that when placed under stressful conditions, individuals with low trait perceived stress showed higher discounting than those with high trait perceived stress, perhaps reflecting greater responsiveness to acute stressors in subjects with low trait stress. In addition acute administration of hydrocortisone, a key hormone involved in stress response, has been found to cause a short-lived increase in discounting (Cornelisse et al., 2013). Further work is required to understand the relationships between baseline and induced stress and their interaction with discounting, as well as to characterize stress in terms of the information content of stressful situations.

The above account has not specifically addressed willpower. Several lines of evidence point to the fact that humans often renege on best-laid plans, in favor of immediate consumption. We propose that this results since people are poor in predicting in advance the effect of conditioned cues and motivational state changes on their behavior (see also Loewenstein, 1996; Metcalfe and Mischel, 1999; Read, 2001; Chapman, 2005; Dayan et al., 2006; Story et al., 2014). Thus, one might plan to abstain from eating dessert as part of a diet plan, but find it

harder to resist when presented with a piece of cake (see for example Read and Van Leeuwen, 1998; Allan et al., 2010) and relapses in drug-taking behavior following abstinence commonly occur after exposure to a previous drug-taking environment (O'Brien et al., 1998). Similarly, people appear poor in predicting their behavior in future motivational states that differ from their current motivational state. For example, in a study of analgesic preferences for childbirth (Christensen-Szalanski, 1984), women asked roughly 1 month in advance of labor preferred to avoid invasive spinal anesthesia in favor of less invasive but less effective pain relief methods, however during active labor women frequently reversed preference and opted for anesthesia. "Battles of will" then consist in the attempt to punish or extinguish existing habitual or Pavlovian responses through the imposition of countervailing model-based (goal-directed) valuations. Hyperbolic discounting theoretically gives rise to similar intertemporal choice conflicts, but considered alone has difficulty accounting for the state-dependence of real world failures of self-control. Thus, in the study of Christensen-Szalanski (1984) it seems likely to be the transition into a painful state that brings about a shift in women's preferences for analgesia, rather than the time preceding childbirth *per se* as hyperbolic discounting would suggest. An interesting direction for future research will be to examine whether individuals with psychiatric disorders, for example borderline personality disorder, exhibit greater choice inconsistency over time, relative to controls. This possibility would accord with a well-esteemed theory that individuals with borderline personality disorder are impaired in modeling mental states (Bateman and Fonagy, 2004).

Another interesting direction not explored here concerns discounting of *past* rewards (Yi et al., 2006; Bickel et al., 2008). Discounting for past rewards has been shown to be systematic and hyperbolic in form, and is correlated with the degree of future discounting across individuals (Yi et al., 2006). Furthermore, cigarette smokers are found to discount past, as well as future, rewards more steeply than non-smokers (Bickel et al., 2008). Symmetry between past and future discounting is in keeping with evidence that remembering the past and imagining the future are both dependent on the hippocampus (Hassabis et al., 2007; Johnson et al., 2007; Schacter et al., 2008; Schacter and Schacter, 2008). Notably past discounting is difficult to directly account for in terms of some of the informational influences suggested in this article. Growth potential for example ought to motivate having received rewards in the distant past, since these should have had time to accrue greater value. Further work is clearly needed to understand the possible normative basis of past discounting. One possibility is that factors tending to foreshorten model-based consideration of future outcomes, such as uncertainty, also diminish retrieval of episodic memories, leading to a narrowing of temporal perspective. Notably, the learning rate in model-free reinforcement learning algorithms corresponds to an exponential discount factor for past reward. Yechiam et al. (2005) have shown that substance misusers and individuals with ventral medial prefrontal cortex lesions both exhibit increased learning rates on the Iowa gambling task, where an excessive focus on recent reinforcement is disadvantageous. This suggests that high learning rates might reflect a form of

"retrospective impulsivity," through assigning too little weight to distant past experience. Further work is required to explore this possibility.

A final consideration is that of how discounting differs between different forms of outcome. Discounting for several forms of appetitive outcome shows consistency across individuals, for example discount rates for money are strongly and significantly correlated with other forms of appetitive outcome, such as the discounting of cigarettes for cigarette smokers, the discounting of heroin for opioid-dependent outpatients and the discounting of food amongst college students (Odum, 2011; Pearson $r = 0.93$; $p = 0.0007$ for money vs. the mean of all other outcomes). However, rates are not *identical* across commodities: people tend to discount primary reinforcers such as food, water and sex more steeply than money (Lawyer et al., 2010; Odum, 2011; Jarmolowicz et al., 2013) and a number of studies have shown that people with substance dependence discount their drug of abuse more steeply than money (e.g., Madden et al., 1997; Bickel et al., 1999; Petry, 2001). Steeper discounting for primary reinforcers might reflect their greater engagement of innate appetitive systems. In other words, deliberative consideration of primary reinforcers might increase attention to the relevant underlying deficit state (drive). Steeper discounting then putatively results due to the negative Pavlovian value associated with prolonging the deficit state. Further research is needed to examine this possibility.

Interesting results have been obtained when discounting choices are made across different commodities, for example in choices between money now vs. cigarettes later, termed cross-commodity discounting (CCD), as opposed to single-commodity discounting (SCD). For instance, Bickel et al. (2011a, 2007) examined discounting in cocaine-dependent individuals between cocaine now vs. cocaine later (C-C), money now vs. money later (M-M), cocaine now vs. money later (C-M), and money now vs. cocaine later (M-C) conditions, where the amounts of money and cocaine across conditions were equated in immediate worth. Consistent with previous findings, C-C discount rates were significantly greater than M-M discount rates; indeed there was a significant main effect of changing the delayed commodity to cocaine, consistent with cocaine being discounted more steeply than money. However, the authors found that, whilst C-M and M-M discounting were statistically indistinguishable, M-C discount rates were significantly higher than C-C discount rates. Wesley et al. (2014) broadly replicate this result, and Jarmolowicz et al. (2014) find a similar pattern of findings for money vs. sex CCD, wherein a M-S condition was associated with the steepest discounting. A possible explanation in terms of the classical economic model would be that cocaine (or sex) is both discounted more steeply and has a *less* concave utility function than money. Bickel et al. (2011a, 2007) illustrate this possibility though favor an explanation in terms of a framing effect. We propose a framing hypothesis whereby primary reinforcers are associated with a steeper implicit hazard rate than money (this might in part underlie their steeper discounting, but is of itself insufficient to explain the above findings); SCD then hypothetically diminishes the implicit hazard rate, by priming the idea that the commodity will definitely be received

sooner-or-later. By contrast, the implicit exchange of money for primary reinforcement in CCD hypothetically amplifies the hazard rate for the delayed commodity, by priming the notion that the delayed commodity is not guaranteed. This hypothesis leads to the observed interaction, with the steepest discounting for CCD in which primary reinforcement is delayed, and is an eminently testable. The possible modulation of such cross-commodity effects in various psychiatric disorders might offer further clues as to the underlying decision mechanisms at play.

In summary we have reviewed motivations for steeper discounting of delayed reward. Discounting tends to be increased across a broad range of disorders, including ADHD, schizophrenia, bipolar disorder, hypomania, depression, borderline personality disorder and substance misuse disorders. We have proposed that these findings can be parsimoniously understood by examining the *reasons* why people should discount the future, namely the opportunity costs of delay, uncertainty associated with future outcomes and the cognitive costs of resolving this uncertainty. We have detailed different types of *information processing* in the brain that can take these factors into account, broadly distinguishing

“parametric discounting,” whereby rewards labeled as delayed are automatically discounted as a function of delay, vs. “planful discounting” where the factors associated with the delay are accounted for in the course of learning. Where possible we have attempted to map these normative influences onto putative, albeit broad neurobiological *mechanisms*. More generally we propose that this approach, that is, attempting to understand the biological substrates of psychiatric disorder in terms of their physiological *function*, and in light of a person’s life history, is key to bridging psychosocial and biological conceptions of mental illness. We accept that our use of this approach here might appear speculative. In essence, we feel is this justified given the emerging nature of the field and await further research developments with eager interest.

FUNDING STATEMENT

This work was supported by the Wellcome Trust [Ray Dolan Senior Investigator Award 098362/Z/12/Z]. The Wellcome Trust Centre for Neuroimaging is supported by core funding from the Wellcome Trust 091593/Z/10/Z. Dr. Moutoussis is also supported by the UCLH Biomedical Research Council.

REFERENCES

- Adler, N. E., and Rehkopf, D. H. (2008). US disparities in health: descriptions, causes, and mechanisms. *Annu. Rev. Public Health* 29, 235–252. doi: 10.1146/annurev.publhealth.29.020907.090852
- Ahn, W.-Y., Rass, O., Fridberg, D. J., Bishara, A. J., Forsyth, J. K., Breier, A., et al. (2011). Temporal discounting of rewards in patients with bipolar disorder and schizophrenia. *J. Abnorm. Psychol.* 120, 911. doi: 10.1037/a0023333
- Ainslie, G. W. (1974). Impulse control in pigeons. *J. Exp. Anal. Behav.* 21, 485–489. doi: 10.1901/jeb.1974.21-485
- Allan, J. L., Johnston, M., and Campbell, N. (2010). Unintentional eating: what determines goal-incongruent chocolate consumption? *Appetite* 54, 422–425. doi: 10.1016/j.appet.2010.01.009
- Anderhub, V., Güth, W., Gneezy, U., and Sonsino, D. (2001). On the interaction of risk and time preferences: an experimental study. *Ger. Econ. Rev.* 2, 239–253. doi: 10.1111/1468-0475.00036
- Andersen, S., Harrison, G. W., Lau, M. I., and Rutström, E. E. (2008). Eliciting risk and time preferences. *Econometrica* 76, 583–618. doi: 10.1111/j.1468-0262.2008.00848.x
- Apperley, L., and Mittal, R. (2013). G207 is there a link between ADHD and social deprivation? *Arch. Dis. Child.* 98, A92–A93. doi: 10.1136/archdischild-2013-304107.219
- Avsar, K. B., Weller, R. E., Cox, J. E., Reid, M. A., White, D. M., and Lahti, A. C. (2013). An fMRI investigation of delay discounting in patients with schizophrenia. *Brain Behav.* 3, 384–401. doi: 10.1002/brb3.135
- Balleine, B. W., and Dickinson, A. (1998). Goal-directed instrumental action: contingency and incentive learning and their cortical substrates. *Neuropharmacology* 37, 407–419. doi: 10.1016/S0028-3908(98)00033-1
- Barke, E. (2014). Temporal discounting in conduct disorder: toward an experience-adaptation hypothesis of the role of psychosocial insecurity. *J. Pers. Disord.* 28, 19–24. doi: 10.1521/pedi.2014.28.1.19
- Barkley, R. A., Edwards, G., Laneri, M., Fletcher, K., and Metevia, L. (2001). Executive functioning, temporal discounting, and sense of time in adolescents with attention deficit hyperactivity disorder (ADHD) and oppositional defiant disorder (ODD). *J. Abnorm. Child Psychol.* 29, 541–556. doi: 10.1023/A:1012233310098
- Bateman, A. W., and Fonagy, P. (2004). Mentalization-based treatment of BPD. *J. Pers. Disord.* 18, 36–51. doi: 10.1521/pedi.18.1.36.32772
- Behrens, T. E., Woolrich, M. W., Walton, M. E., and Rushworth, M. F. (2007). Learning the value of information in an uncertain world. *Nat. Neurosci.* 10, 1214–1221. doi: 10.1038/nn1954
- Berns, G. S., Chappelow, J., Cekic, M., Zink, C. F., Pagnoni, G., and Martin-Skurski, M. E. (2006). Neurobiological substrates of dread. *Science* 312, 754–758. doi: 10.1126/science.1123721
- Bickel, W. K., Jarmolowicz, D. P., Mueller, E. T., Gatchalian, K. M., and McClure, S. M. (2012a). Are executive function and impulsivity antipodes? A conceptual reconstruction with special reference to addiction. *Psychopharmacology* 221, 361–387. doi: 10.1007/s00213-012-2689-x
- Bickel, W. K., Jarmolowicz, D. P., Mueller, E. T., Koffarnus, M. N., and Gatchalian, K. M. (2012b). Excessive discounting of delayed reinforcers as a trans-disease process contributing to addiction and other disease-related vulnerabilities: emerging evidence. *Pharmacol. Ther.* 134, 287–297. doi: 10.1016/j.pharmthera.2012.02.004
- Bickel, W. K., Koffarnus, M. N., Moody, L., and Wilson, A. G. (2014a). The behavioral and neuro-economic process of temporal discounting: a candidate behavioral marker of addiction. *Neuropharmacology* 76, 518–527. doi: 10.1016/j.neuropharm.2013.06.013
- Bickel, W. K., Landes, R. D., Christensen, D. R., Jackson, L., Jones, B. A., Kurth-Nelson, Z., et al. (2011a). Single and cross-commodity discounting among cocaine addicts: the commodity and its temporal location determine discounting rate. *Psychopharmacology* 217, 177–187. doi: 10.1007/s00213-011-2272-x
- Bickel, W. K., Miller, M. L., Yi, R., Kowal, B. P., Lindquist, D. M., and Pitcock, J. A. (2007). Behavioral and neuroeconomics of drug addiction: competing neural systems and temporal discounting processes. *Drug Alcohol Depend.* 90(Suppl. 1), S85–S91. doi: 10.1016/j.drugalcdep.2006.09.016
- Bickel, W. K., Moody, L., Quisenberry, A. J., Ramey, C. T., and Sheffer, C. E. (2014b). A competing neurobehavioral decision systems model of SES-related health and behavioral disparities. *Prev. Med.* 68, 37–43. doi: 10.1016/j.ypmed.2014.06.032
- Bickel, W. K., Odum, A. L., and Madden, G. J. (1999). Impulsivity and cigarette smoking: delay discounting in current, never, and ex-smokers. *Psychopharmacology* 146, 447–454.
- Bickel, W. K., Yi, R., Kowal, B. P., and Gatchalian, K. M. (2008). Cigarette smokers discount past and future rewards symmetrically and more than controls: is

- discounting a measure of impulsivity? *Drug Alcohol Depend.* 96, 256–262. doi: 10.1016/j.drugalcdep.2008.03.009
- Bickel, W. K., Yi, R., Landes, R. D., Hill, P. F., and Baxter, C. (2011b). Remember the future: working memory training decreases delay discounting among stimulant addicts. *Biol. Psychiatry* 69, 260–265. doi: 10.1016/j.biopsych.2010.08.017
- Bitsakou, P., Psychogiou, L., Thompson, M., and Sonuga-Barke, E. J. (2009). Delay aversion in attention deficit/hyperactivity disorder: an empirical investigation of the broader phenotype. *Neuropsychologia* 47, 446–456. doi: 10.1016/j.neuropsychologia.2008.09.015
- Callan, M. J., Shead, N. W., and Olson, J. M. (2009). Foregoing the labor for the fruits: the effect of just world threat on the desire for immediate monetary rewards. *J. Exp. Soc. Psychol.* 45, 246–249. doi: 10.1016/j.jesp.2008.08.013
- Callan, M. J., Shead, N. W., and Olson, J. M. (2011). Personal relative deprivation, delay discounting, and gambling. *J. Pers. Soc. Psychol.* 101, 955. doi: 10.1037/a0024778
- Chao, L. W., Szrek, H., Pereira, N. S., and Pauly, M. V. (2009). Time preference and its relationship with age, health, and survival probability. *Judgm. Decis. Mak.* 4, 1–19.
- Chapman, G. B. (2005). Short-term cost for long-term benefit: time preference and cancer control. *Health Psychol.* 24, S41–S48. doi: 10.1037/0278-6133.24.4.s41
- Christensen-Szalanski, J. J. (1984). Discount functions and the measurement of patients' values. Women's decisions during childbirth. *Med. Decis. Making* 4, 47–58.
- Chung, S. H., and Herrnstein, R. J. (1967). Choice and delay of reinforcement. *J. Exp. Anal. Behav.* 10, 67–74. doi: 10.1901/jeab.1967.10-67
- Coffey, S. F., Schumacher, J. A., Baschnagel, J. S., Hawk, L. W., and Holloman, G. (2011). Impulsivity and risk-taking in borderline personality disorder with and without substance use disorders. *Pers. Disord. Theory Res. Treat.* 2, 128. doi: 10.1037/a0020574
- Conner, M., and Norman, P. (2005). *Predicting Health Behaviour, 2nd Edn.* Maidenhead: McGraw-Hill International.
- Cornelisse, S., Van Ast, V., Haushofer, J., Seinstra, M., and Joels, M. (2013). *Time-Dependent Effect of Hydrocortisone Administration on Intertemporal Choice.* Available online at: <http://ssrn.com/abstract=2294189>
- Critchfield, T. S., and Kollins, S. H. (2001). Temporal discounting: basic research and the analysis of socially important behavior. *J. Appl. Behav. Anal.* 34, 101–122. doi: 10.1901/jaba.2001.34-101
- Crockett, M. J., Clark, L., Apergis-Schoute, A. M., Morein-Zamir, S., and Robbins, T. W. (2012). Serotonin modulates the effects of Pavlovian aversive predictions on response vigor. *Neuropsychopharmacology* 37, 2244–2252. doi: 10.1038/npp.2012.75
- Curtis, C. E., and D'Esposito, M. (2003). Persistent activity in the prefrontal cortex during working memory. *Trends Cogn. Sci.* 7, 415–423. doi: 10.1016/S1364-6613(03)00197-9
- Daw, N. D., Gershman, S. J., Seymour, B., Dayan, P., and Dolan, R. J. (2011). Model-based influences on humans' choices and striatal prediction errors. *Neuron* 69, 1204–1215. doi: 10.1016/j.neuron.2011.02.027
- Daw, N. D., Niv, Y., and Dayan, P. (2005). Uncertainty-based competition between prefrontal and dorsolateral striatal systems for behavioral control. *Nat. Neurosci.* 8, 1704–1711. doi: 10.1038/nn1560
- Daw, N. D., and Touretzky, D. S. (2000). Behavioral considerations suggest an average reward TD model of the dopamine system. *Neurocomputing* 32–33, 679–684. doi: 10.1016/S0925-2312(00)00232-0
- Dayan, P. (2009). Prospective and retrospective temporal difference learning. *Network* 20, 32–46. doi: 10.1080/09548980902759086
- Dayan, P. (2012a). "Models of value and choice," in *Neuroscience of Preference and Choice*, eds R. Dolan and T. Sharot (London: Elsevier), 33–59. doi: 10.1016/B978-0-12-381431-9.00002-4
- Dayan, P. (2012b). Instrumental vigour in punishment and reward. *Eur. J. Neurosci.* 35, 1152–1168. doi: 10.1111/j.1460-9568.2012.08026.x
- Dayan, P., and Balleine, B. W. (2002). Reward, motivation, and reinforcement learning. *Neuron* 36, 285–298. doi: 10.1016/S0896-6273(02)00963-7
- Dayan, P., and Daw, N. D. (2008). Decision theory, reinforcement learning, and the brain. *Cogn. Affect. Behav. Neurosci.* 8, 429–453. doi: 10.3758/CABN.8.4.429
- Dayan, P., Niv, Y., Seymour, B., and Daw, N. (2006). The misbehavior of value and the discipline of the will. *Neural Netw.* 19, 1153–1160. doi: 10.1016/j.neunet.2006.03.002
- Del Giudice, M. (2014). An evolutionary life history framework for psychopathology. *Psychol. Inq.* 25, 261–300. doi: 10.1080/1047840X.2014.884918
- Demurie, E., Roeyers, H., Baeyens, D., and Sonuga-Barke, E. (2012). Temporal discounting of monetary rewards in children and adolescents with ADHD and autism spectrum disorders. *Dev. Sci.* 15, 791–800. doi: 10.1111/j.1467-7687.2012.01178.x
- Denk, F., Walton, M., Jennings, K., Sharp, T., Rushworth, M., and Bannerman, D. (2005). Differential involvement of serotonin and dopamine systems in cost-benefit decisions about delay or effort. *Psychopharmacology* 179, 587–596. doi: 10.1007/s00213-004-2059-4
- Dennhardt, A. A., and Murphy, J. G. (2011). Associations between depression, distress tolerance, delay discounting, and alcohol-related problems in European American and African American college students. *Psychol. Addict. Behav.* 25, 595. doi: 10.1037/a0025807
- Dezfouli, A. (2009). A neurocomputational model for cocaine addiction. *Neural Comput.* 21, 2869. doi: 10.1162/neco.2009.10-08-882
- Dickinson, A., and Balleine, B. (1994). Motivational control of goal-directed action. *Learn. Behav.* 22, 1–18. doi: 10.3758/BF03199951
- Dickinson, A., Balleine, B., Watt, A., Gonzalez, F., and Boakes, R. (1995). Motivational control after extended instrumental training. *Anim. Learn. Behav.* 23, 197–206. doi: 10.3758/BF03199935
- Dierst-Davies, R., Reback, C. J., Peck, J. A., Nuño, M., Kamien, J. B., and Amass, L. (2011). Delay-discounting among homeless, out-of-treatment, substance-dependent men who have sex with men. *Am. J. Drug Alcohol. Abuse* 37, 93–97. doi: 10.3109/00952990.2010.540278
- Diller, J. W., Patros, C. H., and Prentice, P. R. (2011). Temporal discounting and heart rate reactivity to stress. *Behav. Processes* 87, 306–309. doi: 10.1016/j.beproc.2011.05.001
- Dohmen, T., Falk, A., Huffman, D., and Sunde, U. (2010). Are risk aversion and impatience related to cognitive ability? *Am. Econ. Rev.* 100, 1238–1260. doi: 10.1257/aer.100.3.1238
- Dolan, R. J., and Dayan, P. (2013). Goals and Habits in the Brain. *Neuron* 80, 312–325. doi: 10.1016/j.neuron.2013.09.007
- Dom, G., De Wilde, B., Hulstijn, W., Van Den Brink, W., and Sabbe, B. (2006a). Behavioural aspects of impulsivity in alcoholics with and without a cluster-B personality disorder. *Alcohol. Alcohol.* 41, 412–420. doi: 10.1093/alcalc/agl030
- Dom, G., De Wilde, B., Hulstijn, W., Van Den Brink, W., and Sabbe, B. (2006b). Decision-making deficits in alcohol-dependent patients with and without comorbid personality disorder. *Alcohol. Clin. Exp. Res.* 30, 1670–1677. doi: 10.1111/j.1530-0277.2006.00202.x
- Dombrovski, A. Y., Siegle, G. J., Szanto, K., Clark, L., Reynolds, C., and Aizenstein, H. (2012). The temptation of suicide: striatal gray matter, discounting of delayed rewards, and suicide attempts in late-life depression. *Psychol. Med.* 42, 1203–1215. doi: 10.1017/S0033297111002133
- Dombrovski, A. Y., Szanto, K., Siegle, G. J., Wallace, M. L., Forman, S. D., Sahakian, B., et al. (2011). Lethal forethought: delayed reward discounting differentiates high-and low-lethality suicide attempts in old age. *Biol. Psychiatry* 70, 138–144. doi: 10.1016/j.biopsych.2010.12.025
- Domjan, M. (2003). *The Principles of Learning and Behaviour, 5th Edn.* Belmont, CA: Wadsworth.
- Dougherty, D. M., Bjork, J. M., Huckabee, H. C., Moeller, F. G., and Swann, A. C. (1999). Laboratory measures of aggression and impulsivity in women with borderline personality disorder. *Psychiatry Res.* 85, 315–326. doi: 10.1016/S0165-1781(99)00011-6
- DSM V (2013). *Diagnostic and Statistical Manual of Mental Disorders, 5th Edn.* (DSM-5 TM). Washington, DC: American Psychiatric Association.
- Eckel, C., Johnson, C., and Montmarquette, C. (2005). "Savings decisions of the working poor: short- and long-term horizons," in *Field Experiments in Economics, Research in Experimental Economics*, Vol. 10, eds J. Carpenter, G. W. Harrison, and J. A. List (Greenwich, CT: JAI Press), 219–260. doi: 10.1016/S0193-2306(04)10006-9
- Eppinger, B., Walter, M., Heekeren, H. R., and Li, S.-C. (2013). Of goals and habits: age-related and individual differences in goal-directed decision-making. *Front. Neurosci.* 7:253. doi: 10.3389/fnins.2013.00253
- Epstein, L. H., Richards, J. B., Saad, F. G., Paluch, R. A., Roemmich, J. N., and Lerman, C. (2003). Comparison between two measures of delay discounting

- in smokers. *Exp. Clin. Psychopharmacol.* 11, 131–138. doi: 10.1037/1064-1297.11.2.131
- Evans, G. W., and Schamberg, M. A. (2009). Childhood poverty, chronic stress, and adult working memory. *Proc. Natl. Acad. Sci. U.S.A.* 106, 6545–6549. doi: 10.1073/pnas.0811910106
- Ferster, C. B. (1973). A functional analysis of depression. *Am. Psychol.* 28, 857. doi: 10.1037/h0035605
- Field, M., Christiansen, P., Cole, J., and Goudie, A. (2007). Delay discounting and the alcohol Stroop in heavy drinking adolescents. *Addiction* 102, 579–586. doi: 10.1111/j.1360-0443.2007.01743.x
- Fields, S. A., Lange, K., Ramos, A., Thamotharan, S., and Rassu, F. (2014). The relationship between stress and delay discounting: a meta-analytic review. *Behav. Pharmacol.* 25, 434–444. doi: 10.1097/fbp.0000000000000044
- Fields, S., Collins, C., Leraas, K., and Reynolds, B. (2009a). Dimensions of impulsive behavior in adolescent smokers and nonsmokers. *Exp. Clin. Psychopharmacol.* 17, 302–311. doi: 10.1037/a0017185
- Fields, S., Leraas, K., Collins, C., and Reynolds, B. (2009b). Delay discounting as a mediator of the relationship between perceived stress and cigarette smoking status in adolescents. *Behav. Pharmacol.* 20, 455–460. doi: 10.1097/FBP.0b013e328330dcff
- Fishburn, P. C., and Rubinstein, A. (1982). Time preference. *Int. Econ. Rev.* 23, 677–694. doi: 10.2307/2526382
- Floresco, S. B., Maric, T., and Ghods-Sharifi, S. (2008). Dopaminergic and glutamatergic regulation of effort-and delay-based decision making. *Neuropsychopharmacology* 33, 1966–1979. doi: 10.1038/sj.npp.1301565
- Frederick, S., and Loewenstein, G. (2008). Conflicting motives in evaluations of sequences. *J. Risk Uncert.* 37, 221–235. doi: 10.1007/s11166-008-9051-z
- Frederick, S., Loewenstein, G., and O'Donoghue, T. (2002). Time discounting and time preference: a critical review. *J. Econ. Lit.* 40, 351–401. doi: 10.1257/jel.40.2.351
- Friston, K. J., Stephan, K. E., Montague, R., and Dolan, R. J. (2014). Computational psychiatry: the brain as a phantastic organ. *Lancet Psychiatry* 1, 148–158. doi: 10.1016/S2215-0366(14)70275-5
- Friston, K., Schwartenbeck, P., FitzGerald, T., Moutoussis, M., Behrens, T., and Dolan, R. J. (2013). The anatomy of choice: active inference and agency. *Front. Hum. Neurosci.* 7:598. doi: 10.3389/fnhum.2013.00598
- Giordano, L. A., Bickel, W. K., Loewenstein, G., Jacobs, E. A., Marsch, L., and Badger, G. J. (2002). Mild opioid deprivation increases the degree that opioid-dependent outpatients discount delayed heroin and money. *Psychopharmacology* 163, 174–182. doi: 10.1007/s00213-002-1159-2
- Gläscher, J., Daw, N., Dayan, P., and O'Doherty, J. P. (2010). States versus rewards: dissociable neural prediction error signals underlying model-based and model-free reinforcement learning. *Neuron* 66, 585–595. doi: 10.1016/j.neuron.2010.04.016
- Green, L., Fristoe, N., and Myerson, J. (1994). Temporal discounting and preference reversals in choice between delayed outcomes. *Psychon. Bull. Rev.* 1, 383–389. doi: 10.3758/BF03213979
- Green, L., and Myerson, J. (2004). A discounting framework for choice with delayed and probabilistic rewards. *Psychol. Bull.* 130, 769–792. doi: 10.1037/0033-2909.130.5.769
- Green, L., Myerson, J., Lichtman, D., Rosen, S., and Fry, A. (1996). Temporal discounting in choice between delayed rewards: the role of age and income. *Psychol. Aging* 11:79. doi: 10.1037/0882-7974.11.1.79
- Green, L., Myerson, J., and Ostaszewski, P. (1999). Discounting of delayed rewards across the life span: age differences in individual discounting functions. *Behav. Processes* 46, 89–96. doi: 10.1016/S0376-6357(99)00021-2
- Hassabis, D., Kumaran, D., Vann, S. D., and Maguire, E. A. (2007). Patients with hippocampal amnesia cannot imagine new experiences. *Proc. Natl. Acad. Sci. U.S.A.* 104, 1726–1731. doi: 10.1073/pnas.0610561104
- Haushofer, J., Cornelisse, S., Seinstra, M., Fehr, E., Joëls, M., and Kalenscher, T. (2013). No effects of psychosocial stress on intertemporal choice. *PLoS ONE* 8:e78597. doi: 10.1371/journal.pone.0078597
- Hecht, D., Walsh, V., and Lavidor, M. (2013). Bi-frontal direct current stimulation affects delay discounting choices. *Cogn. Neurosci.* 4, 7–11. doi: 10.1080/17588928.2011.638139
- Heerey, E. A., Matveeva, T. M., and Gold, J. M. (2011). Imagining the future: degraded representations of future rewards and events in schizophrenia. *J. Abnorm. Psychol.* 120, 483. doi: 10.1037/a0021810
- Heerey, E. A., Robinson, B. M., McMahon, R. P., and Gold, J. M. (2007). Delay discounting in schizophrenia. *Cogn. Neuropsychiatry* 12, 213–221. doi: 10.1080/13546800601005900
- Hershberger, W. A. (1986). An approach through the looking-glass. *Anim. Learn. Behav.* 14, 443–451. doi: 10.3758/BF03200092
- Hill, E. M., Jenkins, J., and Farmer, L. (2008). Family unpredictability, future discounting, and risk taking. *J. Socio-Econ.* 37, 1381–1396. doi: 10.1016/j.socsec.2006.12.081
- Holt, C. A., and Laury, S. K. (2002). Risk aversion and incentive effects. *Am. Econ. Rev.* 92, 1644–1655. doi: 10.1257/000282802762024700
- Hurst, R. M., Kepley, H. O., McCalla, M. K., and Livermore, M. K. (2011). Internal consistency and discriminant validity of a delay-discounting task with an adult self-reported ADHD sample. *J. Atten. Disord.* 15, 412–422. doi: 10.1177/1087054710365993
- Huys, Q. J., Eshel, N., O'Nions, E., Sheridan, L., Dayan, P., and Roiser, J. P. (2012). Bonsai trees in your head: how the Pavlovian system sculpts goal-directed choices by pruning decision trees. *PLoS Comput. Biol.* 8:e1002410. doi: 10.1371/journal.pcbi.1002410
- Huys, Q. J., Moutoussis, M., and Williams, J. (2011). Are computational models of any use to psychiatry? *Neural Netw.* 24, 544–551. doi: 10.1016/j.neunet.2011.03.001
- Idler, E. L., and Benyamini, Y. (1997). Self-rated health and mortality: a review of twenty-seven community studies. *J. Health Soc. Behav.* 38, 21–37. doi: 10.2307/2955359
- Imhoff, S., Harris, M., Weiser, J., and Reynolds, B. (2014). Delay discounting by depressed and non-depressed adolescent smokers and non-smokers. *Drug Alcohol. Depend.* 135, 152–155. doi: 10.1016/j.drugalcdep.2013.11.014
- Jarmolowicz, D. P., Bickel, W. K., and Gatchalian, K. M. (2013). Alcohol-dependent individuals discount sex at higher rates than controls. *Drug Alcohol. Depend.* 131, 320–323. doi: 10.1016/j.drugalcdep.2012.12.014
- Jarmolowicz, D. P., Landes, R. D., Christensen, D. R., Jones, B. A., Jackson, L., Yi, R., et al. (2014). Discounting of money and sex: effects of commodity and temporal position in stimulant-dependent men and women. *Addict. Behav.* 39, 1652–1657. doi: 10.1016/j.addbeh.2014.04.026
- Johnson, A., van der Meer, M. A., and Redish, A. D. (2007). Integrating hippocampus and striatum in decision-making. *Curr. Opin. Neurobiol.* 17, 692–697. doi: 10.1016/j.conb.2008.01.003
- Jones, B. A., and Rachlin, H. (2009). Delay, probability, and social discounting in a public goods game. *J. Exp. Anal. Behav.* 91, 61–73. doi: 10.1901/jeab.2009.91-61
- Kable, J. W., and Glimcher, P. W. (2010). An “as soon as possible” effect in human intertemporal decision making: behavioral evidence and neural mechanisms. *J. Neurophysiol.* 103, 2513–2531. doi: 10.1152/jn.00177.2009
- Kacelnik, A. (2011). “The evolution of patience,” in *Time and Decision*, eds G. Loewenstein, D. Read, and R. F. Baumeister (New York, NY: Russell Sage), 115–137.
- Kahneman, D., and Tversky, A. (1979). Prospect theory: an analysis of decision under risk. *Econometrica* 47, 263–291. doi: 10.2307/1914185
- Kalenscher, T., and Pennartz, C. M. (2008). Is a bird in the hand worth two in the future? The neuroeconomics of intertemporal decision-making. *Prog. Neurobiol.* 84, 284–315. doi: 10.1016/j.pneurobio.2007.11.004
- Keramati, M., and Gutkin, B. S. (2011). “A reinforcement learning theory for homeostatic regulation,” in *Advances in Neural Information Processing Systems 24 (NIPS 2011)*, eds J. Shawe-Taylor, R. S. Zemel, P. L. Bartlett, F. Pereira, and K. Q. Weinberger. Available online at: <http://papers.nips.cc/paper/4437-a-reinforcement>
- Kimura, K., Izawa, S., Sugaya, N., Ogawa, N., Yamada, K. C., Shiotsuki, K., et al. (2013). The biological effects of acute psychosocial stress on delay discounting. *Psychoneuroendocrinology* 38, 2300–2308. doi: 10.1016/j.psyneuen.2013.04.019
- Kirby, K. N., and Herrnstein, R. J. (1995). Preference reversals due to myopic discounting of delayed reward. *Psychol. Sci.* 6, 83–89. doi: 10.1111/j.1467-9280.1995.tb00311.x
- Kirby, K. N., and Maraković, N. N. (1995). Modeling myopic decisions: evidence for hyperbolic delay-discounting within subjects and amounts. *Organ. Behav. Hum. Decis. Proc.* 64, 22–30. doi: 10.1006/obhd.1995.1086
- Kirby, K. N., and Petry, N. M. (2004). Heroin and cocaine abusers have higher discount rates for delayed rewards than alcoholics or non-drug-using controls. *Addiction* 99, 461–471. doi: 10.1111/j.1360-0443.2003.00669.x

- Kirby, K. N., Petry, N. M., and Bickel, W. K. (1999). Heroin addicts have higher discount rates for delayed rewards than non-drug-using controls. *J. Exp. Psychol. Gen.* 128, 78–87. doi: 10.1037/0096-3445.128.1.78
- Kirk, J., and Logue, A. (1997). Effects of deprivation level on humans' self-control for food reinforcers. *Appetite* 28, 215–226. doi: 10.1006/appe.1996.0071
- Koffarnus, M. N., Jarmolowicz, D. P., Mueller, E. T., and Bickel, W. K. (2013). Changing delay discounting in the light of the competing neurobehavioral decision systems theory: a review. *J. Exp. Anal. Behav.* 99, 32–57. doi: 10.1002/jeab.2
- Kollins, S. H. (2003). Delay discounting is associated with substance use in college students. *Addict. Behav.* 28, 1167–1173. doi: 10.1016/S0306-4603(02)00220-4
- Kuntsi, J., Oosterlaan, J., and Stevenson, J. (2001). Psychological mechanisms in hyperactivity: I response inhibition deficit, working memory impairment, delay aversion, or something else? *J. Child Psychol. Psychiatry* 42, 199–210. doi: 10.1111/1469-7610.00711
- Kurth-Nelson, Z., and Redish, A. D. (2009). Temporal-difference reinforcement learning with distributed representations. *PLoS ONE* 4:e7362. doi: 10.1371/journal.pone.0007362
- Lahav, E., Benzon, U., and Shavit, T. (2011). The effect of military service on soldiers' time preference-Evidence from Israel. *Judgm. Decis. Mak.* 6, 130–138.
- Laibson, D. (1997). Golden eggs and hyperbolic discounting. *Q. J. Econ.* 112, 443–478. doi: 10.1162/003355397555253
- Lawrence, K. A., Allen, J. S., and Chanen, A. M. (2010). Impulsivity in borderline personality disorder: reward-based decision-making and its relationship to emotional distress. *J. Pers. Disord.* 24, 785–799. doi: 10.1521/pedi.2010.24.6.785
- Lawyer, S. R., Williams, S. A., Prihodova, T., Rollins, J. D., and Lester, A. C. (2010). Probability and delay discounting of hypothetical sexual outcomes. *Behav. Proc.* 84, 687–692. doi: 10.1016/j.beproc.2010.04.002
- Leigh, J. P. (1986). Accounting for tastes: correlates of risk and time preferences. *J. Post Keynes. Econ.* 9, 17–31. doi: 10.1080/01603477.1986.11489597
- Leitão, M., Guedes, A., Yamamoto, M. E., and de Araújo Lopes, F. (2013). Do people adjust career choices according to socioeconomic conditions? An evolutionary analysis of future discounting. *Psychol. Neurosci.* 6, 383. doi: 10.3922/j.psns.2013.3.16
- Lempert, K. M., Porcelli, A. J., Delgado, M. R., and Tricomi, E. (2012). Individual differences in delay discounting under acute stress: the role of trait perceived stress. *Front. Psychol.* 3:251. doi: 10.3389/fpsyg.2012.00251
- Lewis, J., and Christopher, J. (1989). Childhood trauma in borderline personality disorder. *Am. J. Psychiatry* 1, 46.
- Loewenstein, G. (1996). Out of control: visceral influences on behavior. *Organ. Behav. Hum. Decis. Proc.* 65, 272–292. doi: 10.1006/obhd.1996.0028
- Loman, M. M. (2012). *Is deprivation-Related ADHD Different from ADHD Among Children Without Histories of Deprivation?* Doctoral Dissertation: University of Minnesota.
- Lorant, V., Croux, C., Weich, S., Deliege, D., Mackenbach, J., and Anseau, M. (2007). Depression and socio-economic risk factors: 7-year longitudinal population study. *Brit. J. Psychiatry* 190, 293–298. doi: 10.1192/bjp.bp.105.020040
- Lucantonio, F., Caprioli, D., and Schoenbaum, G. (2014). Transition from 'model-based' to 'model-free' behavioral control in addiction: involvement of the orbitofrontal cortex and dorsolateral striatum. *Neuropharmacology* 76, 407–415. doi: 10.1016/j.neuropharm.2013.05.033
- Luhmann, C. C. (2013). Discounting of delayed rewards is not hyperbolic. *J. Exp. Psychol. Learn.* 39, 1274. doi: 10.1037/a0031170
- Luman, M., Oosterlaan, J., and Sergeant, J. A. (2005). The impact of reinforcement contingencies on AD/HD: a review and theoretical appraisal. *Clin. Psychol. Rev.* 25, 183–213. doi: 10.1016/j.cpr.2004.11.001
- Lupien, S. J., Maheu, F., Tu, M., Fiocco, A., and Schramek, T. E. (2007). The effects of stress and stress hormones on human cognition: implications for the field of brain and cognition. *Brain Cogn.* 65, 209–237. doi: 10.1016/j.bandc.2007.02.007
- MacKillop, J., and Kahler, C. W. (2009). Delayed reward discounting predicts treatment response for heavy drinkers receiving smoking cessation treatment. *Drug Alcohol. Depend.* 104, 197–203. doi: 10.1016/j.drugalcdep.2009.04.020
- MacKillop, J., and Tidey, J. W. (2011). Cigarette demand and delayed reward discounting in nicotine-dependent individuals with schizophrenia and controls: an initial study. *Psychopharmacology* 216, 91–99. doi: 10.1007/s00213-011-2185-8
- Madden, G. J., Petry, N. M., Badger, G. J., and Bickel, W. K. (1997). Impulsive and self-control choices in opioid-dependent patients and non-drug-using control patients: drug and monetary rewards. *Exp. Clin. Psychopharmacol.* 5, 256.
- Madsen, A., Karle, A., Rubin, P., Cortsen, M., Andersen, H., and Hemmingsen, R. (1999). Progressive atrophy of the frontal lobes in first-episode schizophrenia: interaction with clinical course and neuroleptic treatment. *Acta Psychiatr. Scand.* 100, 367–374. doi: 10.1111/j.1600-0447.1999.tb10880.x
- Mani, A., Mullainathan, S., Shafir, E., and Zhao, J. (2013). Poverty impedes cognitive function. *Science* 341, 976–980. doi: 10.1126/science.1238041
- Mann, J. J., Huang, Y.-Y., Underwood, M. D., Kassir, S. A., Oppenheim, S., Kelly, T. M., et al. (2000). A serotonin transporter gene promoter polymorphism (5-HTTLPR) and prefrontal cortical binding in major depression and suicide. *Arch. Gen. Psychiatry* 57, 729–738. doi: 10.1001/archpsyc.57.8.729
- Mariano, T., Bannerman, D., McHugh, S., Preston, T., Rudebeck, P., Rudebeck, S., et al. (2009). Impulsive choice in hippocampal but not orbitofrontal cortex—lesioned rats on a nonspatial decision-making maze task. *Eur. J. Neurosci.* 30, 472–484. doi: 10.1111/j.1460-9568.2009.06837.x
- Marr, D. (1982). *Vision: A Computational Investigation into the Human Representation and Processing of Visual Information*. San Francisco, CA: W. H. Freeman.
- Mason, L., O'Sullivan, N., Blackburn, M., Bentall, R., and El-Derey, W. (2012). I want it now! Neural correlates of hypersensitivity to immediate reward in hypomania. *Biol. Psychiatry* 71, 530–537. doi: 10.1016/j.biopsych.2011.10.008
- Mathys, C., Daunizeau, J., Friston, K. J., and Stephan, K. E. (2011). A Bayesian foundation for individual learning under uncertainty. *Front. Hum. Neurosci.* 5:39. doi: 10.3389/fnhum.2011.00039
- Mazas, C. A., Finn, P. R., and Steinmetz, J. E. (2000). Decision-making biases, antisocial personality, and early-onset alcoholism. *Alcohol. Clin. Exp. Res.* 24, 1036–1040. doi: 10.1111/j.1530-0277.2000.tb04647.x
- Mazur, J. E. (1987). "An adjusting procedure for studying delayed reinforcement," in *Quantitative Analysis of Behavior*, eds M. Commons, J. A. Nevin, and H. Rachlin (Hillsdale, NJ: Erlbaum), 55–73.
- McClure, S. M., Ericson, K. M., Laibson, D. I., Loewenstein, G., and Cohen, J. D. (2007). Time discounting for primary rewards. *J. Neurosci.* 27, 5796–5804. doi: 10.1523/JNEUROSCI.4246-06.2007
- McClure, S. M., Laibson, D. I., Loewenstein, G., and Cohen, J. D. (2004). Separate neural systems value immediate and delayed monetary rewards. *Science* 306, 503–507. doi: 10.1126/science.1100907
- McDannald, M. A., Lucantonio, F., Burke, K. A., Niv, Y., and Schoenbaum, G. (2011). Ventral striatum and orbitofrontal cortex are both required for model-based, but not model-free, reinforcement learning. *J. Neurosci.* 31, 2700–2705. doi: 10.1523/JNEUROSCI.5499-10.2011
- Meier, S., and Sprenger, C. D. (2012). Time discounting predicts creditworthiness. *Psychol. Sci.* 23, 56–58. doi: 10.1177/0956797611425931
- Metcalfe, J., and Mischel, W. (1999). A hot/cool-system analysis of delay of gratification: dynamics of willpower. *Psychol. Rev.* 106:3. doi: 10.1037/0033-295X.106.1.3
- Michaelson, L., de la Vega, A., Chatham, C. H., and Munakata, Y. (2013). Delaying gratification depends on social trust. *Front. Psychol.* 4:355. doi: 10.3389/fpsyg.2013.00355
- Mischel, W., Shoda, Y., and Rodriguez, M. (1989). Delay of gratification in children. *Science* 244, 933–938. doi: 10.1126/science.2658056
- Moeller, F. G., Barratt, E. S., Dougherty, D. M., Schmitz, J. M., and Swann, A. C. (2001). Psychiatric aspects of impulsivity. *Am. J. Psychiatry* 158, 1783–1793. doi: 10.1176/appi.ajp.158.11.1783
- Moeller, F. G., Dougherty, D. M., Barratt, E. S., Oderinde, V., Mathias, C. W., Harper, R. A., et al. (2002). Increased impulsivity in cocaine dependent subjects independent of antisocial personality disorder and aggression. *Drug Alcohol. Depend.* 68, 105–111. doi: 10.1016/S0376-8716(02)00106-0
- Montague, P. R., Dolan, R. J., Friston, K. J., and Dayan, P. (2012). Computational psychiatry. *Trends Cogn. Sci.* 16, 72–80. doi: 10.1016/j.tics.2011.11.018
- Moore, S. C., and Cusens, B. (2010). Delay discounting predicts increase in blood alcohol level in social drinkers. *Psychiatry Res.* 179, 324–327. doi: 10.1016/j.psychres.2008.07.024
- Moutoussis, M., Bentall, R. P., Williams, J., and Dayan, P. (2008). A temporal difference account of avoidance learning. *Network* 19, 137–160. doi: 10.1080/09548980802192784

- Moutoussis, M., Story, G., and Dolan, R. (2015). The computational psychiatry of reward: broken brains or misguided minds? *Front. Psychol.* 6:1445. doi: 10.3389/fpsyg.2015.01445
- Murphy, F., Sahakian, B., Rubinsztein, J., Michael, A., Rogers, R., Robbins, T., et al. (1999). Emotional bias and inhibitory control processes in mania and depression. *Psychol. Med.* 29, 1307–1321. doi: 10.1017/S0033291799001233
- Myerson, J., and Green, L. (1995). Discounting of delayed rewards: models of individual choice. *J. Exp. Anal. Behav.* 64, 263–276. doi: 10.1901/jeab.1995.64-263
- Myerson, J., Green, L., and Warusawitharana, M. (2001). Area under the curve as a measure of discounting. *J. Exp. Anal. Behav.* 76, 235–243. doi: 10.1901/jeab.2001.76-235
- Nassar, M. R., Rumsey, K. M., Wilson, R. C., Parikh, K., Heasley, B., and Gold, J. I. (2012). Rational regulation of learning dynamics by pupil-linked arousal systems. *Nat. Neurosci.* 15, 1040–1046. doi: 10.1038/nn.3130
- Neal, D. T. (2006). Habits—a repeat performance. *Curr. Dir. Psychol. Sci.* 15, 198. doi: 10.1111/j.1467-8721.2006.00435.x
- Niv, Y., Daw, N., Joel, D., and Dayan, P. (2007). Tonic dopamine: opportunity costs and the control of response vigor. *Psychopharmacology* 191, 507–520. doi: 10.1007/s00213-006-0502-4
- Noble, K. G., McCandliss, B. D., and Farah, M. J. (2007). Socioeconomic gradients predict individual differences in neurocognitive abilities. *Dev. Sci.* 10, 464–480. doi: 10.1111/j.1467-7687.2007.00600.x
- O'Brien, C. P., Childress, A. R., Ehrman, R., and Robbins, S. J. (1998). Conditioning factors in drug abuse: can they explain compulsion? *J. Psychopharmacol.* 12, 15–22. doi: 10.1177/026988119801200103
- O'Doherty, J. P., Dayan, P., Friston, K., Critchley, H., and Dolan, R. J. (2003). Temporal difference models and reward-related learning in the human brain. *Neuron* 38, 329–337. doi: 10.1016/S0896-6273(03)00169-7
- Odum, A. L. (2011). Delay discounting: trait variable? *Behav. Processes* 87, 1–9. doi: 10.1016/j.beproc.2011.02.007
- Odum, A. L., Madden, G. J., Badger, G. J., and Bickel, W. K. (2000). Needle sharing in opioid-dependent outpatients: psychological processes underlying risk. *Drug Alcohol. Depend.* 60, 259–266. doi: 10.1016/S0376-8716(00)00111-3
- Odum, A. L., Madden, G. J., and Bickel, W. K. (2002). Discounting of delayed health gains and losses by current, never- and ex-smokers of cigarettes. *Nicotine Tob. Res.* 4, 295–303. doi: 10.1080/14622200210141257
- Ogata, S. N., Silk, K. R., Goodrich, S., Lohr, N. E., Westen, D., and Hill, E. M. (1990). Childhood sexual and physical abuse in adult patients with borderline personality disorder. *Am. J. Psychiatry* 147, 1008. doi: 10.1176/ajp.147.8.1008
- Orbell, S., and Verplanken, B. (2014). The strength of habit. *Health Psychol. Rev.* 9, 311–317. doi: 10.1080/17437199.2014.992031
- Ostaszewski, P., Green, L., and Myerson, J. (1998). Effects of inflation on the subjective value of delayed and probabilistic rewards. *Psychon. B Rev.* 5, 324–333. doi: 10.3758/BF03212959
- Ouellette, J. A. (1998). Habit and intention in everyday life: the multiple processes by which past behavior predicts future behavior. *Psychol. Bull.* 124:54. doi: 10.1037/0033-2909.124.1.54
- Paloyelis, Y., Asherson, P., and Kuntsi, J. (2009). Are ADHD symptoms associated with delay aversion or choice impulsivity? A general population study. *J. Am. Acad. Child Adolesc. Psychiatry* 48, 837–846. doi: 10.1097/CHI.0b013e3181ab8c97
- Paloyelis, Y., Asherson, P., Mehta, M. A., Faraone, S. V., and Kuntsi, J. (2010a). DAT1 and COMT effects on delay discounting and trait impulsivity in male adolescents with attention deficit/hyperactivity disorder and healthy controls. *Neuropsychopharmacology* 35, 2414–2426. doi: 10.1038/npp.2010.124
- Paloyelis, Y., Stahl, D. R., and Mehta, M. (2010b). Are steeper discounting rates in attention-deficit/hyperactivity disorder specifically associated with hyperactivity-impulsivity symptoms or is this a statistical artifact? *Biol. Psychiatry* 68, e15–e16. doi: 10.1016/j.biopsych.2010.02.025
- Pavlov, I. P. (2003). *Conditioned Reflexes*. Mineola, NY: Courier Dover Publications.
- Paykel, E. S. (1978). Contribution of life events to causation of psychiatric illness. *Psychol. Med.* 8, 245–253. doi: 10.1017/S003329170001429X
- Pepper, G. V., and Nettle, D. (2013). Death and the time of your life: experiences of close bereavement are associated with steeper financial future discounting and earlier reproduction. *Evol. Hum. Behav.* 34, 433–439. doi: 10.1016/j.evolhumbehav.2013.08.004
- Peters, J., and Büchel, C. (2010). Episodic future thinking reduces reward delay discounting through an enhancement of prefrontal-midtemporal interactions. *Neuron* 66, 138–148. doi: 10.1016/j.neuron.2010.03.026
- Petry, N. M. (2001). Delay discounting of money and alcohol in actively using alcoholics, currently abstinent alcoholics, and controls. *Psychopharmacology* 154, 243–250. doi: 10.1007/s002130000638
- Petry, N. M. (2002). Discounting of delayed rewards in substance abusers: relationship to antisocial personality disorder. *Psychopharmacology* 162, 425–432. doi: 10.1007/s00213-002-1115-1
- Petry, N. M. (2003). Discounting of money, health, and freedom in substance abusers and controls. *Drug Alcohol. Depend.* 71, 133–141. doi: 10.1016/S0376-8716(03)00090-5
- Petry, N. M., and Casarella, T. (1999). Excessive discounting of delayed rewards in substance abusers with gambling problems. *Drug Alcohol. Depend.* 56, 25–32. doi: 10.1016/S0376-8716(99)00010-1
- Pine, A., Seymour, B., Roiser, J. P., Bossaerts, P., Friston, K. J., Curran, H. V., et al. (2009). Encoding of marginal utility across time in the human brain. *J. Neurosci.* 29, 9575–9581. doi: 10.1523/JNEUROSCI.1126-09.2009
- Pine, A., Shiner, T., Seymour, B., and Dolan, R. J. (2010). Dopamine, time, and impulsivity in humans. *J. Neurosci.* 30, 8888–8896. doi: 10.1523/JNEUROSCI.6028-09.2010
- Platt, M. L., and Huettel, S. A. (2008). Risky business: the neuroeconomics of decision making under uncertainty. *Nat. Neurosci.* 11, 398–403. doi: 10.1038/nn2062
- Pulcu, E., Trotter, P., Thomas, E., McFarquhar, M., Juhasz, G., Sahakian, B., et al. (2014). Temporal discounting in major depressive disorder. *Psychol. Med.* 44, 1825–1834. doi: 10.1017/S0033291713002584
- Rachlin, H. (2006). Notes on discounting. *J. Exp. Anal. Behav.* 85, 425–435. doi: 10.1901/jeab.2006.85-05
- Ramos, D., Victor, T., Seidl-de-Moura, M. L., and Daly, M. (2013). Future discounting by slum-dwelling youth versus university students in Rio de Janeiro. *J. Res. Adolesc.* 23, 95–102. doi: 10.1111/j.1532-7795.2012.00796.x
- Rapport, M. D., Tucker, S. B., DuPaul, G. J., Merlo, M., and Stoner, G. (1986). Hyperactivity and frustration: the influence of control over and size of rewards in delaying gratification. *J. Abnorm. Child Psychol.* 14, 191–204. doi: 10.1007/BF00915440
- Read, D. (2001). Is time-discounting hyperbolic or subadditive? *J. Risk Uncert.* 23, 5–32. doi: 10.1023/A:1011198414683
- Read, D. (2004). "Intertemporal choice," in *Blackwell Handbook of Judgment and Decision Making*, eds D. J. Koehler and N. Harvey (Oxford: Blackwell), 424–443. doi: 10.1002/9780470752937.ch21
- Read, D., Frederick, S., and Airoldi, M. (2012). Four days later in Cincinnati: longitudinal tests of hyperbolic discounting. *Acta Psychol.* 140, 177–185. doi: 10.1016/j.actpsy.2012.02.010
- Read, D., and Van Leeuwen, B. (1998). Predicting hunger: the effects of appetite and delay on choice. *Organ. Behav. Hum. Decis. Process.* 76, 189–205. doi: 10.1006/obhd.1998.2803
- Reimers, S., Maylor, E. A., Stewart, N., and Chater, N. (2009). Associations between a one-shot delay discounting measure and age, income, education and real-world impulsive behavior. *Pers. Individ. Differ.* 47, 973–978. doi: 10.1016/j.paid.2009.07.026
- Rescorla, R. A., and Solomon, R. L. (1967). Two-process learning theory: relationships between Pavlovian conditioning and instrumental learning. *Psychol. Rev.* 74, 151. doi: 10.1037/h0024475
- Reynolds, B., and Fields, S. (2012). Delay discounting by adolescents experimenting with cigarette smoking. *Addiction* 107, 417–424. doi: 10.1111/j.1360-0443.2011.03644.x
- Reynolds, B., Patak, M., and Shroff, P. (2007). Adolescent smokers rate delayed rewards as less certain than adolescent nonsmokers. *Drug Alcohol Depend.* 90, 301–303. doi: 10.1016/j.drugalcdep.2007.04.008
- Reynolds, B., Richards, J. B., Horn, K., and Karraker, K. (2004). Delay discounting and probability discounting as related to cigarette smoking status in adults. *Behav. Processes* 65, 35–42. doi: 10.1016/S0376-6357(03)00109-8
- Robson, D., and Gray, R. (2007). Serious mental illness and physical health problems: a discussion paper. *Int. J. Nurs. Stud.* 44, 457–466. doi: 10.1016/j.ijnurstu.2006.07.013

- Rossow, I. (2008). Alcohol consumption and discounting. *Addict. Res. Theory* 16, 572–584. doi: 10.1080/16066350801896248
- Rounds, J. S., Beck, J. G., and Grant, D. M. (2007). Is the delay discounting paradigm useful in understanding social anxiety? *Behav. Res. Ther.* 45, 729–735. doi: 10.1016/j.brat.2006.06.007
- Rubinstein, A. (2003). Economics and psychology: the case of hyperbolic discounting. *Econ. Rev.* 44, 1207–1216. doi: 10.1111/1468-2354.t01-1-00106
- Russell, A. E., Ford, T., Williams, R., and Russell, G. (2015). The association between socioeconomic disadvantage and attention deficit/hyperactivity disorder (ADHD): a systematic review. *Child Psychiatry Hum. Dev.* doi: 10.1007/s10578-015-0578-3. [Epub ahead of print].
- Samuelson, P. (1937). A note on the measurement of utility. *Rev. Econ. Stud.* 4, 155–161. doi: 10.2307/2967612
- Schacter, D. L., and Schacter. (2008). Episodic simulation of future events. *Ann. N.Y. Acad. Sci.* 1124, 39. doi: 10.1196/annals.1440.001
- Schacter, D. L., Addis, D. R., and Buckner, R. L. (2008). Episodic simulation of future events: concepts, data, and applications. *Ann. N.Y. Acad. Sci.* 1124, 39–60. doi: 10.1196/annals.1440.001
- Scheres, A., and Hamaker, E. L. (2010). What we can and cannot conclude about the relationship between steep temporal reward discounting and hyperactivity-impulsivity symptoms in attention-deficit/hyperactivity disorder. *Biol. Psychiatry* 68, e17–e18. doi: 10.1016/j.biopsych.2010.05.021
- Scheres, A., Tontsch, C., Thoeny, A. L., and Kaczurkin, A. (2010). Temporal reward discounting in attention-deficit/hyperactivity disorder: the contribution of symptom domains, reward magnitude, and session length. *Biol. Psychiatry* 67, 641–648. doi: 10.1016/j.biopsych.2009.10.033
- Schultz, W. (2006). Behavioral theories and the neurophysiology of reward. *Annu. Rev. Psychol.* 57, 87–115. doi: 10.1146/annurev.psych.56.091103.070229
- Schweighofer, N., Bertin, M., Shishida, K., Okamoto, Y., Tanaka, S. C., Yamawaki, S., et al. (2008). Low-serotonin levels increase delayed reward discounting in humans. *J. Neurosci.* 28, 4528–4532. doi: 10.1523/JNEUROSCI.4982-07.2008
- Schweitzer, J. B., and Sulzer–Azaroff, B. (1995). Self-control in boys with attention deficit hyperactivity disorder: effects of added stimulation and time. *J. Child Psychol. Psychiatry* 36, 671–686. doi: 10.1111/j.1469-7610.1995.tb02321.x
- Seymour, B., and Dolan, R. (2008). Emotion, decision making, and the amygdala. *Neuron* 58, 662–671. doi: 10.1016/j.neuron.2008.05.020
- Seymour, B., O'Doherty, J. P., Koltzenburg, M., Wiech, K., Frackowiak, R., Friston, K., et al. (2005). Opponent appetitive-aversive neural processes underlie predictive learning of pain relief. *Nat. Neurosci.* 8, 1234–1240. doi: 10.1038/nn1527
- Shamosh, N. A., DeYoung, C. G., Green, A. E., Reis, D. L., Johnson, M. R., Conway, A. R., et al. (2008). Individual differences in delay discounting relation to intelligence, working memory, and anterior prefrontal cortex. *Psychol. Sci.* 19, 904–911. doi: 10.1111/j.1467-9280.2008.02175.x
- Smittenaar, P., FitzGerald, T. H., Romei, V., Wright, N. D., and Dolan, R. J. (2013). Disruption of dorsolateral prefrontal cortex decreases model-based in favor of model-free control in humans. *Neuron* 80, 914–919. doi: 10.1016/j.neuron.2013.08.009
- Sonuga-Barke, E., Taylor, E., Sembi, S., and Smith, J. (1992). Hyperactivity and delay aversion—I. The effect of delay on choice. *J. Child Psychol. Psychiatry* 33, 387–398. doi: 10.1111/j.1469-7610.1992.tb00874.x
- Sozou, P. D. (1998). On hyperbolic discounting and uncertain hazard rates. *Proc. Biol. Sci.* 265, 2015–2020. doi: 10.1098/rspb.1998.0534
- Stanger, C., Ryan, S. R., Fu, H., Landes, R. D., Jones, B. A., Bickel, W. K., et al. (2012). Delay discounting predicts adolescent substance abuse treatment outcome. *Exp. Clin. Psychopharmacol.* 20, 205. doi: 10.1037/a0026543
- Steinberg, L., Graham, S., O'Brien, L., Woolard, J., Cauffman, E., and Banich, M. (2009). Age differences in future orientation and delay discounting. *Child Dev.* 80, 28–44. doi: 10.1111/j.1467-8624.2008.01244.x
- Stephan, K. E., and Mathys, C. (2014). Computational approaches to psychiatry. *Curr. Opin. Neurobiol.* 25, 85–92. doi: 10.1016/j.conb.2013.12.007
- Stevens, D. W., and Krebs, J. R. (1986). *Foraging Theory*. New Jersey, NJ: Princeton University Press.
- Story, G., Vlaev, I., Seymour, B., Darzi, A., and Dolan, R. (2014). Does temporal discounting explain unhealthy behavior? A systematic review and reinforcement learning perspective. *Front. Behav. Neurosci.* 8:76. doi: 10.3389/fnbeh.2014.00076
- Story, G. W., Vlaev, I., Dayan, P., Seymour, B., Darzi, A., and Dolan, R. J. (2015). Anticipation and choice heuristics in the dynamic consumption of pain relief. *PLoS Comput. Biol.* 11:e1004030. doi: 10.1371/journal.pcbi.1004030
- Story, G. W., Vlaev, I., Seymour, B., Winston, J. S., Darzi, A., and Dolan, R. J. (2013). Dread and the disvalue of future pain. *PLoS Comput. Biol.* 9:e1003335. doi: 10.1371/journal.pcbi.1003335
- Strotz, R. H. (1957). Myopia and inconsistency in dynamic utility maximisation. *Rev. Econ. Stud.* 23, 165–180. doi: 10.2307/2295722
- Sutton, R. S., and Barto, A. G. (1998). *Introduction to Reinforcement Learning*. Boston, MA: MIT Press.
- Swann, A. C. (2009). Impulsivity in mania. *Curr. Psychiatry Rep.* 11, 481–487. doi: 10.1007/s11920-009-0073-2
- Swirsky-Sacchetti, T., Gorton, G., Samuel, S., Sobel, R., Genetta-Wadley, A., and Burleigh, B. (1993). Neuropsychological function in borderline personality disorder. *J. Clin. Psychol.* 49, 385–396.
- Takahashi, T. (2004). Cortisol levels and time-discounting of monetary gain in humans. *Neuroreport* 15, 2145–2147. doi: 10.1097/00001756-200409150-00029
- Takahashi, T., Ikeda, K., and Hasegawa, T. (2007). A hyperbolic decay of subjective probability of obtaining delayed rewards. *Behav. Brain Funct.* 3:52. doi: 10.1186/1744-9081-3-52
- Takahashi, T., Oono, H., Inoue, T., Boku, S., Kako, Y., Kitaichi, Y., et al. (2008). Depressive patients are more impulsive and inconsistent in intertemporal choice behavior for monetary gain and loss than healthy subjects—an analysis based on Tsallis' statistics. *Neuro Endocrinol. Lett.* 29, 351.
- Takahashi, T., Shinada, M., Inukai, K., Tanida, S., Takahashi, C., Mifune, N., et al. (2009). Stress hormones predict hyperbolic time-discount rates six months later in adults. *Neuro Endocrinol. Lett.* 31, 616–621.
- Tanaka, S. C., Schweighofer, N., Asahi, S., Shishida, K., Okamoto, Y., Yamawaki, S., et al. (2007). Serotonin differentially regulates short-and long-term prediction of rewards in the ventral and dorsal striatum. *PLoS ONE* 2:e1333. doi: 10.1371/journal.pone.0001333
- Thaler, R. H. (1981). Some empirical evidence on dynamic inconsistency. *Econ. Lett.* 8, 201–207. doi: 10.1016/0165-1765(81)90067-7
- Thaler, R. H., and Shefrin, H. M. (1981). An economic theory of self-control. *J. Polit. Econ.* 89, 392–406. doi: 10.1086/260971
- Thorndike, E. L. (1927). The law of effect. *Am. J. Psychol.* 39, 212–222. doi: 10.2307/1415413
- Townsend, J., Bookheimer, S. Y., Folland-Ross, L. C., Sugar, C. A., and Altschuler, L. L. (2010). fMRI abnormalities in dorsolateral prefrontal cortex during a working memory task in manic, euthymic and depressed bipolar subjects. *Psychiatry Res.* 182, 22–29. doi: 10.1016/j.psychres.2009.11.010
- Trepel, C., Fox, C. R., and Poldrack, R. A. (2005). Prospect theory on the brain? Toward a cognitive neuroscience of decision under risk. *Brain Res. Cogn. Brain Res.* 23, 34–50. doi: 10.1016/j.cogbrainres.2005.01.016
- Tricomi, E., Balleine, B. W., and O'Doherty, J. P. (2009). A specific role for posterior dorsolateral striatum in human habit learning. *Eur. J. Neurosci.* 29, 2225–2232. doi: 10.1111/j.1460-9568.2009.06796.x
- Tripp, G., and Alsop, B. (2001). Sensitivity to reward delay in children with attention deficit hyperactivity disorder (ADHD). *J. Child Psychol. Psychiatry* 42, 691–698. doi: 10.1111/1469-7610.00764
- Van den Bergh, B., Dewitte, S., and Warlop, L. (2008). Binkies instigate generalized impatience in intertemporal choice. *J. Consum. Res.* 35, 85–97. doi: 10.1086/525505
- van der Pol, M., and Cairns, J. (2002). A comparison of the discounted utility model and hyperbolic discounting models in the case of social and private intertemporal preferences for health. *J. Econ. Behav. Organ.* 49, 79–96. doi: 10.1016/S0167-2681(02)00059-8
- van Haren, N. E., Pol, H. E. H., Schnack, H. G., Cahn, W., Brans, R., Carati, I., et al. (2008). Progressive brain volume loss in schizophrenia over the course of the illness: evidence of maturational abnormalities in early adulthood. *Biol. Psychiatry* 63, 106–113. doi: 10.1016/j.biopsych.2007.01.004
- Van Oers, J., Bongers, I., Van de Goor, L., and Garretsen, H. (1999). Alcohol consumption, alcohol-related problems, problem drinking, and socioeconomic status. *Alcohol. Alcohol.* 34, 78–88. doi: 10.1093/alcac/34.1.78
- Velakoulis, D., Stuart, G. W., Wood, S. J., Smith, D. J., Brewer, W. J., Desmond, P., et al. (2001). Selective bilateral hippocampal volume loss

- in chronic schizophrenia. *Biol. Psychiatry* 50, 531–539. doi: 10.1016/S0006-3223(01)01121-0
- Wang, M., Rieger, M. O., and Hens, T. (2010). How time preferences differ: evidence from 45 countries. *Swiss Finance Institute Research Paper*, 09–47. Available online at: http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1481443 doi: 10.2139/ssrn.1481443
- Wang, X.-J., and Krystal, J. H. (2014). Computational psychiatry. *Neuron* 84, 638–654. doi: 10.1016/j.neuron.2014.10.018
- Wang, X. T., and Dvorak, R. D. (2010). Sweet future fluctuating blood glucose levels affect future discounting. *Psychol. Sci.* 21, 183–188. doi: 10.1177/0956797609358096
- Washio, Y., Higgins, S. T., Heil, S. H., McKerchar, T. L., Badger, G. J., Skelly, J. M., et al. (2011). Delay discounting is associated with treatment response among cocaine-dependent outpatients. *Exp. Clin. Psychopharmacol.* 19, 243–248. doi: 10.1037/a0023617
- Watkins, C. J., and Dayan, P. (1992). Q-learning. *Mach. Learn.* 8, 279–292. doi: 10.1007/BF00992698
- Weich, S., and Lewis, G. (1998). Poverty, unemployment, and common mental disorders: population based cohort study. *Brit. Med. J.* 317, 115–119. doi: 10.1136/bmj.317.7151.115
- Weller, R. E., Avsar, K. B., Cox, J. E., Reid, M. A., White, D. M., and Lahti, A. C. (2014). Delay discounting and task performance consistency in patients with schizophrenia. *Psychiatry Res.* 215, 286–293. doi: 10.1016/j.psychres.2013.11.013
- Wesley, M. J., Lohrenz, T., Koffarnus, M. N., McClure, S. M., De La Garza, R., Salas, R., et al. (2014). Choosing money over drugs: the neural underpinnings of difficult choice in chronic cocaine users. *J. Addict.* 2014:189853. doi: 10.1155/2014/189853
- Williams, D. R., and Williams, H. (1969). Auto-maintenance in the pigeon: sustained pecking despite contingent non-reinforcement. *J. Exp. Anal. Behav.* 12, 511–520. doi: 10.1901/jeab.1969.12-511
- Williams, G. C. (1957). Pleiotropy, natural selection, and the evolution of senescence. *Evolution* 11, 398–411. doi: 10.2307/2406060
- Wilson, R. C., Nassar, M. R., and Gold, J. I. (2010). Bayesian online learning of the hazard rate in change-point problems. *Neural Comput.* 22, 2452–2476. doi: 10.1162/NECO_a_00007
- Wilson, V. B., Mitchell, S. H., Musser, E. D., Schmitt, C. F., and Nigg, J. T. (2011). Delay discounting of reward in ADHD: application in young children. *J. Child Psychol. Psychiatry* 52, 256–264. doi: 10.1111/j.1469-7610.2010.02347.x
- Wing, V. C., Moss, T. G., Rabin, R. A., and George, T. P. (2012). Effects of cigarette smoking status on delay discounting in schizophrenia and healthy controls. *Addict. Behav.* 37, 67–72. doi: 10.1016/j.addbeh.2011.08.012
- Winstanley, C. A., Theobald, D. E., Dalley, J. W., and Robbins, T. W. (2005). Interactions between serotonin and dopamine in the control of impulsive choice in rats: therapeutic implications for impulse control disorders. *Neuropsychopharmacology* 30, 669–682. doi: 10.1038/sj.npp.1300610
- Wunderlich, K., Smittenaar, P., and Dolan, R. J. (2012). Dopamine enhances model-based over model-free choice behavior. *Neuron* 75, 418–424. doi: 10.1016/j.neuron.2012.03.042
- Yechiam, E., Busemeyer, J. R., Stout, J. C., and Bechara, A. (2005). Using cognitive models to map relations between neuropsychological disorders and human decision-making deficits. *Psychol. Sci.* 16, 973–978. doi: 10.1111/j.1467-9280.2005.01646.x
- Yi, R., Gatchalian, K. M., and Bickel, W. K. (2006). Discounting of past outcomes. *Exp. Clin. Psychopharmacol.* 14:311. doi: 10.1037/1064-1297.14.3.311
- Yu, A. J., and Dayan, P. (2003). “Expected and unexpected uncertainty: ACh and NE in the neocortex,” in *Advances in Neural Information Processing Systems 15* (Cambridge, MA).
- Yu, A., and Dayan, P. (2005). Uncertainty, neuromodulation, and attention. *Neuron* 46, 681–692. doi: 10.1016/j.neuron.2005.04.026
- Zanarini, M. C., Williams, A. A., Lewis, R. E., and Reich, R. B. (1997). Reported pathological childhood experiences associated with the development of borderline personality disorder. *Am. J. Psychiatry* 154:1101. doi: 10.1176/ajp.154.8.1101

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2016 Story, Moutoussis and Dolan. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

