OPEN ACCESS

King's College London

King's Research Portal

DOI:
10.1007/s11097-017-9508-0

Document Version
Peer reviewed version

Link to publication record in King's Research Portal

Citation for published version (APA):
Fridland, E., & Wiers, C. (2017). Addiction and Embodiment. *Phenomenology & the Cognitive Sciences*. Advance online publication. https://doi.org/10.1007/s11097-017-9508-0

Citing this paper
Please note that where the full-text provided on King's Research Portal is the Author Accepted Manuscript or Post-Print version this may differ from the final Published version. If citing, it is advised that you check and use the publisher's definitive version for pagination, volume/issue, and date of publication details. And where the final published version is provided on the Research Portal, if citing you are again advised to check the publisher's website for any subsequent corrections.

General rights
Copyright and moral rights for the publications made accessible in the Research Portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognize and abide by the legal requirements associated with these rights.

•Users may download and print one copy of any publication from the Research Portal for the purpose of private study or research.
•You may not further distribute the material or use it for any profit-making activity or commercial gain
•You may freely distribute the URL identifying the publication in the Research Portal

Take down policy
If you believe that this document breaches copyright please contact librarypure@kcl.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.

Download date: 13. Sep. 2024

Addiction and Embodiment

Word count: 11,301

Ellen Fridland, PhD [1]
Corinde E. Wiers, PhD [2]

[1] Department of Philosophy, King's College London, United Kingdom
[2] Berlin School of Mind and Brain, Humboldt-Universität zu Berlin, Germany

Corresponding author:
Ellen Fridland
King's College London
Department of Philosophy
Room 507, Philosophy Building
Strand
London
WC2R 2LS

ellenfridland@gmail.com
ellen.fridland@kcl.ac.uk

Addiction and Embodiment

Abstract

Recent experiments have shown that when drug-dependent individuals are confronted with drug-related cues, they exhibit an automatically activated tendency to approach these cues (i.e., drug approach bias). The strength of the drug approach bias has been associated with clinically relevant measures, such as increased drug craving and relapse, as well as activations in brain reward areas. Moreover, retraining the approach bias by means of Cognitive Bias Modification has been demonstrated to decrease relapse rates in alcohol-dependent patients and to reduce alcohol cue-evoked limbic activity. Here, we review empirical and theoretical literature on the drug approach bias and explore two distinct models of how the drug approach bias may be embodied. First, we consider the "biological meaning" hypothesis, which grounds the automatic approach bias in the natural meaning of the body. Second, we consider the "sensorimotor hypothesis," which appeals to the specific sensorimotor loops involved in the instantiation of addictive behaviors as the basis of the automatic approach bias. In order to differentiate between the adequacies of these competing explanations, we present specific, predictions that each model should make.

Keywords addiction, approach bias, embodiment, implicit cognition

Contents

1. Introduction

Drug addiction is characterized by a core paradox: the continuation of drug consumption despite the awareness of negative consequences. Though initial drug taking may be a fully conscious choice, recent studies have shown that automatic processes play a large role in the pathology of addiction and the high risk of relapse. That is, there is evidence that drug cues capture automatic attention (e.g., Field, Mogg, Mann, Bennett, & Bradley, 2013), evoke activation in the midbrain dopamine reward system (see Heinz, Beck, Grusser, Grace, & Wrase, 2009 for a review) and engender automatic approach responses in drug-dependent individuals (e.g., C. E. Wiers et al., 2014), which may take place largely outside of conscious awareness. These automatically activated processes have been associated with increased drug consumption, despite the fact that individuals may express an explicit wish to quit.

The focus of the current paper is the drug approach bias: the automatic tendency to approach faster rather than avoid drug cues in drug-dependent individuals. It remains unclear what the underlying mechanisms of the approach bias are. For example, the bias has been investigated in light of cognitive theories, such as Pavlovian conditioning (i.e., classical conditioning of the drug with drug cues) and habit-formations (i.e., pairing of drug stimulus and the response) or dual process models which hypothesize an imbalance between overactive impulsive processes and less active control processes in drug consumption, resulting in imbalanced approach/avoid responses to drug cues (e.g., Watson, de Wit, Hommel, & Wiers, 2012; R. W. Wiers, Gladwin, Hofmann, Salemink, & Ridderinkhof, 2013).

In contrast to purely cognitive explanations which rely on rational or associative mental representations to account for cognition, embodiment theorists highlight the possibility of situated bodily and sensorimotor processes structuring and possibly even constituting cognitive states (Paulus & Stewart, 2014). The role of the body in explaining cognitive judgment and behvior has been investigated in a varity of experimental paradigms. For instance, the classic exmperiment of Strack, Martin, and Stepper (1988) showed that participants rate cartoons as funnier when a pen is held in between their teeth. Holding a pen in one's teeth results in the activation of the same muscles that are used when smiling. In contrast, when a pen is held using the lips, which uses the same muscles as when frowning, particants rate cartoons as less funny. Also, arm movements have been shown to be related to evaluative outcomes of tasks: pushing objects has been associated with the avoidance of undesired objects whereas pulling has been associated with the approach of something positive (Chen & Bargh, 1999; Markman & Brendl, 2005; Phaf, Mohr, Rotteveel, & Wicherts, 2014).

As yet, it remains unexplored whether the drug approach bias reflects an embodied reaction towards drug cues. It is this possibility that we would like to explore in this paper. To this end, we first review empirical evidence for the drug approach bias and its relation to clinical measures, such as craving and relapse. In addition, we provide an overview of studies that research the effect of retraining the approach bias by means of Cognitive Bias Modification (CBM), which have recently shown to be clinically effective in alcohol dependent patients (Eberl et al., 2013; R. W. Wiers, Eberl, Rinck, Becker, & Lindenmeyer, 2011). After a short paragraph on cognitive models that have tried to explain the drug approach bias, in the second half of the paper we explore two models of how the drug approach bias may be embodied. We shall not present global arguments defending the virtues of understanding automatic approach tendencies in drug-dependent individuals as embodied rather than in cognitive or associative terms. Though we do not deny that a firm argument in favor of embodied approaches would be useful and relevant to our discussion, we take it for granted that there are important considerations within the embodied cognition paradigm that are worthy of attention in the absence of a lengthy, antecedent defense of the virtues of the paradigm as a whole.

In general, this paper is an exercise in empirically informed philosophy of mind. Our main goal is to clarify the conceptual connections between various theoretical paradigms and their empirical support, and also to present predictions that are theoretically consistent with the distinct options presented. We do not aim to take a position on which of the options is to be favored.

2. Empirical evidence for automatic approach tendencies in addiction

Over the last years computerized "implicit" or automatic tasks have been developed to assess behavioral reactions of humans that lie largely outside of conscious awareness. For example, studies have assessed reactions of avoidance to pictures of spiders (Rinck & Becker, 2007), motivational reactions towards pictures of food (Loeber et al., 2011) and drugs (Barkby, Dickson, Roper, & Field, 2012; Ernst et al., 2014), and even implicit attitudes towards race (Stanley, Sokol-Hessner, Banaji, & Phelps, 2011). The tasks that measure these reactions are considered implicit or automatic if the instructions of the task are indirect (i.e., subjects are unaware of what the task measures) or if the outcome measures are reaction times that are fast, goal-independent and not controllable (De Houwer, 2006). In drug addiction, in particular, it has been shown that drug-related cues evoke attention in dependent individuals (i.e., attentional bias), such as drug-related words in a drug Stroop task (Cox, Fadardi, & Pothos, 2006) or pictorial cues in the Visual Probe task (e.g., Field et al., 2013). Moreover, there is cumulative evidence that drug-dependent individuals approach rather than avoid pictorial drug cues in comparison to non-addicted control groups. This is called the drug approach bias.

The approach bias can be measured on several tasks: the Stimulus-Response Compatibility (SRC) task and the joystick-based Approach Avoidance Task (AAT). In the SRC task, subjects move a manikin that is depicted on a computer screen towards or away from cues (drug-related or neutral) using button presses. Drug-dependent individuals have been sown to move the manikin faster towards drug-related cues than neutral cues (Blumstein & Schardt, 2009; Bradley, Field, Healy, & Mogg, 2008; Mogg, Bradley, Field, & De Houwer, 2003; Mogg, Field, & Bradley, 2005; Solomon et al.). In the AAT, subjects push and pull pictures away and towards one self with a joystick. The approach bias is calculated by the difference score of push minus pull trials per cue type: the larger this number, the stronger the tendency to approach rather than avoid a cue type. Though the outcome measures of both the SRC and AAT are labeled approach bias, there is no evidence for a correlation between the two measures (Krieglmeyer & Deutsch, 2010). Since we are interested in automatic action tendencies towards drugs, rather than in 3rd person perspectives on approach/avoidance tendencies, in this article we limit ourselves to the drug approach bias measured with the AAT.

2.1 Variations of the Approach Avoidance Task

There are implicit and explicit versions of the AAT, using indirect and direct instructions respectively. The most frequently used indirect instruction is when participants are asked to push and pull cues according to the format of the cue, rather than the cue itself. For example, often the landscape or portrait format of a picture is the feature to which participants are asked to respond (e.g., R. W. Wiers et al., 2011) but cues tilted slightly left or right have also been used (e.g., Cousijn, Goudriaan, & Wiers, 2011). On the indirect task it has been shown that heavy drinkers (R. W. Wiers, Rinck, Dictus, & van den Wildenberg, 2009), alcohol-dependent patients (C. E. Wiers et al., 2014; R. W. Wiers et al., 2011), heroin abusers (Zhou et al., 2012), heavy tobacco smokers (C. E. Wiers et al., 2013) and cannabis users (Cousijn et al., 2011) approach drug cues faster compared to non-addicted control groups. In the task with explicit instructions, the AAT consists of blocks where participants are asked to push away drug cues and pull neutral cues, and vice versa. For example, Ernst et al. (2014) found that alcohol-dependent inpatients are faster in approaching than avoiding alcohol cues on the direct AAT, an effect with a comparable size as found on implicit tasks (e.g., C. E. Wiers et al., 2014). So even though the indirect instruction was initially thought to be advantageous in reducing the controllability of the outcome measure, it does not seem to be a necessary feature for measuring the drug approach bias.

A second important feature of the AAT is a zooming effect, introduced by Rinck and Becker (2007). The zooming effect involves cues zooming in on the screen when participants pull the joystick, whereas cues zoom out while pushing the joystick. This feature has shown to be of importance, since joystick movements alone are ambiguous: for example reaching out one's arm may represent avoidance in some situations (i.e., pushing something away from the body), but approach in others (i.e., grasping to reach out for a drink). It has been shown in various studies that the interpretations of the movements depend on the outcome of the action (Krieglmeyer, Deutsch, De Houwer, & De Raedt, 2010; Lavender & Hommel, 2007). While the approach bias was initially thought solely to be a motor reaction to drug cues, these movements have been shown to be meaningful and subject to interpretation. However, a study of adolescent drinkers showed an approach bias for alcohol cues using button presses rather than a joystick (Peeters et al., 2012). Although a direct comparison of joysticks versus button presses is lacking, the button press result demonstrates that the role of arm movements in the approach bias is far from clear.

2.2 Associations of the approach bias with craving and drug use

In addition to the increasing evidence that drug users' tendencies of approaching cues are stronger compared to control groups, it has been explored whether the strength of these measures are associated with drug craving, drug use, and other clinical measures. In smokers, for example, the strength of drug approach tendencies correlated with drug craving in various studies (Mogg et al., 2005; Watson, de Wit, Cousijn, Hommel, & Wiers, 2013; C. E. Wiers et al., 2013). Moreover, the alcohol approach bias (measured on the SRC) was correlated with drinking consumption and self-reported alcohol approach preferences (Barkby et al., 2012). Though this finding has not been reported in AAT studies, C. E. Wiers et al (2014) found that alcohol craving was correlated with amygdala activations while approaching versus avoiding alcohol cues in alcohol-dependent patients.

First insight into neural mechanisms involved in the alcohol approach bias, shows that reward-related brain areas (i.e., the nucleus accumbens, the medial prefrontal cortex and amygdala) were associated with approaching versus avoiding alcohol cues in alcohol-dependence (Ernst et al., 2014; C. E. Wiers et al., 2014). In contrast, the dorsolateral prefrontal cortex, an area usually involved in cognitive control, was more active for avoiding alcohol cues, measured on a direct task only (Ernst et al., 2014). This finding was in line with approach/avoidance studies on emotional processing that found that the dlPFC was active when stimulus and response are incongruent (approach sad faces) than during congruent (approach happy faces) trials (Roelofs, Minelli, Mars, van Peer, & Toni, 2009; Volman, Toni, Verhagen, & Roelofs, 2011).

Currently, no AAT study has reported a direct relation with relapse in clinical populations. Nevertheless, the strength of a cannabis approach bias has been shown to predict changes in cannabis consumption in heavy cannabis users (Cousijn et al., 2011) and it has been demonstrated that former heavy smokers, who had not smoked for a long time, did not have a bias on the AAT (C. E. Wiers et al., 2013). In sum, these findings indicate that the alcohol approach bias is of clinical importance in drug addiction.

2.3 Approach/avoidance tendencies beyond drug addiction

Approach/avoidance reactions are not limited to drug cues in drug-consuming populations. Some authors suggest that approach/avoidance tendencies may be general bodily reactions to positive and negatively stimuli, respectively (Phaf et al., 2014). For example, Chen and Bargh (1999) demonstrated that participants pull positive words faster than negative words on a lever, whereas negative words are pushed faster than positive words. In line with this, people who feared spiders had stronger avoidance tendencies for spiders than for neutral pictures (Rinck & Becker, 2007) and socially anxious people avoided smiling and angry faces faster compared to non-anxious controls (Heuer, Rinck, & Becker, 2007). Depressed patients were recently shown to not show any approach/avoidance reactions towards emotional faces on the AAT, in comparison to a healthy control group (Radke, Guths, Andre, Muller, & de Bruijn, 2014). In addition, schizophrenic patients with higher levels of oxytocin had quicker avoidance tendencies for angry faces, suggesting that the effects of oxytocin may influence the avoidance of negative or threatening emotions. In this study, stronger avoidance of angry faces correlated with more severe psychotic symptoms and paranoia (Brown et al., 2014). Using food cues, anorexia-nervosa patients showed a decreased approach bias for food cues compared to controls, suggesting decreased motivational saliency for food or a deliberate avoidance of food cues (Veenstra & de Jong, 2011).

But is there evidence that approach reactions are motor responses essentially related to the body? In a clever experiment, Markman and Brendl (2005) showed that responses to approach positive words and avoid negative words were faster for participants' names on a computer screen than to participants' bodies. In fact, approach and avoidance in relation to the body was reversed: that is, participants were faster in pulling negative words towards their bodies than pushing negative words away from their bodies, and vice versa for positive words. Markman and Brendl (2005) conclude that approach/avoidance movements depend on people's representation of their selves in space rather than on their physical location. This result poses a challenge to embodied approaches that attempt to ground natural meaning in the actual, physical body. However, it has yet to be determined if the reversed approach/avoidance response is the result of cognitive inhibition mechanisms that override the natural reaction to pull towards ones' body for approach and push away for avoidance or whether approach and avoidance tendencies are essentially unrelated to the physical location of the body. In order to determine this, one would have to compare the speed of pull/push for one's represented or projected location in space with the speed of pull/push towards one's actual, physical body in the absence of a concurrent projected representation of self.

2.4. Effects of retraining the approach bias

A growing research field is the retraining of the approach bias with a cognitive bias modification (CBM). The first CBM was an adaptation of the visual probe task, retraining the attentional bias for anxious pictures (MacLeod, Rutherford, Campbell, Ebsworthy, & Holker, 2002). This scheme has been adapted for drug addiction research and lets participants selectively disengage attention away from drug cues. The drug attentional bias has been shown to be modifiable with such a CBM scheme in tobacco smokers (Attwood, O'Sullivan, Leonards, Mackintosh, & Munafo, 2008; Field, Duka, Tyler, & Schoenmakers, 2009), heavy drinkers (Fadardi & Cox, 2009; T. Schoenmakers, Wiers, Jones, Bruce, & Jansen, 2007) and alcohol-dependent patients (T. M. Schoenmakers et al., 2010). Some of these studies found generalization to new stimuli: training reduced behavioral approach bias scores for both trained and new stimuli (Fadardi & Cox, 2009; T. M. Schoenmakers et al., 2010). However, others failed to find this effect (Field et al., 2009; T. Schoenmakers et al., 2007).

In the AAT-based version of CBM, participant systematically push away drug cues with a joystick in the majority of trials. In heavy drinking students, CBM has been shown to decrease the strength of the approach bias and reduce post-training alcohol intake in heavy drinking students (R. W. Wiers, Rinck, Kordts, Houben, & Strack, 2010). In two randomized-controlled trials in alcohol-dependent patients, CBM reduced relapse rates up to 13% in alcohol-dependent patients as compared with a placebo-training group (R. W. Wiers et al., 2011) and compared with a non-training group (Eberl et al., 2013). Interestingly, R. W. Wiers et al. (2011) compared an implicit and explicit instruction (i.e., either push away according to picture format or according to drink type) and showed this did not have consequences for the effects on relapse rates or decreased approach bias scores. Nevertheless, unconscious or conscious, these findings demonstrate a strong clinical potential of CBM in alcohol addiction.

It remains as yet unknown how CBM leads to decreased relapse in patients. For example, it may be that CBM decreases the incentive salience to drug cues (R. W. Wiers, Gladwin, & Rinck, 2013). In line with this, we recently found that CBM reduces alcohol cue-evoked activity in the amygdala, which correlated with decreases in subjective ratings of alcohol craving (C.E. Wiers, 2014).

2.5. Cognitive models of addiction

A wide variety of cognitive models on drug addiction have been developed that propose working hypotheses on the drug approach bias. By cognitive, we mean models that are committed to explaining the approach bias in psychological or representational terms. That is, these models emphasize rational or associative connections between psychological states and/or emphasize cognitive control processes involved in regulating automatic approach tendencies. Excellent reviews on details of these models have been published elsewhere (e.g., Phaf et al., 2014; Watson et al., 2012)). Here, we provide a brief overview of the most important cognitive models on the basis of which we provisionally address some important predictions about what we would expect to see were these models correct. We go on to argue that embodied components of cognition ought to be addressed in an adequate account of the drug approach bias.

Cognitive models on addiction can be roughly divided into three types: (1) models that stress high motivation to consume drugs (e.g., Robbins & Everitt, 1999; Robinson & Berridge, 2003), (2) models that propose a lack of control processes to resist drug consumption (e.g., Jentsch & Taylor, 1999; Koob & Volkow, 2010), and (3) dual process models of addiction which propose an imbalance of strong motivational processes and weak control to resist drug-taking (e.g., Bechara, 2005; Gladwin, Figner, Crone, & Wiers, 2011; R. W. Wiers et al., 2007).

The process of classical conditioning of motivational reactions to drugs and drug cues are important for motivational models of addiction. Over the course of drug consumption, drug paraphernalia or drug contextual cues become associated with the effects of the drug (Siegel, 1999) and become conditioned stimuli (CS) to drug-effects. The conditioned response (CR) then leads to increased attention as well as approach reactions to drug cues (R. W. Wiers et al., 2007). According to the incentive sensitisation theory of Robinson and Berridge (1993), repeated use of drugs leads to "incentive sensitisation": the neural responses to drugs found in brain regions related to reinforcement and motivation become enhanced. This neural response causes drug cues associated with this brain response to the drug to acquire "incentive salience": the property of, first, attracting attention and, second, of acting as a "motivational magnet". That is, becoming attractive and evoking approach behaviour. In other terms, drugs and drug cues evoke increasing "wanting" (as distinguished from hedonic impact, or "liking"), with dopamine in the midbrain as a key neurobiological substrate of drug-cue learning (Robinson & Berridge, 1993, 2003). Central to habit-theories of addiction is the pairing of drug stimulus and the response: drug taking and drug approach responses then becomes automatic and outside of conscious awareness (Robbins & Everitt, 1999; Tiffany, 1990).

Addiction has also been described as a disorder of disrupted self-control over automatically triggered impulses to use (Baler & Volkow, 2006). The dorsolateral prefrontal cortex (dlPFC) has been shown to be structurally and functionally impaired in drug-dependent individuals, which makes it an important region for the theorized lack of cognitive control (Baler & Volkow, 2006; Bechara, 2005; Hayashi, Ko, Strafella, & Dagher, 2013; Jentsch & Taylor, 1999; Kalivas, 2004; Park et al., 2010; Volkow et al., 2010). Dual process models of addiction are focused on the interaction between top-down and bottom-up processes. There is a wide variety of such models, some of which posit dual systems - an associative, impulsive system in which incentive sensitisation would be located, and a deliberative, reflective system that controls behaviour in order to achieve long-term goals by delaying gratification and inhibiting impulsive behaviour such as drug taking - while others describe different dualities, e.g., between states of processing that bias response selection towards impulsive versus reflective response selection (Bechara, 2005; Gladwin et al., 2011; R. W. Wiers et al., 2007). Despite their differences, dual process models share the common feature of possibly antagonistic interactions between multiple processes or systems that may explain the conflict that typifies addiction: persistent drug taking, even when the individual appears to have an explicit desire to quit (C.E. Wiers, 2014).

2.5 Interim conclusion

In sum, there is increasing evidence that drug-dependent individuals show automatic action tendencies towards drugs cues, which have shown to be of clinical importance. Since the approach bias has also been found for generally positive and anxious pictures, approach reactions to stimuli may be general bodily reactions to positively and negatively valenced stimuli (Phaf et al., 2014). However, motor representations alone may not be sufficient to account for approach and avoidance reactions as (Markman & Brendl, 2005).

The drug approach bias may be explained by motivational mechanisms to approach the drug, or a lack of control to successfully avoid it, or a combination of both, as cognitive models of addiction have previously argued. However, relevant questions on the drug approach bias still remain open and the specificity of the AAT is not known. For example, it remains unexplored whether button presses rather than joysticks are also successful in measuring an approach bias. In the next section, we explore how the drug approach bias may be embodied.

3. Embodied versus cognitive models of explaining addiction

Recent years have seen a wave of interest in embodied explanations of cognition as alternatives to classic cognitive accounts (Varela Thompson & Rosch, 1991; Damasio, 1994, 1999; Hurley, 1998, 2006; Barsalou, 1999, 2002, 2008, 2009, 2010; Bechara, Damasio & Damasio, 2000, 2003; Prinz & Barsalou 2000; O'Regan & Noë 2001; Noë 2005, 2009; Gallagher, 2005). Embodied models of cognition emphasize the role of the body, the environment, and action in grounding psychological states and mechanisms. In contrast to classic computational theories, which rely on abstract, amodal, symbolic information processing as the basis of cognition (Church 1936, Turing, 1936; Fodor, 1975, 1981, 1987; Haugeland, 1978, 1981; Pylyshyn, 1980, 1984; Marr 1983),[1] embodied theories highlight the contribution of situated, bodily and sensorimotor processes in structuring and possibly even constituting cognitive states.[2]

We shall begin by assuming that there are good reasons for seriously considering the role of the body in constituting or substantively influencing implicit, automatic biases in alcohol-dependent patients. Rather than arguing for embodied over cognitive approaches for explaining AAT and the retraining results detailed above, in the following section, we shall begin by examining the evidence that one *should* observe, if an embodied explanation is indeed superior to a cognitive one. Then, in section 4, we shall sketch two distinct accounts, both within an embodied cognition paradigm, that could be used to explain the automatic approach and avoidance tendencies observed in alcohol-dependent patients.

---

[1] We are not claiming that all cognitivist theories are computational (just like we are not claiming that all embodied theories are non-representational) but merely appealing to the computational theory of mind to draw a conceptual contrast with embodied theories.

[2] To be clear, we do not hold it to be conceptually necessary that embodied theories of cognition are anti-representationalist or anti-information-processing. In fact, most embodied theorists are firmly rooted in representationalist, information-processing paradigms (e.g., Barsalou, 1999, 2002, 2008, 2009, 2010; Damasio, 1994, 1999).

Though we will not argue for the virtues of embodied theories over cognitive explanations of automatic approach and avoidance tendencies, we would like to offer the following proposal to researchers interested in differentiating between embodied and cognitive approaches to addiction. We propose that if cognitivist models are best suited to explain AAT then there should be very little or no advantage to using naturally meaningful bodily movements like pushing or pulling towards the actual body to detect or retrain implicit automatic biases. This is because on a cognitive account of the AAT, the quicker reaction times of alcohol-dependent patients to alcoholic stimuli are significant in that they reveal the psychological saliency of alcohol-related cues. That is, there is nothing in the particular pushing or pulling behaviors often used to test approach and avoidance biases that would necessarily be tied to the automatic bias itself. Rather, the bias is coded in the associative or rational cognitive connections between representational states and could become manifest in a variety of equally revealing ways. In short, the connection between the bias and the pushing and pulling movements would remain contingent in a way that it should not on an embodied theory of addiction, which should understand the approach/pulling and avoidance/pushing movements as themselves at least partially constitutive of the automatic biases.

Accordingly, to determine if embodied theories are superior for explaining the above data, one should investigate if verbal report or arbitrary movements such as button presses (e.g., Peeters et al., 2012) are less effective at identifying and retraining AAT than paradigms that exploit the naturally meaningful movements of the body or actions that are connected to the specific activity being investigated. One should also investigate if pushing/pulling towards the actual location of the body is more effective than pushing/pulling towards a projected representation of the self in space (Markman & Brendl, 2005)**.** Specifically, we propose that if the embodied model is correct, then we should find a significant advantage to using joystick pushing/pulling plus visual zoom-out/zoom-in as opposed to button presses that are arbitrarily assigned an "approach" or "avoid" meaning. [3] Moreover, we should see that using the actual physical location of the body will be more effective than pushing or pulling toward a projected representation of oneself, e.g., on a computer screen. Likewise, if the embodied approach is correct, then we should observe a significant advantage in identifying or retraining automatic approach and avoidance biases by using

[3] See Peeters et al. (2012), who use button presses to detect automatic approach and avoidance tendencies. This study is not decisive in choosing between cognitive and embodied approaches since the embodied approach should not be committed to the idea that nothing can be gleaned by using arbitrary movements. The reasonable embodied approach need only be committed to the fact that using movements or actions that are naturally associated with the task being investigated, or the general class of approach and avoidance behaviors, should have a significant advantage over paradigms that do not use this embodied component. The comparison between button presses and a joystick has not yet been investigated.

movements that are closely related to the specific activities being probed or treated.[4] For example, if embodied models best explain the AAT results detailed above, then one should expect that the regular movements involved in grasping a bottle or glass and bringing the vessel to one's mouth and tilting it back to drink could be relevant for identifying or retraining approach and avoidance biases in alcohol-dependent patients. Whether such predictions are born out has yet to be determined.

4. Two accounts of how the approach bias may be embodied

In this section, we explore two accounts of how automatic approach and avoidance biases may be embodied. These accounts are rooted in two distinct approaches to embodiment. It is not our contention that any particular theory of embodiment is necessarily committed to the details of our account. Rather, we hope to show how different approaches to embodiment will entail different predictions for explaining the automatic approach and avoidance biases of alcohol-dependent patients. Moreover, our goal is not endorse either one of these models over the other but, rather, to highlight the conceptual and empirical differences between the two hypotheses as potential explanations of AAT. We will also suggest directions that future research could explore in order to decide between various embodied explanations of automatic approach and avoidance tendencies. Importantly, answering these questions will help to specify the scope of potentially effective therapeutic treatment options for a variety of addictions.

4.1 The biological meaning model

One option for understanding the automatic approach tendencies of drug-addicted patients is what we will call the "biological meaning" model. As embodied theorists have noticed, in virtue of the structures of our bodies and the layout of our worlds, certain gestures and movements may have intrinsic, natural meaning.

As George Lakoff and Mark Johnson argue in their seminal book, "Metaphors We Live By", metaphors are not added to thoughts for aesthetic or rhetorical flourish but, rather, are themselves the basis of conceptual thought. That is, for Lakoff and Johnson (1980a, 1980b), concepts are inherently metaphorical and since we think in and structure experience according to our concepts, both thinking and experiencing are essentially metaphorical as well. They write:

> The concepts that govern our thought are not just matters of the intellect. They also govern our everyday functioning, down to the most mundane details. Our concepts structure

[4] The movements we have in mind are arbitrary but not random or awkward. Obviously, if entirely counterintuitive movements or symbols were used to test the approach/avoidance bias we would see a decrease in accuracy detecting automatic biases simply in virtue of the fact that subjects would have difficulty learning the connection between the arbitrary task and its assigned meaning.

> what we perceive, how we get around in the world and how we relate to other people. Our conceptual system thus plays a central role in defining our everyday realities. If we are right in suggesting that our conceptual system is largely metaphorical, then the way we think, what we experience, and what we do every day is very much a matter of metaphor (1980b, p. 454).

Importantly, according to Lakoff and Johnson, the basic meaning of many metaphors is organized around the body and its natural activities or states such as sleeping or waking, health and death. This is especially relevant for orientation metaphors.[5] As Lakoff and Johnson write, "These spatial orientations arise from the fact that we have bodies of the sort we have and that they function as they do in our physical environment…[s]uch metaphorical orientations are not arbitrary. They have a basis in our physical and cultural experience" (1980a, p.15). Examples of the way the body informs the meaning of concepts is illustrated in the following way:

> HAPPY IS UP; SAD IS DOWN
> I'm feeling *up.* That *boosted* my spirits. My spirits *rose.* You're in *high* spirits. Thinking about her always gives me a *lift.* I'm feeling *down.* I'm *depressed.* He's really *low* these days. I *fell* into a depression. My spirits *sank.*
> Physical basis: Drooping posture typically goes along with sadness and depression, erect posture with a positive emotional state.
>
> CONSCIOUS IS UP; UNCONSCIOUS IS DOWN
> Get *up.* Wake *up.* I'm *up* already. He *rises* early in the morning. He *fell* asleep. He *dropped* off to sleep. He's *under* hypnosis. He *sank* into a coma.
> Physical basis: Humans and most other mammals sleep lying down and stand up when they awaken.
>
> HEALTH AND LIFE ARE UP; SICKNESS AND DEATH ARE DOWN
> He's at the *peak* of health. Lazarus *rose* from the dead. He's in *top* shape.
> As to his health, he's way *up* there. He *fell* ill. He's *sinking* fast. He came *down* with the flu. His health is *declining.* He *dropped* dead.
> Physical basis: Serious illness forces us to lie down physically. When you're dead, you are physically down (1980a, p.16)

---

[5] For more on the spatial organization of human cognition, see Tversky (2009). As she explains, "spatial thinking forms the foundation for other thought" (p.202) and spatial thinking is necessarily tied to the boundaries, shape, and size and function of our bodies, the space of possible action, and our access to the distal layout of the environment which surrounds us.

Similarly, In *How the Body Shapes the Mind,* Shaun Gallagher draws our attention to the fact that both phenomenal experience and intentional thought are constrained and shaped by our physical bodies. Gallagher "explores the ways the body plays an active role in shaping perception and action, its functional roles in enabling intentionality, and the constraints and possibilities defined by the shape and structure of the human body" (2005, p. 134). For instance, Gallagher explains the intimate relationship between bodily posture and cognitive function. He writes:

> Teachers often say to young students: "Now sit up straight and pay attention." There is some truth in the teacher's coaxing. Perception and attention cannot be uncoupled from the body's postural attitudes. Consider the experiments conducted by Kinsbourne and others on the effects of body posture on judgment and attention (Grubb and Reed 2002; Kinsbourne 1975; Lempert and Kinsbourne 1982), which show that the lateral position of head and eyes or whole body influences cognitive performance (Gallagher, 2005, p.140-141).

Further, in considering the global impacts of upright posture on human cognitive capacities, Gallagher writes:

> with the upright posture we gain distance and independence. Distance from the ground; distance from things; independence from other people. In standing, the range of vision is extended, and accordingly the environmental horizon is widened and distanced. The spatial frameworks for perception and action are redefined. Things are less close, less encountered as one crawls among them… Standing frees the hands for gnostic touching, manipulation, carrying, tool use and for pointing (a social gesture), all of which transcend grasping. At the same time, these functional changes introduce complexities into a brain structure that is being redesigned for rational thought. Standing also brings us 'face to face' with each other, and this profoundly transforms sexuality from strict animality to something human (2005, p. 148).

Gallagher's point is that basic morphological and structural features of the body, as well as movements and actions that are characteristic of the human species have robust impacts on how we encounter the world and the cognitive capacities that are available to us as a result. According to Gallagher, the structure of the body is an essential ingredient for understanding the nature of human cognition. From this perspective, the possibility of a disembodied brain-in-a-vat with cognitive states and capacities identical to ours is a philosopher's fiction. This is because the disembodied brain would not have the natural constraints of the body to organize and ground cognition and so it could not have the same cognitive capacities and states that we enjoy.

Taking inspiration from these kinds of approaches to the embodied organization of perceptual and intentional meaning, when we return to consider the automatic approach and avoidance biases of alcohol-dependent patients, we should attend to the structure and function of the human body. On a “biological meaning” model, we should notice that, as humans, the vast majority of our internal organs are located toward the center of our bodies, at our core, and not at our peripheries. That is, our vital organs are not in our arms or legs but nearer to the center of our bodies. Furthermore, without the protection of a substantial bone mass, much of the area that houses our internal organs turns out to be remarkably vulnerable.[6] From a biological perspective, it would be reasonable for humans to develop a strong predisposition to protect this area from harm and to limit contact with the center body to nonthreatening objects and persons. Of course, humans should have a predisposition to protect every bodily area from harm. The point here is to highlight the fact that there should be a particularly strong tendency associated with the center body for its protection.

Embodied cognition theorists of the “biological meaning” kind should hold that the morphology of the human body makes it natural that pulling something towards one’s core or approaching it would be laden with positive affect, a sign, potentially, of liking, wanting, needing and trusting. They should also notice that pushing something away would naturally be assigned a negative valence, meaning potentially that it is disliked, rejected or not to be trusted. Since pulling something towards oneself is a costly maneuver, it stands to reason that approach or pulling behaviors will be reserved for those things that are desired and needed. Likewise, pushing something away or avoiding would signify rejection. In this way, we can see how our bodies shape the meaning of movements and endow movement with an intrinsic affective component, which is grounded in our evolutionary, biological history.

We should note that since the entire body is vulnerable, one could argue that there is nothing noteworthy about the most vulnerable area at the center of the body that gives it priority for grounding the natural meaning of pushing and pulling, approach and avoidance tendencies. However, it is important to notice that various other constraints and demands make it likely that this protective tendency is especially strong for the center body. After all, demands for finding food or investigating one’s surroundings make it likely that one’s arms and hands should sometimes move towards or approach objects, even in conditions of uncertainty. Similarly with legs and feet, which will likely make contact with uninspected objects on a regular basis as a result of the demands of mobility. As such, the cost of not protecting one’s own organs coupled with the lack of a competing function for the center body makes it especially likely to ground the natural meaning of pull and push, approach and avoidance behaviors.

---

[6] Obviously, this applies less to the organs that are protected by ribs.

Another biological consideration that could be seen as relevant for explaining automatic approach biases comes from the fact that accessing nourishment for humans requires placing external parts of the world into our mouths. That is, survival requires pulling nourishing objects towards oneself. The fact that survival requires the grasping of an external object and pulling it towards oneself in order to avoid caloric deficit suggests that pulling, as a motion, should be strongly positively valenced. We should notice, however, that the avoid or pushing aspect of natural meaning would not be easily explained if we take nourishment as central to the biological meaning of approach and avoidance behaviors. This is not to say that eating contributes nothing to the biological meaning of pulling behaviors but it is to say that if we are looking for a symmetrical understanding of pushing and pulling, approach and avoidance, the location of the internal organs seems to account for both while nourishment can be seen as strengthening the natural meaning of approach or pulling.

Further, since for the purposes of procreation and child rearing, the proximity of other individuals to one's own body is required, pulling towards oneself should be strongly positively valenced. From an evolutionary perspective, it is likely that approach or pulling something towards oneself would require a differentiation between what is liked or loved, what is one's own, and what is rejected or neutral. Again, we see that pulling towards oneself, or approach, would be naturally meaningful for the purpose of selecting mates, nourishing and nurturing infants.

If this is the correct way to understand approach and avoidance we should expect that pulling towards oneself and pushing away from oneself are universal across cultures and generations. If meaning is grounded in the basic biological features of the human body, such as the location of our vital organs, the necessity to ingest food for nourishment, and the requirement of physical proximity for procreation and child-rearing, we can see how it is that approach or pulling behaviors and avoidance or pushing ones would evolve as gestures or movements containing intrinsic positive and negative affect, respectively.

When it comes to explaining the enhanced automatic approach biases of alcohol-dependent patients in response to alcohol-related cues, on the biological meaning model, we should say that alcohol elicits an acquired and reinforced like/ want/ need/pull/ approach bias. This explanation would fuse wanting or liking with approaching at the level of the biological body. That is, the biological meaning explanation would posit that the positive, affective state connected to alcoholic cues—which is learned through reinforcement by alcohol consumption—just is a state which moves one towards alcohol-related cues, or pulls alcohol-related cues towards oneself. The approach behavior, on this model, would not be a response to wanting or liking, but intrinsic to liking and wanting. That is, liking or wanting would not be seen as triggers of approach behavior but, rather, approach behavior would itself be swaddled with meaning: meaning like desiring, wanting, needing, etc. So, when the alcohol-dependent patient wants alcohol that wanting is itself a movement towards the alcohol. The movement is not added to the wanting but a constituent part of it.

This interpretation of wanting as fundamentally tied to approach behaviors is in line with Antonio Damasio's (1994, 1999, 2001) theory of emotions, where emotions are essentially responses meant to direct animals toward advantage and away from harm. As Damasio writes:

> [E]motional responses are a mode of reaction of brains that are prepared by evolution to respond to certain classes of objects and events with certain repertoirs of action… emotions allow organisms to cope successfully with objects and situations that are potentially dangerous or advantageous…[they are] the most visible part of a huge edifice of undeliberated biological regulation that includes homeostatic reaction that maintain metabolism, pain signaling, and drives such as hunger and thirst (2001, p. 781).

Taking this account seriously, we should not be surprised to see that wanting or liking is necessarily tied to approaching or pulling. That is, because desire plays the functional role of moving an animal towards the object desired, it is natural to interpret a heightened desire with a faster response in approaching the desired object.

Combining the identification of wanting and approaching with the structural considerations of the body discussed above, we have reason to expect that pushing or pulling away from the front, center body will be more naturally meaningful than pushing or pulling movements related to the sides of the body, or head and feet. That is, the biological meaning model would predict that training paradigms that exploit the natural meaning of approach and avoidance relative to the morphological vulnerabilities of the center body would be more illuminating and effective in the retraining paradigms than non-naturally meaningful movements. That is, if the biological meaning account of AAT is correct, one should expect that pushing and pulling towards and away form the center body to retrain AAT should be more effective then the arbitrary assignment of push/pull to a symbolic context such as assigning one letter or symbol on a keyboard to "approach" and another to "avoid". We should also expect that pushing and pulling towards the actual physical body should be more effective than pushing or pulling towards a projected representation of the body in space. This is not to predict that no such methods could gain any significant result in CBM. Presumably, it would still be possible to form various temporary associations with arbitrary symbols and representations and, through those associations, glean an informative and potentially effective set of data. It is only to say that if the biological meaning story is correct, one would expect using naturally meaningfully movements to be more efficient and effective than assigning a contingent meaning to arbitrary symbols or locations in space.

Moreover, one would expect that if the AAT and retraining via CBM of alcohol-dependent patients is effective due to its exploitation of natural, biological meaning, then we should expect that most other addictions or disorders that distort motivations in favor of unhealthy or harmful automatic implicit approach and avoidance biases could also effectively be treated using similar methods. That is, the specific addiction or disorder that one is looking to treat would not need to share a common set of epistemic or doxastic states, or sensorimotor routines or practices in order for a treatment method based in biological meaning to be effective. As such, if the biological meaning model is the correct explanation of why AAT and retraining of AAT is effective in alcohol addiction, then we would predict that retraining of AAT by using the joystick push and pull plus zoom, would be effective in smoking, gambling, anxiety disorders, anorexia, etc. That is, if the reason for AAT and the effectiveness of retraining of automatic approach and avoidance impulses piggybacks on the general, universal, evolutionary meaning of gestures such as pushing as negative and pulling as positive, then one should expect that using these very same movements can be effective for the retraining of many distorted impulses that share little more in common than the fact of their distortion of automatic impulses to approach what is harmful and/or avoid what is beneficial.

4.2 The sensorimotor hypothesis

A second embodied explanation of the existence of the automatic approach bias observed in alcohol-dependent patients and the subsequent impact of retraining on the bias through CBM is what we will call "the sensorimotor hypothesis". In contrast to the biological meaning model, the sensorimotor hypothesis accounts for automatic approach and avoidance biases by appeal to the specific sensorimotor loops that are established through the regular instantiation of alcohol-consumption and alcohol-related activities. That is, as opposed to appealing to the evolutionary, biological meaning of approach and avoidance behaviors, which become fused with alcohol in alcohol addiction, the sensorimotor hypothesis focuses on the particular sensoriomotor connections that are acquired and canalized through alcohol addiction to explain why approach and avoidance biases appear in alcohol-dependent patients.

Our sensorimotor explanation of automatic approach and avoidance tendencies takes its inspiration from Kevin O'Regan and Alva Noë's (2001) account of enactive perception, Lawrence Barsalou's (1999, 2002, 2008, 2009, 2010) theory of grounded cognition, and Friedemann Pulvermuller's (1999, 2008) theory of embodied language. The most salient aspect of these diverse accounts for our purposes is their emphasis on the particularity or situatedness of sensorimotor processes in explaining perception, concepts, and language, respectively.

Enactive perception theorists such as Susan Hurley (1998), O'Regan and Noë (2001), and Noë (2005) hold that sensorimotor skill or the tacit understanding of sensorimotor contingencies is constitutive of the qualitative character of a perceptual event. On their view, implicit understanding of the constantly changing sensorimotor dependencies between a moving agent and a perceptual array determine the phenomenal character of perceptual experience. That is, what we expect about how the look of an object will change relative to our perspective and position, determines what we perceive. This knowledge is based on our embodied skills and interactions with the world. As Noë writes,

> Perceptual experience acquires content thanks to our possession of bodily skills. *What we perceive* is determined by *what we do* (or what we know how to do); it is determined by what we are *ready* to do…To be a perceiver is to understand, implicitly, the effects of movement on sensory stimulation…*the enactive approach* is that our ability to perceive not only depends on, but is constituted by, our possession of this sort of sensorimotor knowledge (2005, p. 1). *Emphasis in original.*

The enactive view of perception relies on sensorimotor skills in order to account for the determinate content and character of perceptual experience. For our purposes, an important feature of this view is that our particular sensorimotor skills, acquired through particular experiences, and dependent on our particular body types and the particular way in which those body-types can navigate or traverse space, are crucial for determining the character of our perceptual states. On the enactive view, it is the particular sensorimotor skills rather than the general nature of sensorimotor interaction that is relevant for understanding the nature of perceptual experience. As O'Regan and Noë write:

> For two systems to have the same knowledge of sensorimotor contingencies all the way down they will have to have bodies that are identical all the way down (at least in relevant respects). For only bodies that are alike in low-level detail can be functionally alike in the relevant ways (2001, p. 1015).

The focus on the specificity of the body, which constrains experience and knowledge, is the feature of the sensorimotor view of perception that we'd like to emphasize. Our explanation of the automatic approach and avoidance bias in alcohol-dependent patients, of course, is not an investigation into why a perceptual stimulus of, e.g., a beer bottle or wine glass, looks a certain way to an alcohol dependent patient, but the emphasis on the particularity of experience in constituting perception is nonetheless relevant for our version of the sensorimotor hypothesis.

This feature becomes even more important when we connect it to a theory of grounded cognition, like the one defended by Barsalou (1999, 2002, 2009, 2010) who claims that conceptual knowledge is not abstract or amodal, but perceptual in nature. That is, we can incorporate the specificity of enactive perception with the specificity of concept formation, described by Barsalou, to produce a general sensorimotor picture that highlights the particularity of experience in constituting cognitive processes.

According to Barsalou, concepts are grounded in our perceptual experiences of objects, persons, and events. That is, concepts are not abstract, amodal, symbols, but retain their sensorimotor origins. As Barsalou (2002) writes, "rather than being detached from events, knowledge remains grounded in them"(p.1). Further,

> Various perceptual experiences with the same object, event, or individual are stored together in associated neural networks. When the same or similar object is subsequently observed, imagined, or considered in thought, the sensorimotor systems that were active in the original perceptual experience run a simulation or a reenactment of the experience. As the simulations are run in various circumstances, additional associations are formed that further shape or expand the concept or simulator…Further, not every simulation or aspect of a simulation is active in any one experience or encounter with an instance of a category (Barsalou, 1999 p. 585-7).

Importantly, particular conceptualizations or simulations that constitute a concept are always context or situation-dependent. That is, "the concept does not represent the category in isolation, independently of the situation in which it occurs" Rather, "situations are fundamental to cognition" (Barsalou, 2002 p.1-3). For example, the specific background where, e.g., a chair is observed, say a living room or an office, is retained in the conceptualization (Barsalou, 2002). These features are not lost or abstracted away in the formation of the concept. From this, it follows that each individual's experience with the conceptual category will determine the shape and scope of the concept that she possesses. The particulars of the perceptual experience are retained as part of the concept and relevant for the inferences and predictions that will be most salient when a person deliberates, plans, or engages in other cognitive processes that involve the concept (Barsalou, 1999).

Additionally, the relevant perceptual experiences, according to Barsalou, are not only multimodal but can be sensory, proprioceptive, and introspective.[7] As such, concepts are not limited to coding objective features of the external world but can incorporate various internal bodily feelings, emotional states, and/or cognitive processes. This means that the perceptual experiences that become associated with a particular category will often be holistic in nature. As Barsalou writes:

> Besides visual information, the event sequence [of

---

[7] By proprioceptive experiences, Barsalou (1999) has in mind the internal bodily sensations involved in, e.g., lifting or running. Introspective experiences also include the experience of emotions or moods such as, happiness or hunger.

> encountering a car] might include the proprioceptive experience of pressing the pedal, the auditory experience of hearing the engine roar, the haptic experience of feeling the car vibrating, and mild excitement about the power experienced… After processing many cars, a tremendous amount of multimodal information becomes established that specifies what it is like to experience cars sensorially, proprioceptively, and introspectively. In other words, the frame for car contains extensive multimodal information of what it is like to experience this type of thing (1999, p. 587).

As such, the richer one's experience with a particular category of object, action, event, or individual, the richer the associative network of simulations related to those categories will become. Again, as with the enactive theory of perception, what we see in Barsalou's theory of grounded concepts is that the particular, situated, sensorimotor experiences of an individual with a category are central to concept formation. Once again, the specifics of experience are vital for understanding the formation of cognitive categories.

In a similar vein, when providing a theory of embodied language, Pulvermuller (1999, 2008) draws on the particular sensorimotor experiences that one has when learning a word to ground its semantic value. Pulvermuller relies on a Hebbian framework (1949) to explain embodied language, appealing to the now popular notion that "neurons that fire together, wire together." As Pulvermuller writes,

> If a particular object is frequently being visually perceived, a set of neurons in the visual cortices will repeatedly become active a the same time. Therefore, cell assemblies will form representing the shape of the object… if motor behavior co-occurs with sensory stimulation cell assemblies may form including neurons in motor and sensory cortices. To put it in a more general way: the cortical localization of a representation is a function of where in the cortex simultaneous activity occurred when the representation was learned (1999, p. 259).

By applying the Hebbian framework to learning situations involved in language acquisition, Pulvermuller, like O'Regan and Noë (2001) and Barsalou (1999, 2008, 2009), focuses on the actual, embodied experiences of an agent in order to account for cognitive capacities. Pulvermuller draws on the particular sensorimotor processes active during language learning in order to ground the semantic meaning of words in their sensorimotor associations. Pulvermuller writes:

> If language is not learned through vocal and auditory modalities, but through the manual and visual modalities (sign language) cortical localization of cell assemblies representing meaningful elements should be different (1999, p. 261).

The position that Pulvermuller is advocating is not just emphasizing the general modality specific features that are associated with world-learning and use (such as audition or vision), but goes further to posit that fine-grained differences in semantic meanings should be reflected in fine-grained differences in sensorimotor activity. For instance, the words “kick” and “lick” should activate distinct brain areas dedicated to the feet and the mouth respectively.

Again, we are confronted with the idea that specific experiences that involve sensorimotor channels are essential for understanding cognitive function. In each of the above accounts, whether of perception, concepts, or language, we see that the sensorimotor views ground cognitive function in the particular experiences of the cognitive agent. That is, in each of the above accounts, the particular experiences that an agent has with a category or stimulus constitute the content and character of that agent’s cognitive processes.

Therefore, when we move to explaining the automatic approach and avoidance tendencies of alcohol-dependent patients on a sensorimotor view, we should focus on the particular sensorimotor experiences that alcoholic-dependent individuals will have had with alcohol-related stimuli. According to the sensorimotor hypothesis, we should not only appeal to the natural, species-wide, biological meaning of approach and avoidance behaviors that become fused with alcohol-related cues, but rather, attend to the particular sensorimotor loops that are enacted and reenacted in the habitual activity of alcohol consumption.

It follows that the sensorimotor account should appeal to the regular range of movements involved in drinking, the visual and olfactory features of alcohol, together with the affective and emotional states that alcohol evokes, as the basis of the approach bias in alcoholic patients. We should focus on the fact that the particular perception of a beer bottle or wine glass or highball glass would be tied to the particular movement of lifting the bottle or glass and bringing it to one’s lips, lifting one’s chin back, raising one’s elbow, replacing the glass etc. This sensorimotor loop would then be laden with various affective states as a result of the reinforcement received by the addict when consuming alcoholic beverages.

In this way, the particular actions and perceptions that are involved in alcohol dependency are appealed to in order to ground the automatic approach and avoidance biases of alcohol-dependent patients. It is vital to notice that the automatic approach bias on the sensorimotor story must be accounted for by the entire sensorimotor loop, or, at the very least, by the fact that drinking alcohol requires bringing alcoholic beverages towards oneself in order to consume them, that is, pulling alcohol towards oneself or approaching it.

4.3 Contrasting the biological meaning model with the sensorimotor hypothesis

Let us now contrast the sensorimotor hypothesis with the biological meaning story. On the sensorimotor view, the alcohol-dependent patient's sensorimotor habits should account for the automatic approach bias. However, if the component of the sensorimotor loop that does the work in explaining AAT is the affective state and its natural approach tendency then we are back to the biological meaning story. This is not to say that a biological meaning story could not be tied to the sensorimotor hypothesis as an account of AAT in alcohol-dependent patients. It is to say, however, that if it were the natural meaning of approach/desiring that accounts for the automatic approach bias, then the sensorimotor hypothesis would not be the relevant factor in explaining AAT, even though sensorimotor loops could still be relevant for understanding how affective states get attached to particular objects or events.

Accordingly, it seems plausible that if the sensorimotor hypothesis is the correct interpretation of the automatic approach bias then the fact that approach is observed in alcohol-dependent patients should be contingent upon the fact that drinking alcohol requires lifting and pulling beverages towards oneself. Likewise, when considering the retraining of automatic approach biases through the use of the push/pull joystick, one would expect that the effectiveness of the CBM joystick paradigm would be explained by appeal to either the breaking of a sensorimotor loop that has been previously reinforced and ingrained into the repertoires of alcohol-dependent patients or in the formation of a new competing sensorimotor loop.

Specifically, since perception of alcohol in alcohol-dependent patients habitually elicits a motor response similar enough to the pulling of the joystick (that is, reaching and grasping a glass and moving it towards one's mouth) it is postulated that one could identify the addictive behavior by presenting an alcohol-related stimulus and measuring automatic approach tendencies with a joystick pull. Likewise, a pushing response to the alcoholic cues, instead of pulling response could either break the reinforced sensorimotor loop or create a new, competing sensorimotor loop that is tied to the perception of alcohol-related cues.

Notably, the predictions that one would make if the sensorimotor hypothesis accounted for the automatic approach bias and it's retraining would be importantly distinct from the predictions made by biological meaning model. Generally, the sensorimotor hypothesis should have particular, fine-grained requirements for detecting, breaking, and creating new sensorimotor loops. This would require attention to both the particular perceptual stimulus involved in the habitual behavior and the particular motor response that the stimulus elicits. If the sensorimotor hypothesis is correct, then, presumably, it should be more difficult to detect addictions to gambling or exercise by using the AAT. This is because, presumably, the motor routines involved in gambling or exercises such as running and cycling, do not involve anything like the pulling motions that alcohol consumption requires.

Likewise, if the sensorimotor hypothesis is correct then the scope of addictions and disorders that one would expect the joystick CBM paradigm to be effective in treating will be much more narrow than on the biological meaning model. For instance, it isn't obvious that the joystick CBM is sufficiently similar to the motor routines involved in shooting heroin or snorting cocaine for this method to be effective in retraining the automatic biases of patients addicted to those drugs. Further, for addictions or disorders that do not involve the ingestion or consumption of substances, such as gambling or sex-addiction, the joystick would seem to be a poor method of retraining since the sensorimotor loops involved in manually pushing and pulling images would presumably be sufficiently unlike the sensorimotor routines involved in gambling and sex.

One would expect, however, that addictions to cigarettes and other inhalable drugs could be retrained using the joystick paradigm. Though, one would expect that retraining using the particular motor response involved in the target behavior should be more effective than using a sensorimotor loop that only loosely resembles the actual motor routines involved in the addictive behavior. Further, one would expect that similarity of perceptual stimuli should likewise enhance the effectiveness of the CBM paradigm since the automatic approach bias is hypothesized to be rooted in the particular experiences and habits of the addicted individuals. As such, one would predict that an alcohol dependent patient who prefers beer to wine or spirits will respond better to pushing images of beer away from herself rather than to pushing away other kinds of alcoholic stimuli.

It is worth noting that on the biological meaning model, it isn't clear if the perceptual stimulus of a particular alcoholic beverage or container should elicit a stronger approach bias or a more effective stimulus for retraining than any member of the class of alcoholic beverages. It seems to us that the biological meaning approach is neutral on this question. That is, there is nothing about the model that prevents it from being filled out in either a more general or fine-grained direction. However, we also think it is worth highlighting that on a cognitivist explanation of AAT, belonging to the general class of alcoholic stimuli should likely suffice for eliciting an approach bias. This is because on a cognitive model, it is not simply a particular set of stored representations or episodic memories that are linked to an affective state, but an association between the affective state and a conceptual category that triggers the automatic approach response in alcohol dependent patients.[8] This difference could be useful in beginning to differentiate between competing paradigms.

5. Breaking old loops or creating new ones?

[8] Of course, a cognitivist could appeal to particular representations and memories as the basis for associative links with affective states. Still, we should notice that if a high degree of match between a particular stimulus and the actual perceptual experience of the agent were relevant for explaining AAT and retraining, then this would eliminate at least one form of cognitive explanation from adequately accounting for the above results.

One last distinction that we address in terms of its conceptual implications for understanding the AAT concerns the way to interpret the results of retraining using the joystick paradigm on the sensorimotor hypothesis. As we stated above, on the sensorimotor view, retraining can be understood in one of two ways: either retraining replaces or weakens existing sensorimotor loops that are responsible for automatic approach biases or retraining creates new and competing sensorimotor loops. To our knowledge, the correct way to interpret the effects of retraining on the sensorimotor hypothesis are still open.

Considerations presented by both Pulvermuller (1999) and Barsalou (1999) weigh in favor of interpreting retraining of automatic biases as a matter of replacing or weakening existing sensorimotor connections. As Pulvermuller (1999) discusses, the proper way to understand Hebbian learning is not only to think of the strengthened connections that occur during synchronous activation of neuron populations (the LTP or long term potentiation) but also in terms of the weakening of connections that occur when neural populations do not fire synchronously (long term depression or LTD). That is, to understand how assemblies of cells become fused together into functional units, one should consider both potentiation and depression. This way of considering the retraining results achieved with pushing the joystick in response to alcohol-related cues would indicate that the approach bias that is wired into the sensorimotor loop of alcohol-dependent patients weakens when the automatic approach or pulling behavior is not activated by alcoholic stimuli.

On Barsalou's account, conceptual categories are dynamic in that they continuously incorporate new experiences or conceptualizations and adjust saliency of expected features as a result of subsequent perceptions. As such, the shape or extent of a category is adjusted in light of new experiences and information. As Barsalou writes, "the subsequent storage of additional perceptual symbols in the same association area may alter connections in the original pattern causing subsequent activations to differ" (1999 p. 585). If the retraining of automatic approach and avoidance biases is to be understood on Barsalou's theory, then we should conclude that pushing as opposed to pulling in response to alcohol-related cues replaces or weakens connections between existing sensorimotor networks.

However, we should notice that the replacing or weakening of sensorimotor loops is not the only conceptually possible explanation of the joystick-retraining paradigm. As Robinson and Berridge (2003) argue, automatic associative affective states are at least semi-permanent and hence relatively fixed or unbreakable. Once created, they are there to stay. If this view is correct then one could posit that the retraining of automatic biases, instead of weakening existing connections creates new competing ones. In this way, the alcohol-addicted patient does not lose their automatic approach bias but gains an automatic avoidance bias. These two biases compete. Hence, retraining cultivates in alcohol-dependent patients a resistance to the automatic approach bias that ruled unchecked prior to the CBM.

We should also notice that these two ways of interpreting the retraining of automatic biases fit naturally with two different general theories of addiction. The replacement or weakening option is coherent with the view that addiction is largely an implicit affective, motivational disorder (e.g., Robinson & Berridge, 2003; R.W. Wiers et al, 2013). As such, changing or replacing one's disordered implicit, automatic affective and motivational states would be sufficient for countering the force of the addiction. The competing automatic bias account is more in line with the dual-process theory of addiction since, presumably, the resistance to the original automatic approach bias that is offered by the competing automatic avoidance bias buys time for the rational control system to step in and decide to act in line with the alcohol-dependent patient's considered judgments. Both options open up opportunities for effective treatment and for understanding the nature of addiction.

6. Conclusions

In the first half of the paper we reviewed empirical data on the drug approach bias in various drug-dependent individuals. Behaviorally retraining the approach bias with bias modification training has been demonstrated to reduce craving, relapse and brain activations in reward-related areas. Approach reactions to stimuli may be general bodily reactions, although motor representations alone have been shown to not be sufficient to explain the bias.

In the second half of the paper, we presented two embodied accounts of the AAT. According to the biological meaning model, the natural, biological structure and function of the human body grounds meaning for a variety of gestures and actions, such as pushing and pulling. We predicted that if the biological meaning model best explains the mechanisms underlying the AAT and associated retraining effects, then the joystick/zoom paradigm should be clinically effective in treating a variety of addictions, just so long as they exhibit disordered biases. On the other hand, we argued that if the sensorimotor model was the best explanation of the AAT, then the specific activities involved in a particular addiction, that is, the specific sensorimotor loops constituting addictive behavior, would be relevant for both explaining the AAT and the retraining results. As such, successful treatment would have to involve retraining or replacing the particular sensorimotor loops that have become canalized through the process of addiction.

In spelling out the conceptual commitment and implications of these models, we hope that we have both clarified the foundations of the embodied theories as well as presented useful empirical predictions for differentiating between them. Lastly, we have also gestured to some important empirical work that could decide whether any embodied model of addiction is better at explaining the AAT than competing cognitive models. We trust that differentiating between competing models will engender effective applications for addiction treatment in the future.

Acknowledgements

CEW was funded by the Berlin School of Mind and Brain and Humboldt Graduate School.

References

Attwood, A. S., O'Sullivan, H., Leonards, U., Mackintosh, B., & Munafo, M. R. (2008). Attentional bias training and cue reactivity in cigarette smokers. *Addiction, 103*(11), 1875-1882. doi: 10.1111/j.1360-0443.2008.02335.x

Baler, R. D., & Volkow, N. D. (2006). Drug addiction: the neurobiology of disrupted self-control. *Trends Mol Med, 12*(12), 559-566. doi: 10.1016/j.molmed.2006.10.005

Barkby, H., Dickson, J. M., Roper, L., & Field, M. (2012). To approach or avoid alcohol? Automatic and self-reported motivational tendencies in alcohol dependence. *Alcohol Clin Exp Res, 36*(2), 361-368. doi: 10.1111/j.1530-0277.2011.01620.x

Barsalou, L.W. (1999). Perceptual symbol systems. *Behavioral and Brain Sciences,* 22 (4), 577-660.

Barsalou, L.W. (2002). Being there conceptually: Simulating categories in preparation for situated action. In N.L. Stein, P.J. Bauer, & M. Rabinowitz (Eds.), *Representation, memory, and development: Essays in honor of Jean Mandler* (pp. 1-19). Mahwah, NJ: Erlbaum.

Barsalou, L.W. (2008). Grounded cognition. *Annual Review of Psychology*, 59, 617-645.

Barsalou, L.W. (2009). Simulation, situated conceptualization, and prediction. *Philosophical Transactions of the Royal Society of London: Biological Sciences,* 364, 1281-1289.

Barsalou, L.W. (2010). Grounded cognition: Past, present, and future. T*opics in Cognitive Science,* 2, 716-724.

Bechara, A., Damasio, H., & Damasio, A.R. (2000). Emotion, decision making and the orbitofrontal cortex. *Cerebral Cortex,* 10 (3), 295-307.

Bechara, A., Damasio, H., & Damasio, A.R. (2003). Role of the amygdala in decision-making. *Annals of the New York Academy of Sciences,* 985, 356-69.

Bechara, A., Damasio, H., Tranel, D., & Damasio, A.R. (2005). The Iowa gambling task and the somatic marker hypothesis: some questions and answers. *Trends in Cognitive Sciences*, 9 (4), 159-162.

Bechara, A. (2005). Decision making, impulse control and loss of willpower to resist drugs: a neurocognitive perspective. *Nat Neurosci, 8*(11), 1458-1463. doi: 10.1038/nn1584

Blumstein, H., & Schardt, S. (2009). Utility of radiography in suspected ventricular shunt malfunction. *J Emerg Med, 36*(1), 50-54. doi: S0736-4679(07)00843-8 [pii] 10.1016/j.jemermed.2007.06.044

Bradley, B. P., Field, M., Healy, H., & Mogg, K. (2008). Do the affective properties of smoking-related cues influence attentional and approach biases in cigarette smokers? *J Psychopharmacol, 22*(7), 737-745. doi: 10.1177/0269881107083844

Brown, E. C., Tas, C., Kuzu, D., Esen-Danaci, A., Roelofs, K., & Brune, M. (2014). Social approach and avoidance behaviour for negative emotions is modulated by endogenous oxytocin and paranoia in schizophrenia. *Psychiatry Res*. doi: 10.1016/j.psychres.2014.06.038

Chen, M., & Bargh, J. A. (1999). Consequences of automatic evaluation: Immediate behavioral predispositions to approach or avoid the stimulus. *Personality and Social Psychology Bulletin, 25*(2), 215-224. doi: Doi 10.1177/0146167299025002007

Cousijn, J., Goudriaan, A. E., & Wiers, R. W. (2011). Reaching out towards cannabis: approach-bias in heavy cannabis users predicts changes in cannabis use. *Addiction, 106*(9), 1667-1674. doi: 10.1111/j.1360-0443.2011.03475.x

Cox, W. M., Fadardi, J. S., & Pothos, E. M. (2006). The addiction-stroop test: Theoretical considerations and procedural recommendations. *Psychol Bull, 132*(3), 443-476. doi: 10.1037/0033-2909.132.3.443

Damasio, A.R. (1994). *Descartes' error: Emotion, reason, and the human brain*. Putnam.

Damasio, A.R. (2001). Fundamental feelings, *Nature*, 413, 781. doi:10.1038/35101669

Damasio, A.R. (1999). *The feeling of what happens: Body and emotion in the making of consciousness*. Harcourt Brace and Co.

De Houwer, J. (2006). What are implicit measures and why are we using them. In R. W. Wiers & A. W. Stacy (Eds.), *The handbook of implicit cognition and addiction* (pp. 11–28). Thousand Oaks: Sage.

Eberl, C., Wiers, R. W., Pawelczack, S., Rinck, M., Becker, E. S., & 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

Ernst, L. H., Plichta, M. M., Dresler, T., Zesewitz, A. K., Tupak, S. V., Haeussinger, F. B., Ehlis, A. C. (2014). Prefrontal correlates of approach preferences for alcohol stimuli in alcohol dependence. *Addict Biol, 19*(3), 497-508. doi: 10.1111/adb.12005

Fadardi, J. S., & Cox, W. M. (2009). Reversing the sequence: reducing alcohol consumption by overcoming alcohol attentional bias. *Drug Alcohol Depend, 101*(3), 137-145. doi: 10.1016/j.drugalcdep.2008.11.015

Field, M., Duka, T., Tyler, E., & Schoenmakers, T. (2009). Attentional bias modification in tobacco smokers. *Nicotine Tob Res, 11*(7), 812-822. doi: 10.1093/ntr/ntp067

Field, M., Mogg, K., Mann, B., Bennett, G. A., & Bradley, B. P. (2013). Attentional Biases in Abstinent Alcoholics and Their Association With Craving. *Psychology of Addictive Behaviors, 27*(1), 71-80. doi: Doi 10.1037/A0029626

Fodor, J. (1975). *The language of thought.* New York: Thomas Crowell.

Fodor, J. (1981). *Representations*. Cambridge, MA: Bradford Books/MIT Press.

Fodor, J, (1987). *Psychosemantics*. Cambridge, MA: Bradford Books.

Gallagher, S. (2005). *How the body shapes the mind*. Oxford: Clarendon Press.

Gladwin, T. E., Figner, B., Crone, E. A., & Wiers, R. W. (2011). Addiction, adolescence, and the integration of control and motivation. *Dev Cogn Neurosci, 1*(4), 364-376. doi: 10.1016/j.dcn.2011.06.008
Haugeland, J. (1978). The nature and plausibility of cognitivism. *Behavioral and Brain Sciences*, 2, 215–226.
Haugeland, J. (1981). *Mind design*. Cambridge, MA: MIT Press/Bradford Books.
Hayashi, T., Ko, J. H., Strafella, A. P., & Dagher, A. (2013). Dorsolateral prefrontal and orbitofrontal cortex interactions during self-control of cigarette craving. *Proc Natl Acad Sci U S A*. doi: 10.1073/pnas.1212185110
Hebb, D.O. (1949). *The organization of behavior*. New York: Wiley & Sons.
Heinz, A., Beck, A., Grusser, S. M., Grace, A. A., & Wrase, J. (2009). Identifying the neural circuitry of alcohol craving and relapse vulnerability. *Addict Biol, 14*(1), 108-118. doi: 10.1111/j.1369-1600.2008.00136.x
Heuer, K., Rinck, M., & Becker, E. S. (2007). Avoidance of emotional facial expressions in social anxiety: The Approach-Avoidance Task. *Behav Res Ther, 45*(12), 2990-3001. doi: 10.1016/j.brat.2007.08.010
Hurley, S.L. (1998). *Consciousness in action*. Cambridge, MA: Harvard University Press.
Hurley, S.L. (2006). Active perception and perceiving action: The shared circuits model. In T. Szabo Gendler & J. Hawthorne (Eds.), *Perceptual experience*. Oxford: Oxford University Press.
Hyman, S. E., & Malenka, R. C. (2001). Addiction and the brain: the neurobiology of compulsion and its persistence. *Nat Rev Neurosci, 2*(10), 695-703. doi: 10.1038/35094560
Jentsch, J. D., & Taylor, J. R. (1999). Impulsivity resulting from frontostriatal dysfunction in drug abuse: implications for the control of behavior by reward-related stimuli. *Psychopharmacology (Berl), 146*(4), 373-390.
Kalivas, P. W. (2004). Glutamate systems in cocaine addiction. *Curr Opin Pharmacol, 4*(1), 23-29. doi: 10.1016/j.coph.2003.11.002
Koob, G. F., & Volkow, N. D. (2010). Neurocircuitry of addiction. *Neuropsychopharmacology, 35*(1), 217-238. doi: 10.1038/npp.2009.110
Krieglmeyer, R., & Deutsch, R. (2010). Comparing measures of approach-avoidance behaviour: The manikin task vs. two versions of the joystick task. *Cogn Emot, 24*(5), 810-828. doi: Pii 917897230
Doi 10.1080/02699930903047298
Krieglmeyer, R., Deutsch, R., De Houwer, J., & De Raedt, R. (2010). Being Moved: Valence Activates Approach-Avoidance Behavior Independently of Evaluation and Approach-Avoidance Intentions. *Psychol Sci, 21*(4), 607-613. doi: Doi 10.1177/0956797610365131
Lakoff, G., & Johnson, M. (1980a). *Metaphors we live by.* Chicago: University of Chicago Press.
Lakoff, G. & Johnson, M. (1980b). Conceptual metaphor in everyday language. *Journal of Philosophy,* 77 (8), 453-486.
Lavender, T., & Hommel, B. (2007). Affect and action: Towards an event-coding account. *Cogn Emot, 21*(6), 1270-1296. doi: Doi 10.1080/02699930701438152

Loeber, S., Grosshans, M., Korucuoglu, O., Vollmert, C., Vollstadt-Klein, S., Schneider, S., . . . Kiefer, F. (2011). Impairment of inhibitory control in response to food-associated cues and attentional bias of obese participants and normal-weight controls. *Int J Obes (Lond)*. doi: 10.1038/ijo.2011.184

MacLeod, C., Rutherford, E., Campbell, L., Ebsworthy, G., & Holker, L. (2002). Selective attention and emotional vulnerability: assessing the causal basis of their association through the experimental manipulation of attentional bias. *J Abnorm Psychol, 111*(1), 107-123.

Markman, A. B., & Brendl, C. M. (2005). Constraining theories of embodied cognition. *Psychol Sci, 16*(1), 6-10. doi: 10.1111/j.0956-7976.2005.00772.x

Marr, D., (1983). *Vision: A computational investigation into the human representation and processing of visual information*. New York: W. H. Freeman and Company.

Mogg, K., Bradley, B. P., Field, M., & De Houwer, J. (2003). Eye movements to smoking-related pictures in smokers: relationship between attentional biases and implicit and explicit measures of stimulus valence. *Addiction, 98*(6), 825-836.

Mogg, K., Field, M., & Bradley, B. P. (2005). Attentional and approach biases for smoking cues in smokers: an investigation of competing theoretical views of addiction. *Psychopharmacology (Berl), 180*(2), 333-341. doi: 10.1007/s00213-005-2158-x

Noë, A. (2005). *Action in perception.* Cambridge, MA: The MIT Press.

Noë, A. (2009). *Out of our heads: Why you are not your brain, and other lessons from the biology of consciousness*. Hill and Wang.

O'Regan, J.K., & Noë, A. (2001). A sensorimotor account of vision and visual consciousness. *Behavioral and Brain Sciences,* 24 (5), 883-917.

Park, S. Q., Kahnt, T., Beck, A., Cohen, M. X., Dolan, R. J., Wrase, J., & Heinz, A. (2010). Prefrontal cortex fails to learn from reward prediction errors in alcohol dependence. *J Neurosci, 30*(22), 7749-7753. doi: 10.1523/JNEUROSCI.5587-09.2010

Paulus, M. P., & Stewart, J. L. (2014). Interoception and drug addiction. *Neuropharmacology, 76 Pt B*, 342-350. doi: 10.1016/j.neuropharm.2013.07.002

Peeters, M., Wiers, R. W., Monshouwer, K., van de Schoot, R., Janssen, T., & Vollebergh, W. A. (2012). Automatic processes in at-risk adolescents: the role of alcohol-approach tendencies and response inhibition in drinking behavior. *Addiction, 107*(11), 1939-1946. doi: 10.1111/j.1360-0443.2012.03948.x

Phaf, R. H., Mohr, S. E., Rotteveel, M., & Wicherts, J. M. (2014). Approach, avoidance, and affect: a meta-analysis of approach-avoidance tendencies in manual reaction time tasks. *Front Psychol, 5*, 378. doi: 10.3389/fpsyg.2014.00378

Prinz, J.J., & Barsalou, L.W. (2000). Steering a course for embodied representation. In E. Dietrich & A. Markman (Eds.), *Cognitive dynamics: Conceptual change in humans and machines,* pp. 51-77. NJ: Lawrence Erlbaum.

Pulvermüller, F. (1999). Words in the brain's language. *Behavioral and Brain Sciences,* 22 (2), 253-279.

Pulvermuller, F. (2008). Grounding language in the brain. In M. de Vega, A. Glenberg & A. Graesser (Eds.), *Symbols and embodiment: Debates on meaning and cognition*. Oxford: Oxford University Press.

Pylyshyn, Z. (1980). Computation and cognition: Issues in the foundation of cognitive science. *The Behavioral and Brain Sciences*, 3, 111–132.

Pylyshyn, Z. (1984). *Computation and cognition: Toward a foundation for cognitive science*. Cambridge, MA: Bradford Books/MIT Press.

Radke, S., Guths, F., Andre, J. A., Muller, B. W., & de Bruijn, E. R. (2014). In action or inaction? Social approach-avoidance tendencies in major depression. *Psychiatry Res*. doi: 10.1016/j.psychres.2014.07.011

Rinck, M., & Becker, E. S. (2007). Approach and avoidance in fear of spiders. *J Behav Ther Exp Psychiatry, 38*(2), 105-120. doi: 10.1016/j.jbtep.2006.10.001

Robbins, T. W., & Everitt, B. J. (1999). Drug addiction: bad habits add up. *Nature, 398*(6728), 567-570. doi: 10.1038/19208

Robinson, T. E., & Berridge, K. C. (1993). The neural basis of drug craving: an incentive-sensitization theory of addiction. *Brain Res Brain Res Rev, 18*(3), 247-291.

Robinson, T. E., & Berridge, K. C. (2003). Addiction. *Annu Rev Psychol, 54*, 25-53. doi: 10.1146/annurev.psych.54.101601.145237

Roelofs, K., Minelli, A., Mars, R. B., van Peer, J., & Toni, I. (2009). On the neural control of social emotional behavior. *Soc Cogn Affect Neurosci, 4*(1), 50-58. doi: 10.1093/scan/nsn036

Schoenmakers, T., Wiers, R. W., Jones, B. T., Bruce, G., & Jansen, A. T. (2007). Attentional re-training decreases attentional bias in heavy drinkers without generalization. *Addiction, 102*(3), 399-405. doi: 10.1111/j.1360-0443.2006.01718.x

Schoenmakers, T. M., de Bruin, M., Lux, I. F., Goertz, A. G., Van Kerkhof, D. H., & Wiers, R. W. (2010). Clinical effectiveness of attentional bias modification training in abstinent alcoholic patients. *Drug Alcohol Depend, 109*(1-3), 30-36. doi: 10.1016/j.drugalcdep.2009.11.022

Siegel, S. (1999). Drug anticipation and drug addiction. The 1998 H. David Archibald Lecture. *Addiction, 94*(8), 1113-1124.

Solomon, S. B., Wang, D., Sun, J., Kanias, T., Feng, J., Helms, C. C., . . . Natanson, C. Mortality increases after massive exchange transfusion with older stored blood in canines with experimental pneumonia. *Blood, 121*(9), 1663-1672. doi: blood-2012-10-462945 [pii] 10.1182/blood-2012-10-462945

Turing, A. (1936). On computable numbers. *Proceedings of the London Mathematical Society*, 24, 230–265.

Tversky, B (2009). Spatial cognition: Embodied and situated. In M. Aydede & P. Robbins (Eds.), *The Cambridge handbook of situated cognition*, pp. 201-217. Cambridge: Cambridge University Press.

Varela, F., Thompson, E., & Rosch, E. (1991). *The embodied mind: Cognitive science and human experience.* Cambridge, MA: MIT Press.

Stanley, D. A., Sokol-Hessner, P., Banaji, M. R., & Phelps, E. A. (2011). Implicit race attitudes predict trustworthiness judgments and economic trust decisions. *Proc Natl Acad Sci U S A, 108*(19), 7710-7715. doi: 10.1073/pnas.1014345108
Strack, F., Martin, L. L., & Stepper, S. (1988). Inhibiting and facilitating conditions of the human smile: a nonobtrusive test of the facial feedback hypothesis. *J Pers Soc Psychol, 54*(5), 768-777.
Tiffany, S. T. (1990). A cognitive model of drug urges and drug-use behavior: role of automatic and nonautomatic processes. *Psychol Rev, 97*(2), 147-168.
Veenstra, E. M., & de Jong, P. J. (2011). Reduced automatic motivational orientation towards food in restricting anorexia nervosa. *J Abnorm Psychol, 120*(3), 708-718. doi: 10.1037/a0023926
Volkow, N. D., Wang, G. J., Fowler, J. S., Tomasi, D., Telang, F., & Baler, R. (2010). Addiction: decreased reward sensitivity and increased expectation sensitivity conspire to overwhelm the brain's control circuit. *Bioessays, 32*(9), 748-755. doi: 10.1002/bies.201000042
Volman, I., Toni, I., Verhagen, L., & Roelofs, K. (2011). Endogenous testosterone modulates prefrontal-amygdala connectivity during social emotional behavior. *Cereb Cortex, 21*(10), 2282-2290. doi: 10.1093/cercor/bhr001
Watson, P., de Wit, S., Cousijn, J., Hommel, B., & Wiers, R. W. (2013). Motivational mechanisms underlying the approach bias to cigarettes. *Experimental Psychopathology*.
Watson, P., de Wit, S., Hommel, B., & Wiers, R. W. (2012). Motivational Mechanisms and Outcome Expectancies Underlying the Approach Bias toward Addictive Substances. *Front Psychol, 3*, 440. doi: 10.3389/fpsyg.2012.00440
Wiers, C. E. (2014). *The automatic pull of drug cues in addiction: Neural correlates and effects of re-training.* (Doctoral thesis), Freie Universität Berlin, Berlin.
Wiers, C. E., Kuhn, S., Javadi, A. H., Korucuoglu, O., Wiers, R. W., Walter, H., . . . Bermpohl, F. (2013). Automatic approach bias towards smoking cues is present in smokers but not in ex-smokers. *Psychopharmacology (Berl), 229*(1), 187-197. doi: 10.1007/s00213-013-3098-5
Wiers, C. E., Stelzel, C., Park, S. Q., Gawron, C. K., Ludwig, V. U., Gutwinski, S., . . . Bermpohl, F. (2014). Neural correlates of alcohol-approach bias in alcohol addiction: the spirit is willing but the flesh is weak for spirits. *Neuropsychopharmacology, 39*(3), 688-697. doi: 10.1038/npp.2013.252
Wiers, R. W., Bartholow, B. D., van den Wildenberg, E., Thush, C., Engels, R. C., Sher, K. J., . . . Stacy, A. W. (2007). Automatic and controlled processes and the development of addictive behaviors in adolescents: a review and a model. *Pharmacol Biochem Behav, 86*(2), 263-283. doi: 10.1016/j.pbb.2006.09.021
Wiers, R. W., Eberl, C., Rinck, M., Becker, E. S., & Lindenmeyer, J. (2011). Retraining automatic action tendencies changes alcoholic patients' approach bias for alcohol and improves treatment outcome. *Psychol Sci, 22*(4), 490-497. doi: 10.1177/0956797611400615
Wiers, R. W., Gladwin, T. E., Hofmann, W., Salemink, E., & Ridderinkhof, K. R. (2013). Cognitive Bias Modification and Cognitive Control Training in Addiction and

Related Psychopathology: Mechanisms, Clinical Perspectives, and Ways Forward. *Clinical Psychological Science, 1*(2), 192-212

Wiers, R. W., Gladwin, T. E., & Rinck, M. (2013). Should we train alcohol-dependent patients to avoid alcohol? *Front Psychiatry, 4*, 33. doi: 10.3389/fpsyt.2013.00033

Wiers, R. W., Rinck, M., Dictus, M., & van den Wildenberg, E. (2009). Relatively strong automatic appetitive action-tendencies in male carriers of the OPRM1 G-allele. *Genes Brain Behav, 8*(1), 101-106. doi: 10.1111/j.1601-183X.2008.00454.x

Wiers, R. W., Rinck, M., Kordts, R., Houben, K., & Strack, F. (2010). Retraining automatic action-tendencies to approach alcohol in hazardous drinkers. *Addiction, 105*(2), 279-287. doi: 10.1111/j.1360-0443.2009.02775.x

Zhou, Y., Li, X., Zhang, M., Zhang, F., Zhu, C., & Shen, M. (2012). Behavioural approach tendencies to heroin-related stimuli in abstinent heroin abusers. *Psychopharmacology (Berl), 221*(1), 171-176. doi: 10.1007/s00213-011-2557-0