



# Motivational mechanisms and outcome expectancies underlying the approach bias toward addictive substances

P. Watson<sup>1,2</sup>, S. de Wit<sup>2,3</sup>, Bernhard Hommel<sup>4,5</sup> \* and R. W. Wiers<sup>1,2</sup>

<sup>1</sup> ADAPT Lab, Department of Developmental Psychology, University of Amsterdam, Amsterdam, Netherlands

<sup>2</sup> Cognitive Science Center Amsterdam, University of Amsterdam, Amsterdam, Netherlands

<sup>3</sup> Department of Clinical Psychology, University of Amsterdam, Amsterdam, Netherlands

<sup>4</sup> Cognitive Psychology Unit, Leiden University, Leiden, Netherlands

<sup>5</sup> Leiden Institute for Brain and Cognition, Leiden, Netherlands

## Edited by:

Dorit Wenke, Humboldt University at Berlin, Germany

## Reviewed by:

Giovanni Pezzulo, National Research

Council of Italy, Italy

Andreas B. Eder, University of

Würzburg, Germany

## \*Correspondence:

Bernhard Hommel, Cognitive Psychology Unit, Department of Psychology, Leiden University, Room 2B05, Wassenaarseweg 52, 2333 AK Leiden, Netherlands.  
e-mail: hommel@fsw.leidenuniv.nl

Human behavior can be paradoxical, in that actions can be initiated that are seemingly incongruent with an individual's explicit desires. This is most commonly observed in drug addiction, where maladaptive behavior (i.e., drug seeking) appears to be compulsive, continuing at great personal cost. Approach biases toward addictive substances have been correlated with actual drug-use in a number of studies, suggesting that this measure can, in some cases, index everyday maladaptive tendencies. At present it is unclear whether this bias to drug cues is a Pavlovian conditioned approach response, a habitual response, the result of a Pavlovian-instrumental transfer process, or a goal-directed action in the sense that expectancy of the rewarding effects of drugs controls approach. We consider this question by combining the theoretical framework of associative learning with the available evidence from approach bias research. Although research investigating the relative contributions of these mechanisms to the approach bias is to date relatively limited, we review existing studies and also outline avenues for future research.

**Keywords:** approach, dual-process theory, addiction, associative learning, motivation, goal-directed action, habit, Pavlovian-instrumental transfer

## INTRODUCTION

To what extent is human behavior under voluntary control? Drug addiction is an extreme example, where drug seeking continues despite negative social or interpersonal consequences. Although many drug users are fully aware of the negative consequences and seek treatment in order to abstain from drug use, risk of relapse remains high. This highlights the paradoxical, destructive characteristic of addiction: that drug-seeking behavior persists despite explicit motivations to the contrary. Understanding the cognitive and motivational mechanisms that maintain such behaviors may allow us to better understand action control in general.

The approach bias is a behavioral inclination to approach rather than avoid certain stimuli. Experimental research into the approach bias has provided evidence for correlations with actual drug use and it is theorized that an approach bias may contribute to problematic drug-related behavior (Stacy and Wiers, 2010). An important question that remains to be addressed, however, is what the approach bias represents and how it relates to other features of drug use such as craving. Specifically, it is not clear whether the approach bias has the characteristics of being a goal-directed behavior, controlled by the expectancy of a rewarding outcome. Alternatively, it may better fit the profile of a Pavlovian conditioned response, or of a persistent, habitual response to drug cues, or it may be driven by Pavlovian-instrumental interactions.

In the present paper, we provide a comprehensive overview of these possible mechanisms that may facilitate the approach bias toward addictive substances. To this end, we relate the approach bias to theories of addiction as well as theoretical concepts of

associative learning theory (based on fundamental animal as well as human behavioral research). Furthermore, we critically review the experimental measurement of approach bias and the evidence that it can underlie maladaptive behaviors. The existing literature does not allow firm conclusions to be drawn about the relative contributions of different mechanisms, although we hope that this manuscript will inspire empirical investigations of this issue. To further stimulate such investigations, we will outline several possible avenues for future research at the end of this article.

## THEORETICAL BACKGROUND

Addiction has been described as a progressive neurological disorder of learning and memory whereby everyday associative learning processes become pathological (Hyman, 2005; Koob and Volkow, 2010). Teenagers experimenting with alcohol, for example, may discover that it makes them feel sociable and lively. The initial learning of associations between rewarding outcomes and the contexts or behaviors that lead to them, allows for the emergence of goal-directed behavior (e.g., approaching the bar at a party *in order* to feel sociable). For many teenagers, this goal-directed behavior can over time become habitual, such that simply being in a party context is the impetus to move toward the bar, regardless of any consideration of possible (pleasant) outcomes. For some individuals it may ultimately also persist when this behavior has undesirable consequences. Those individuals that continue to consume alcohol despite pervasive negative consequences and sometimes even explicit intentions to abstain may be regarded as *compulsive* drug-users (Koob and Volkow, 2010).

A number of learning and reward processes underlie this behavioral transition from voluntary drug use to clinically recognized drug dependency whereby the consumption is maintained at increasingly high cost. Associative learning theories, arising from the systematic study of animal behavior and neurobiology, offer a coherent framework for defining and dissociating these processes (see Balleine et al., 2008; Balleine and O'Doherty, 2009). Initial empirical investigations in humans support the case for applying the associative framework to human behavior (review: de Wit and Dickinson, 2009). The psychological constructs arising from associative learning theory are also paralleled in some neurocomputational models of decision-making, and we refer the interested reader to recent publications on this topic for an in-depth discussion (Daw et al., 2005; Dayan et al., 2006; Balleine et al., 2008). We shall discuss approach behavior within the associative learning context, because this context provides the most useful concepts to understand and investigate approach and avoidance behavior. At the same time, however, we emphasize that a purely non-propositional, associative framework can be argued to be insufficient as an exhaustive account of human decision-making (see for example Mitchell et al., 2009, but see also associated commentaries).

### PAVLOVIAN OR CLASSICAL CONDITIONING

Associations between environmental cues (e.g., beer at a party) and motivationally relevant events (e.g., feeling lively and sociable) are acquired over time and may play an important role in guiding everyday decision-making. The conditions under which such associations emerge have been studied by means of Pavlovian (or classical) conditioning methods that establish a predictive relationship between these (Pavlov, 1927). Conditioned stimuli (CS) are cues that were once irrelevant (e.g., a food bowl) and that through repetitive pairing with a motivationally relevant stimulus (e.g., food) come to elicit conditioned responses. These conditioned responses can be consummatory in nature (e.g., salivation) or preparatory (approach toward the food bowl; Konorski, 1967). Furthermore, it has been shown in humans that Pavlovian conditioning can lead to acquired likes and dislikes of previously neutral objects and places (evaluative conditioning; Hermans et al., 2002; Hofmann et al., 2010). These processes may play an important role in drug-seeking behavior as contexts that were previously paired with the rewarding experience of drug taking become preferred and will elicit conditioned approach responses that may support drug seeking.

### PAVLOVIAN VERSUS INSTRUMENTAL LEARNING

While conditioned approach may well contribute to the approach bias, there may also be an instrumental component. Whereas Pavlovian behavior results from the contingencies between stimuli and motivationally relevant events, instrumental behavior arises from the contingency between a response and a motivationally relevant outcome (Skinner, 1938). If the outcome is rewarding, the instrumental agent will acquire the responses that lead to this outcome.

Embedded within any instrumental contingency is also a Pavlovian relationship between the context and the outcome. For example, where a light stimulus (S) may signal that a lever

press response (R) will produce a food pellet outcome (O) there is an inherent, parallel S–O relationship between the light and the rewarding food pellet being conditioned. Often a conditioned Pavlovian response will facilitate instrumental behavior (e.g., salivation or approach to the food bowl will facilitate eating behavior). There are occasions, however, when contradicting instrumental and Pavlovian responses (whether preparatory or consummatory) can cause conflict (Sheffield, 1965; Hershberger, 1986). Hershberger (1986), for example, created a “looking glass world” in which a food bowl receded with twice the speed at which hungry chicks ran toward it, and drew near at twice the speed at which the chicks ran away from it. To gain access to the food the chicks had to learn to overcome the Pavlovian bias to approach the food bowl, which acted as a CS for the food it contained. Most chicks continued to run toward the bowl, however, and thereby lost the available food. The approach bias of these chicks was clearly controlled predominantly by Pavlovian conditioning, as sensitivity to the R–O contingency should have allowed them to learn to make the opposite response of running away from the food bowl. On the other hand, it is well-known that animals, as well as humans, are capable of instrumental behavior. In a later section we will review the evidence for an instrumental component of the approach bias in humans as measured in the laboratory.

### FROM GOAL-DIRECTED ACTIONS TO HABITUAL RESPONSES

Instrumental conditioning could contribute to the approach bias by giving rise to either goal-directed approach or to habitual approach that is triggered directly by environmental stimuli. Goal-directed actions are performed in order to achieve desirable outcomes (e.g., approaching the bar at a party to feel more lively and sociable) and are thus flexibly modulated by the incentive value of the outcome (Adams and Dickinson, 1981). Over time, however, these appetitive outcomes gradually reinforce S–R associations, that give rise to habitual responding that is directly evoked by the context. In this scenario, the party context triggers approach, rather than consideration of the drinking outcome. Overtraining of an instrumental action is one way to bring about habitual responding (Adams, 1982; Dickinson, 1985; Tricomi et al., 2009). In the early stages of drug use, drug-seeking behavior appears to meet the criteria for goal-direction action. Habitual drug seeking triggered by certain cues and contexts may, however, help to maintain drug-seeking behavior, even when the drug is no longer desired.

In animal studies, the degree to which behavior is goal-directed or habitual is formally assessed by means of the outcome devaluation procedure. In this procedure instrumental training is followed by devaluation of the instrumental outcome (e.g., through satiation on a particular food reward). Subsequently, an extinction test is conducted to assess instrumental responding for the devalued outcome. If behavior is predominantly under goal-directed control, responding for the devalued outcome should be immediately reduced. In contrast to goal-directed behavior, S–R habits are not sensitive to devaluation of outcomes and such behavior will persist.

### PAVLOVIAN-TO-INSTRUMENTAL TRANSFER

A popular beer brand logo can prompt thoughts of beer drinking, which may increase the probability that an individual will

head to the nearest bar and realize that outcome. This anticipatory effect is formally described as Pavlovian-to-instrumental transfer (PIT). In lab demonstrations of this phenomenon, a common outcome (such as a food reward) functions as both a Pavlovian reward and also an instrumental reward, in separate training phases. This training allows for the separate development of both Pavlovian S–O expectancies and instrumental O–R associations. The interaction effect is then assessed by presenting the Pavlovian cue whilst the subject is given the opportunity to perform the instrumental response for that outcome. Many studies have shown that although the Pavlovian cue was never directly paired with a response, the expectancy elicited by the cue increases the likelihood of instrumental responding for that specific reward (outcome-specific PIT) or in some cases boosts responding generally (general PIT; Estes, 1948; Rescorla and Solomon, 1967; Corbit and Balleine, 2005). PIT effects are now well documented in humans (Hogarth et al., 2007; Bray et al., 2008; Talmi et al., 2008; Hogarth and Chase, 2011; Huys et al., 2011; Nadler et al., 2011; Hogarth, 2012) and both forms of transfer (specific and general) could play a role in instrumental approach behavior. For example, the sight of a beer brand logo may remind one of beer drinking which may activate approach behavior that is previously been instrumental in obtaining beer, via S–O–R associations. General PIT, on the other hand, can only further strengthen a pre-existing bias. For example, if there is already an approach bias toward alcohol, then any reward-associated cue (such as a cigarette logo for smokers) may further increase that bias by boosting the dominant approach response. Intriguingly, in animals, PIT effects have been shown to be insensitive to outcome devaluation (Rescorla, 1994; Holland, 2004) suggesting that these could play an important role in addiction relapse. Two recent studies in smokers provided evidence that Pavlovian cues predicting smoking outcomes increase the likelihood of responding for cigarettes and that furthermore, this can occur regardless of the current incentive value of the smoking outcome (Hogarth and Chase, 2011; Hogarth, 2012). In these studies, Pavlovian cues increased responding for cigarettes even after participants had read health warnings about cigarettes (Hogarth and Chase, 2011) or been treated with nicotine replacement therapy (Hogarth, 2012) and, crucially, had decreased responding in the absence of the cues. It seems paradoxical that this behavior, controlled by the anticipation of the outcome it produces, is not modulated by the current incentive value of that outcome, but this effect has been convincingly demonstrated in animals and humans (Rescorla, 1994; Holland, 2004; Hogarth and Chase, 2011; Hogarth, 2012; but see Allman et al., 2010). It appears, therefore, that in outcome-specific Pavlovian-instrumental interactions, the representation of the outcome contains sensory, but not motivationally relevant information (see Delamater and Oakeshott, 2007). The result is that the perceptual characteristics of the outcome prompt the associated response, regardless of the current incentive value of that outcome. Future research should elucidate the exact mechanism that mediates outcome-specific PIT, but on the basis of these outcome-reevaluation studies, we will make a distinction in the remainder of this manuscript between goal-directed action and outcome-specific PIT, with the latter also being mediated by

anticipation of the outcome but occurring independently of the incentive goal status of the outcome.

## THEORIES OF ADDICTION

The motivation driving destructive drug-seeking behavior is a key component of all major theories of addiction and most provide some explanation for why environmental cues can trigger relapse, even after long periods of sobriety. Some theories make clear predictions about the mechanisms that could facilitate an approach bias toward drug cues and whether such a bias is goal-directed, stimulus-bound habits, or due to PIT anticipatory processes. Relevant to the discussion at hand are incentive sensitization, theories based on the role of expectancy, various dual-process models (including habit theories of addiction), and negative reinforcement models.

### INCENTIVE SENSITIZATION

The incentive sensitization model proposes that repeated drug use causes neuroadaptations in mesolimbic dopaminergic systems controlling the incentive values assigned to drug stimuli (Robinson and Berridge, 1993; Berridge, 2007). Over time a pathological incentive value becomes attributed to drug cues and contexts prompting compulsive drug-taking. This incentive sensitization increases even whilst levels of subjective pleasure decrease over the course of addiction (defined as increased “wanting” even in the absence of “liking” a drug). Thus whilst drugs can become disliked and an individual may have explicit motivations to avoid them, cues remain extremely salient and continue to elicit craving and motivate approach behavior (Robinson and Berridge, 1993, 2000, 2001).

### EXPECTANCY THEORIES

Some models propose that expectancies of drug outcomes play a crucial role in motivating drug-seeking behavior. Goldman and colleagues argue that drug use is a goal-directed choice, based on the expectation of the hedonic effect of the drug outcome (Goldman et al., 1987; Goldman, 2002). Following a meta-analysis of conditioning studies using tobacco rewards, Hogarth and Duka (2006) found evidence for the role of expectancies in drug-seeking behaviors. Recently, however, this view was extended, given demonstrations that cue-elicited anticipation of a cigarette reward prompted responding for that reward, even when incentive value was low. The authors suggested therefore that parallel goal-directed expectancies and PIT anticipatory processes jointly determine action control (Hogarth and Chase, 2011; Hogarth, 2012).

### DUAL-PROCESS MODELS

There are various dual-process models of addiction. Similarly to Tiffany’s habit theory of addiction (Tiffany, 1990; Tiffany and Conklin, 2000), associative theories make a distinction between goal-directed and stimulus-bound behaviors (Everitt et al., 2001; Everitt and Robbins, 2005; de Wit and Dickinson, 2009; Hogarth and Chase, 2011; Hogarth, 2012). Another group of dual-process models describe an automatic, appetitive system opposed by an

executive control system (Strack and Deutsch, 2004; Wiers et al., 2007; Hofmann et al., 2008, 2009; Stacy and Wiers, 2010). These will be discussed in turn.

### **Habitual versus goal-directed control in dual-process models**

Tiffany's habit theory of addiction (Tiffany, 1990; Tiffany and Conklin, 2000) proposes that over time, drug-taking "rituals" become automatic behavioral schema, prompted by the environment. Whilst there are many unique features within Tiffany's model, this transition from goal-directed to habitual behavior is also captured by associative dual-process models of addiction (Everitt et al., 2001; Everitt and Robbins, 2005). According to this view, the reinforcing effects of drugs lead to strong S–R associations between contextual stimuli and drug-seeking behaviors. Over time, approach behavior toward drugs becomes a habitual response, triggered by environmental cues. This behavioral transition appears to be paralleled by impaired functioning of cortico-striatal networks supporting goal-directed behavior (Porriño et al., 2004; Everitt and Robbins, 2005). In a slightly different vein, Hogarth and Chase (2011) argue that goal-directed and PIT anticipatory processes operate in an additive manner, jointly determining behavior.

### **Implicit versus explicit control in dual-process models**

Another set of dual-process models are centered upon the notion that appetitive behavior can be automatically triggered by a variety of cues and that this behavior needs to be regulated by executive control processes, maintaining goal focus, and motivation to resist use and abuse of drugs (Strack and Deutsch, 2004; Wiers et al., 2007; Hofmann et al., 2008, 2009; Stacy and Wiers, 2010). Individual differences in impulsivity and cognitive control modulate the effectiveness of this regulation (Dawe et al., 2004; Hofmann et al., 2009; Peeters et al., 2012). The impulsive and executive dual processes can be mapped fairly well onto the stimulus-bound and goal-directed distinction of associative models, although it should be noted that the reflective control system is often argued to be propositional in nature, not associative (Strack and Deutsch, 2004).

### **Negative reinforcement theories**

Alleviation of a negative affective state – either withdrawal symptoms or more generally, depression or stress – is a commonly cited cause of relapse (Carey and Correia, 1997; Shiffman and Waters, 2004; Kuntsche et al., 2005). Negative reinforcement theories highlight the role of internal cues (negative affective states) in prompting drug use (Koob and Le Moal, 1997; Baker et al., 2004; Eissenberg, 2004; Ahmed and Koob, 2005). Koob and colleagues propose that addiction is the result of dysfunction in not only the reward system but also the anti-reward system, driving aversive states (Ahmed and Koob, 2005; Koob and Le Moal, 2005). It is beyond the scope of the present paper to provide a detailed overview of negative reinforcement theories, but for the purposes of the current discussion we would like to note that similar motivational mechanisms may underlie drug-seeking based on the rewarding properties of drugs versus avoidance of aversive states (Baker et al., 2004; Eissenberg, 2004). Avoidance behavior in the context of drug use could be either a goal-directed strategy based on expectancies of the alleviating outcome, or a stimulus-response

habit reinforced by alleviation of negative states, or the result of Pavlovian-instrumental interactions.

### **INTERIM CONCLUSIONS**

Assessing the characteristics of the motivational mechanisms underlying addictive approach behavior, should provide evidence in favor of different models of addiction. These models overlap with respect to several common predictions. They all provide an explanation for why relapse can be triggered by environmental cues, whether this is due to S–R associations, incentive sensitization, or triggering of goal-directed expectancies. There are nonetheless a number of subtle distinctions. Many implicit/explicit dual-process models propose that the approach bias represents an automatic, positive evaluation of drug cues, which is argued to be distinct from explicit processes. Goal-directed expectancy theories on the other hand would argue that the bias arises due to positive expectancy of the drug outcome. Some unique predictions derive from these models, which can be empirically tested. For example, associative dual-process models predict that approach behavior will eventually be resistant to outcome devaluation as behavior transitions to habitual control. This is in stark contrast to goal-directed expectancy theories, which predict that decreases in outcome value will continue to reduce responding for drug outcomes.

### **EMPIRICAL DATA**

A longstanding idea in psychological science is that a considerable amount of behavior is driven by rapid, evolutionary relevant, affective evaluations of stimuli. These affective evaluations classify all stimuli as either "negative" or "positive," facilitating in the latter case, approach behavior (Bindra, 1974; Dickinson and Dearing, 1979; Chen and Bargh, 1999; Fazio, 2001; Krieglmeier et al., 2010).

Motivations and affective attitudes are commonly assessed via questionnaires. Unfortunately, however, conclusions from such explicit measures can be difficult as participants may lack insight into the driving forces behind their actions and choices. In the case of addiction, invalid self-reports may result from self-presentational strategies or self-deception in an attempt to maintain a positive self-image. Moreover, introspection has been argued to not be a reliable and objective method of assessing motivational states (Berridge et al., 2009; Schooler and Mauss, 2009; Wood and Neal, 2009; Neal et al., 2012). As behavior becomes more habitual over time, there may not be a corresponding shift in subjective awareness of this fact. In the case of drug relapse, *post hoc* evaluation of one's behavior may lead an individual to conclude that their behavior was motivated by a craving for the drug as opposed to being prompted by the external environment.

To overcome these difficulties and problems, a number of indirect, speeded reaction-time tasks have been developed to assess the valence and strength of affective evaluative associations (and the resulting approach behavior) without the need for explicit reflection on the part of the subject (Fazio, 2001; De Houwer, 2006; De Houwer et al., 2009). We focus here on measures of action tendencies, although it should be noted that varieties of the Implicit Association Task have also been used to assess approach and avoidance associations of a target category such as alcohol (Palfai and Ostafin, 2003; Ostafin and Palfai, 2006).

## MEASUREMENTS OF APPROACH BIAS

In order to directly assess approach tendencies, a number of tasks have been developed that measure speed of approach toward (generally pictorial) stimuli. Approach bias is generally measured as the difference in reaction time on trials where participants make an approach movement (such as pulling a joystick) versus an avoid movement (pushing a joystick) to the pictorial stimuli on a computer screen. A number of tasks have been developed, and whilst they all measure approach bias, they are confusingly and interchangeably labeled as either stimulus-response compatibility (SRC) tasks (utilizing either a manikin or a joystick), approach avoidance tasks (AATs), or affective Simon tasks. To avoid confusion, we will use the explicit paradigm labels when describing these tasks in later sections.

### THE MANIKIN TASK

The manikin task provides an indirect measure of approach and avoidance behaviors. Approach tendencies are assessed by calculating the difference in reaction times across two blocks of experimental trials. In the first block the participant moves a computerized manikin toward one category of stimuli (e.g., alcohol) and away from other stimuli (e.g., soft drinks). In a subsequent block this assignment reverses. Using this task, participants have been seen to approach positive words faster than they are avoided, with the reverse effect for negative words (De Houwer et al., 2001). In addition, the manikin task has been used to assess approach behaviors in studies focusing on eating disorders (Woud et al., 2011), obesity (Havermans et al., 2011), Pavlovian conditioning of neutral stimuli (Thewissen et al., 2007; van Gucht et al., 2008), as well as approach tendencies toward alcohol (Field et al., 2005b, 2008; Schoenmakers et al., 2008; van Hemel-Ruiter et al., 2011; Barkby et al., 2012), cigarettes (Mogg et al., 2003, 2005; Bradley et al., 2004; Field et al., 2005a), and cannabis (Field et al., 2006; Cousijn et al., 2012).

### THE JOYSTICK TASK

The joystick task can be used to measure differences in reaction times when the participant pushes a joystick away from his/her body in response to stimuli as opposed to pulling the joystick toward his/her body on a subsequent block. This task has been used to study phobias and anxiety (see Roefs et al., 2011), lifestyle and fitness goals (Fishbach and Shah, 2006), and food deprivation manipulations (Seibt et al., 2007).

Whilst it was originally suggested that approach and avoidance movements are represented as stored motor patterns, triggered by automatic, affective stimuli evaluations, it has become increasingly clear that motor actions *per se* do not represent either approach or avoidance. The same motor response (such as arm flexion) may represent approach in one situation (moving something toward oneself) but avoidance in another situation (quickly moving hand away from a stimulus to be avoided; Chen and Bargh, 1999). Indeed many studies have now shown that it is an individual's interpretation of the result of the behavior that is important (i.e., is the stimulus moved closer or further away) and as such, neutral body movements can be interpreted as approach and avoidance actions depending on the outcome (Lavender and Hommel, 2007; Seibt et al., 2008; van Dantzig et al., 2008; Krieglmeier et al., 2010). The

zooming joystick task (ZJT) is thus a disambiguated version of the joystick task, designed to avoid misinterpretation and recategorization of pushing and pulling movements. The introduction of a zooming feature ensures that participants experience the illusion of stimuli moving away from them and coming toward them when they push or pull the joystick (this is achieved by reducing or enlarging the size of the picture). This zooming feature reduces the possibility of participants interpreting pulling movements as avoid rather than approach. Using the zooming version of the task (negative) approach tendencies to spiders have been assessed in spider phobia (Rinck and Becker, 2007).

By asking participants to respond to an irrelevant task feature such as orientation or location of the picture on screen instead of the content of the picture, the task may be rendered more implicit (De Houwer, 2003). Using the ZJT with *irrelevant feature instructions*, approach bias has been examined in heavy drinkers (Wiers et al., 2009, 2010), alcoholic patients (Wiers et al., 2011), at-risk adolescents (Peeters et al., 2012), and heavy cannabis users (Cousijn et al., 2011).

### RELIABILITY AND VALIDITY

Studies correlating approach bias measures to real-life behavior, as opposed to self-report measurements, are limited in number (due in part to the complexity of such designs). Rinck and Becker (2007) found that an approach bias on the ZJT predicted actual approach behaviors to live spiders, over and above that which was predicted with spider-phobia questionnaires. More importantly for the present discussion, approach bias on an irrelevant feature version of the ZJT was correlated with the amount of alcohol drunk in what was described to participants as an unrelated consumer "taste test" following the task (Wiers et al., 2010).

The split-half reliability of the manikin and joystick tasks is variable but generally good when using task-relevant feature instructions (Rinck and Becker, 2007; Krieglmeier and Deutsch, 2010; Field et al., 2011). The advantage of instructing participants to respond on the basis of a task irrelevant feature is that it makes the task less susceptible to explicit control on the part of the participant, the practical drawback is that compatibility effects tend to be smaller (Krieglmeier and Deutsch, 2010; Field et al., 2011). This reduced effect may be due to the fact that attention is not drawn to the affective properties of the stimuli in the irrelevant feature version. Several studies have, nonetheless, demonstrated robust approach (or avoidance) biases using irrelevant feature instructions (De Houwer et al., 2001; Rinck and Becker, 2007; Seibt et al., 2007; Wiers et al., 2009, 2010; Veenstra and de Jong, 2010; Cousijn et al., 2011). The reliability of the irrelevant feature instruction version was found to be poor in one study (Krieglmeier and Deutsch, 2010) whilst another reported reasonably good reliability (Cousijn et al., 2011).

There are, evidentially, pros and cons to the various approach/avoidance task versions. Two studies, conducted with both the manikin and the standard joystick tasks, unexpectedly failed to find evidence for a correlation between the two approach bias scores (Krieglmeier and Deutsch, 2010; van Hemel-Ruiter et al., 2011). The reasons for this are not immediately clear although there are major differences in how approach and avoidance are conceptualized within the tasks. Both the standard

joystick task (without the zooming feature) and the manikin task are susceptible to recategorization – the manikin can be recategorized as someone other than the self and the approach/avoid movements in the standard joystick task are relatively ambiguous. Future studies should carefully select the task paradigm, depending on the research question, a point that we will return to in a later section “Outstanding Questions and Future Directions.”

### APPROACH BIASES IN ADDICTION

Using these aforementioned approach/avoidance paradigms, addiction researchers have provided substantial evidence for a relationship between drug-approach bias and drug use. That is, although the approach bias is measured experimentally in a lab, with superficial key press or lever movements, the behavioral tendency to be faster at approaching rather than avoiding drug stimuli, does seem to confer information about drug behavior more generally. Approach tendencies have been demonstrated in heavy (non-clinical) users of alcohol (Schoenmakers et al., 2008; Wiers et al., 2009, 2010), social drinkers (Field et al., 2005b), and cigarette smokers (Field et al., 2005a; Thewissen et al., 2007). Cigarette smokers have been seen to show a greater approach bias than non-smokers (Mogg et al., 2003; Bradley et al., 2004), as do cannabis users versus non-users (Field et al., 2006; Cousijn et al., 2011). Hazardous (non-clinical) drinkers were seen to have a stronger approach bias compared to light drinkers (Field et al., 2008). These results suggest a reliable relationship between drug use and drug-approach bias, particularly when examining healthy participants with moderate levels of dependence.

It should be noted that patterns of results can differ depending on the populations studied. Whilst lighter drinkers showed a weaker approach bias (Field et al., 2008), this pattern was reversed in one study investigating light versus heavy cigarette smokers (Mogg et al., 2005). In addition, in contrast to the aforementioned alcohol studies (with students), three studies involving patients receiving treatment for alcoholism did not find stronger approach tendencies for alcohol pictures compared to soft drink pictures (Wiers et al., 2011; Barkby et al., 2012; Spruyt et al., in press). These studies are small in number, but differing patterns of results in different populations at different stages of addiction can likely tell us something about the role of explicit motivations in approach bias measurements. We will discuss this in further detail, in the later section “Outstanding Questions and Future Directions.”

### APPROACH TENDENCIES IN ADDICTION AND THE UNDERLYING MOTIVATIONS

The question remains as to which of the cognitive and motivational mechanisms outlined earlier in the “Theoretical Background” section contribute to the approach bias as measured in the lab. Evidence for the contribution of a Pavlovian component to experimental measures of the approach bias comes from studies that have shown that CS-reward learning quickly engenders an approach bias toward these novel CSs. Using different variations of AAT, approach bias has been conditioned toward novel Pavlovian stimuli predicting cigarette outcomes (Thewissen et al., 2007) and chocolate outcomes (van Gucht et al., 2008). A direct association between these stimuli and an instrumental approach response cannot mediate the approach bias to novel CSs,

indicating therefore that Pavlovian mechanisms do play a role in the approach bias.

There is evidence, however, suggesting that the approach bias cannot be completely reduced to a purely Pavlovian conditioned response. As discussed previously, Hershberger (1986) showed that under conditions where chicks needed to make a withdrawal response in order to make a food bowl move toward them, they were unable to suppress the urge to approach the food bowl. This behavioral inflexibility provides evidence that the approach behavior of the chicks was predominantly controlled by a Pavlovian mechanism. In contrast to chicks, however, humans are perfectly well able to adapt their approach behavior. To our knowledge, the human equivalent of Hershberger’s experiment has not been conducted yet, but a recent study did employ a similar design. In a manikin task, participants were required to make an initial brief avoid movement in order to approach positive words and an initial brief approach movement to avoid negative words. Krieglmeyer et al. (2011) showed that even when the initial movement is avoidance, participants will still react faster if the final outcome is that the manikin approaches positive words. The reverse was true for avoiding negative words such that even if the initial movement was to approach a negative word, participants reacted faster if the final outcome was avoidance. This study suggests that the approach bias is more complex than being a mere Pavlovian approach response, as the final outcome (is the stimulus further away or closer) and not the initial direction of movement (toward or away from the stimuli), seems to influence reaction times.

Retraining studies with the joystick task provide further evidence for instrumental control over approach behavior (Kawakami et al., 2007, 2008; Wiers et al., 2010, 2011). For example, Wiers et al. (2010) presented the vast majority of a set of alcohol pictures in the push rather than pull format and found that retraining reduced the approach bias toward these pictures. The observation that participants can modify the bias following training (avoiding appetitive pictures) suggests that the bias is more than a conditioned response and shows a degree of flexibility that is in line with an instrumental account of the approach bias.

The results discussed above suggest that the approach response is not a purely Pavlovian response, although it is challenging to disentangle the relative contributions of Pavlovian and instrumental mechanisms using these paradigms. The question remains, nonetheless, as to whether the approach bias is flexibly modulated by changes in incentive value of the outcome or whether it is merely triggered by the drug stimuli. A number of studies have found that approach bias measurements increase in line with self-report craving scores, a result that is generally interpreted to suggest that approach behavior is sensitive to the current incentive value of the outcome (Field et al., 2005b, 2008; van Gucht et al., 2008). However, this correlative finding should be interpreted with caution as it does not necessarily imply a causal relationship between craving and the approach bias. Furthermore, other studies failed to find evidence for a relationship between craving and the approach bias (Mogg et al., 2003; Thewissen et al., 2007; Wiers et al., 2010; Cousijn et al., 2011). However, none of those studies really address whether behavior is *immediately* sensitive to a change in the incentive outcome value. So far, such outcome-reevaluation designs

have yielded mixed results. Two studies, using the manikin task, manipulated craving by giving participants a placebo drink in one session and a dose of alcohol in another. Approach bias scores to alcohol pictures (Schoenmakers et al., 2008) and smoking pictures (Field et al., 2005a) were then compared between the two sessions (alcohol or placebo). Both studies found that self-reported craving was higher in the alcohol session but there was no difference in the approach bias scores – a null effect that although difficult to interpret, is more in line with the habitual account. These results are in contrast to a study using the standard joystick task, that examined the effects of satiety on the approach bias (Seibt et al., 2007). Participant's responses to images of food were measured either before or after lunch and satiety did appear to reduce the bias in the non-deprived group, suggesting that approach behavior was driven by the current desire for food. Unfortunately, however, this study failed to include a neutral control picture condition, and we can therefore not ascertain whether hunger increased the approach movement toward food pictures specifically, or approach behavior generally. Still, these results suggest that this line of research should be extended further to critically assess motivational modulation of approach.

As discussed previously, some dual-process theories suggest that the approach bias results from an interaction between associative learning processes and explicit cognitive control processes. Barkby et al. (2012) provided correlational evidence for the importance of cognitive control, by testing patients receiving treatment for alcohol addiction on the manikin task. Their critical finding was that approach bias scores on the manikin task correlated with individual differences in explicit approach/avoidance intentions. Further evidence, that behavioral intentions can influence approach behavior, comes from a study using a variant of the ZJT (Sharbanee et al., in press). Rather than calculating approach bias as a difference score between the push and pull reaction times, this study made a distinction between “pull alcohol picture” trials and “push alcohol picture” trials – the former trial type assumed to be congruent with an appetitive tendency and the latter incongruent. Only incongruent trials, therefore, should demand recruitment of executive control processes (to overcome the appetitive tendency and push the alcohol picture away). As expected, results showed that working memory scores modulated reaction times in problem drinkers attempting (unsuccessfully) to control their alcohol consumption, but this effect was only observed on incongruent “push alcohol” trials. This suggests that the approach bias arises due to a complex interaction between the strength of the approach tendency and the ability to inhibit this tendency when required.

We should point out, however, that many other studies suggest that approach tendencies are not always under intentional control. These studies report seemingly “automatic” approach biases that are not in line with instrumental withdrawal intentions: participants scoring higher on a restrained eating scale showed a greater approach bias toward food cues (Veenstra and de Jong, 2010); smokers showed an approach bias toward smoking cues that they reported as unpleasant (Bradley et al., 2008); and appetitive Pavlovian stimuli inhibited instrumental withdrawal in situations where the instrumental withdrawal behavior was rewarded with money (Huys et al., 2011). It appears, therefore, that explicit intentions can sometimes influence approach, but a complete account

of the approach bias will also have to encapsulate the role of associative learning processes.

To summarize, the evidence surveyed suggests that both Pavlovian and instrumental mechanisms play a role in facilitating the approach bias, but it is not yet clear how these processes interact or sum to produce this behavioral tendency. Furthermore it remains to be seen whether the approach bias is flexibly modulated by outcome value or has the characteristics of a habitual response to drug cues. Recent research examining instrumental responding for cigarette outcomes has argued that goal-directed and PIT processes operate in parallel, summing in an additive manner (Hogarth and Chase, 2011; Hogarth, 2012). However, the role of PIT in the approach bias remains to be empirically addressed. Furthermore, next to associative mechanisms, behavioral intentions may also modulate the approach bias, with one study suggesting that the approach bias measures some combination of both appetitive and regulatory control processes (Sharbanee et al., in press).

### OUTSTANDING QUESTIONS AND FUTURE DIRECTIONS

Approach bias tasks offer a fast and simple manner of measuring approach tendencies to drug-related stimuli and appear to tell us something about drug use, given that a number of studies have correlated approach bias scores with actual drug use (Field et al., 2008; Wiers et al., 2010; Cousijn et al., 2011) and shown group differences between heavier versus lighter/non-users (Mogg et al., 2003; Bradley et al., 2004; Field et al., 2006; Cousijn et al., 2011). Whilst understanding the mechanisms that underlie the approach bias is an important theoretical question, it should be noted that these tasks are not ideally suited to dissociating the various motivation mechanisms introduced in the earlier “Theoretical Background” section. To isolate goal-directed approach from Pavlovian approach for example, requires a task where the instrumental actions are bidirectional (i.e., left and right). In such a task, the relationships between the stimulus and the outcomes are held constant whilst the relationships between the *direction* of action and the outcomes are manipulated. The relative contribution of Pavlovian processes is equal to both actions, and hence controlled for (Dickinson et al., 1996). In addition, it has been observed that outcome devaluation modulates both conditioned Pavlovian responses and goal-directed instrumental responses (Colwill and Rescorla, 1988), and as such we cannot differentiate between these two in an outcome-reevaluation study if the approach bias is the dependent measure. Nonetheless, whether the approach bias is flexibly modulated by outcome reevaluation or is directly triggered by the drug stimuli is an outstanding question. The studies that have employed outcome-reevaluation paradigms have yielded mixed results (Field et al., 2005a; Seibt et al., 2007; Schoenmakers et al., 2008) and this line of research within the context of addictive substances, should be continued.

Specifically, outcome-reevaluation studies conducted with individuals at different levels of dependency, could address the question of whether the approach bias becomes more habitual over the course of addiction. Given the observation that users receiving clinical treatment may not show a very strong approach bias (Wiers et al., 2011; Barkby et al., 2012; Spruyt et al., in press), this method could also be used to assess whether control over approach behavior is regained during (successful) treatment. The work of

Hogarth and colleagues in the field of smoking addiction suggests that there are a number of ways to reevaluate addictive substances, namely health warnings, temporary satiety through consumption, and treatments aimed at alleviating withdrawal symptoms (Hogarth and Chase, 2011; Hogarth, 2012). Another way may be to pair the consumption of an appetitive substance with an aversive flavor (Howard, 2001; van Gucht et al., 2010). In order to conduct an outcome-reevaluation test, an approach bias measurement would first be taken with neutral and category of interest (e.g., smoking) pictures. Then the smoking outcome would be devalued (e.g., through satiety) and the approach bias measurement would be repeated. If the approach bias measurement is not reduced following outcome devaluation, this would suggest that the approach bias is a stimulus-bound response to drug stimuli. Different versions of the approach bias tasks, as reviewed earlier, may be better suited to reevaluation studies given that ideally a repeated measures design is employed and that the second measurement, following outcome devaluation, should be conducted in extinction (without presentation of the outcome). The standard joystick or manikin tasks therefore, with irrelevant feature version, would be preferable in such a paradigm – as participant awareness of the study aims should be reduced as much as possible.

Pavlovian-instrumental interactions are thought to play an important role in addictive behavior (Hogarth and Chase, 2011; Hogarth, 2012). However, whether PIT processes *can* confer a specific or general motivating effect on approach/avoidance movements on these tasks has not been assessed. This could be investigated using, for instance, the manikin task. In an initial instrumental (O–R) learning phase, participants would make an approach movement to earn one specific outcome and avoidance to earn another (e.g., approach is rewarded with beer; avoidance rewarded with wine). This would be followed by Pavlovian (S–O) training where participants would learn the predictive relationship between neutral stimuli and these same outcomes. During the subsequent transfer test, occasional stimulus presentations would be expected to facilitate/speed up the response associated with a common outcome. For example, a stimulus for beer would be expected to facilitate approach, while a stimulus for wine should facilitate avoidance. General PIT, on the other hand, could be assessed by comparing the influence of a stimulus associated with a third outcome (e.g., whiskey) versus one associated with no alcoholic drink.

fMRI research of the approach bias is scarce, but given the wealth of knowledge that exists concerning the neural correlates of the various motivational mechanisms highlighted in this review (see Balleine and O'Doherty, 2009 for a detailed overview), this could be a very fruitful avenue for investigating many of the questions raised thus far. Approach bias to cannabis using a manikin task was recently investigated in an fMRI study with heavy cannabis users (Cousijn et al., 2012). Results suggested that ventromedial prefrontal cortex (vmPFC) was recruited during congruent “approach cannabis” blocks as opposed to incongruent “avoid cannabis” blocks. The vmPFC/orbitofrontal cortex, along with the caudate, have been consistently implicated in goal-directed action (Valentin et al., 2007; de Wit et al., 2009, 2012b). Similar prefrontal regions are argued to encode Pavlovian outcome values (Gottfried et al., 2002, 2003). The results of Cousijn and colleagues suggest,

therefore, that the approach bias is driven by mechanisms flexibly modulated by outcome value, as opposed to habits – the latter being mediated not by prefrontal regions, but instead networks involving the posterior putamen and premotor cortex (Tricomi et al., 2009; Ashby et al., 2010; de Wit et al., 2012b). We should point out however that another fMRI study reported contradictory findings – namely that vmPFC was recruited on incongruent and not congruent trials (Roelofs et al., 2009). Important methodological differences could account for these differential results – Roelofs and colleagues used a standard joystick task with affective, non-drug-related facial stimuli (happy/unhappy faces) and the effects disappeared when participants were instructed to approach/avoid on the basis of gender rather than facial expression (irrelevant feature version). Future studies should hopefully be able to resolve these findings.

Research examining the neural correlates of approach tendencies would be best suited to approach bias paradigms with high internal reliability – such as the relevant feature version of the manikin task. A question of interest is whether networks implicated in goal-directed control versus habitual control are recruited, and whether this is different between groups who are at different stages of addiction. In humans, differential regions of the amygdala are thought to mediate general and outcome-specific PIT (Prévost et al., 2012) and their engagement in approach bias tasks could also be examined. fMRI can also be used to assess the role of brain regions such as the anterior cingulate, known to be important in overcoming response conflict and cognitive control more generally (reviews: Botvinick et al., 2004; Ridderinkhof et al., 2004), during approach bias tasks. Although the fMRI studies mentioned earlier examined contrasts of either congruent > incongruent or incongruent > congruent – reporting the results of both contrasts would be beneficial to test hypotheses relating to the interaction between appetitive responses and explicit cognitive control.

Pharmacological manipulations could also be employed to investigate the effect of neurotransmitter depletion on approach tendencies, with dopamine being an obvious candidate. Females submitted to dopamine precursor depletion, for example, were seen to rely more on habitual S–R knowledge at the expense of goal-directed O–R knowledge in a task designed to assess the relative balance of these two systems (de Wit et al., 2012a). A GABA antagonist (Baclofen), also thought to have effects via mediation of dopaminergic systems, reduces craving and consumption in alcohol, and cigarette addiction (Franklin et al., 2009; Gorsane et al., 2012) yet the effect on the approach bias has not been studied. Studies of this type would help with attempts to understand what, exactly, the approach bias represents and how it relates to other measures such as craving.

Finally, the extent to which these manipulations selectively affect approach versus avoidance (rather than the composite approach bias score) is worth investigation. Some studies have started to tease apart the relative contributions of “congruent” approach responses to appetitive stimuli as opposed to “incongruent” avoidance movements away from appetitive stimuli, with interesting insights (Roelofs et al., 2009; Barkby et al., 2012; Cousijn et al., 2012; Sharbanee et al., in press). By looking at these processes separately we can gain a better understanding of



what the approach bias actually measures and what role various cognitive and motivational mechanisms play in producing this effect.

## CONCLUSION

Human behavior can be paradoxical, in that actions are initiated that are seemingly incongruent with an individual's explicit motivations. This is most commonly observed in addiction, where maladaptive behavior (i.e., drug seeking) appears to be compulsive. Different theoretical approaches attempt to explain this behavior in different ways, with some suggesting that positive expectancies drive such behavior and others arguing that environmental stimuli can trigger behaviors incongruent with current goals and behavioral intentions. A number of studies have observed correlations between problematic drug use and approach bias scores (Field et al., 2008; Cousijn et al., 2011) suggesting that approach bias measurements can index everyday behaviors. Understanding the cognitive and motivational mechanisms that drive such an approach bias may therefore provide insight into both adaptive and maladaptive action control.

Determining the mechanisms underlying approach may have clinical implications. Cognitive therapy may be useful if expectancies and cognitive control are important determinants of approach. On the other hand, alternative approaches such as exposure response prevention therapy or counter-conditioning, that target the behavior directly, may be more appropriate if approach is stimulus-bound (e.g., van Gucht et al., 2010). Retraining the approach bias using the ZJT has been shown to be effective in

reducing approach behavior toward alcohol cues (Wiers et al., 2010), and in a clinical population such retraining of the approach bias leads to a significantly smaller relapse rate at one-year follow-up compared to individuals receiving sham training (Wiers et al., 2011). A better understanding of the motivational mechanisms that underlie the approach bias, will also provide a better understanding of what exactly is being trained in this novel paradigm and how it can be better improved as a viable treatment.

Integration of the literature on approach bias, motivation, and associative learning provides a clear framework with which to identify and disentangle the relative contributions of various cognitive and motivational mechanisms underlying such maladaptive behavior. Whilst the literature surveyed suggests that both Pavlovian and instrumental mechanisms contribute to the approach bias, it remains to be elucidated exactly how they interact and sum to produce approach behavior. Hopefully further research assessing these questions will be forthcoming, within the limits that are inherent to such a task paradigm. Understanding the mechanisms that underlie an approach bias will provide a better understanding of the complex interplay of automatic processes, outcome expectancies, and behavioral intentions underlying human action control. This is not only theoretically important but ultimately has implications for clinical treatment.

## ACKNOWLEDGMENTS

This PhD studentship is made possible through funding received from the Netherlands Organization for Scientific Research (NWO; 433-09-243).

## REFERENCES

- Adams, C. D. (1982). Variations in the sensitivity of instrumental responding to reinforcer devaluation. *Q. J. Exp. Psychol. Sect. B Comp. Physiol. Psychol.* 34, 77–98.
- Adams, C. D., and Dickinson, A. (1981). Instrumental responding following reinforcer devaluation. *Q. J. Exp. Psychol. Sect. B Comp. Physiol. Psychol.* 33, 109–121.
- Ahmed, S. H., and Koob, G. F. (2005). Transition to drug addiction: a negative reinforcement model based on an allostatic decrease in reward function. *Psychopharmacology (Berl.)* 180, 473–490.
- Allman, M. J., DeLeon, I. G., Cataldo, M. F., Holland, P. C., and Johnson, A. W. (2010). Learning processes affecting human decision making: an assessment of reinforcer-selective Pavlovian-to-instrumental transfer following reinforcer devaluation. *J. Exp. Psychol. Anim. Behav. Process.* 36, 402–408.
- Ashby, F. G., Turner, B. O., and Horvitz, J. C. (2010). Cortical and basal ganglia contributions to habit learning and automaticity. *Trends Cogn. Sci. (Regul. Ed.)* 14, 208–215.
- Baker, T. B., Piper, M. E., McCarthy, D. E., Majeskie, M. R., and Fiore, M. C. (2004). Addiction motivation reformulated: an affective processing model of negative reinforcement. *Psychol. Rev.* 111, 33–51.
- Balleine, B. W., Daw, N. D., and O'Doherty, J. P. (2008). "Multiple forms of value learning and the function of dopamine," in *Neuroeconomics: Decision Making and the Brain*, ed. P. W. Glimcher (Waltham, MA: Academic Press), 367–385.
- Balleine, B. W., and O'Doherty, J. P. (2009). Human and rodent homologies in action control: corticostriatal determinants of goal-directed and habitual action. *Neuropsychopharmacology* 35, 48–69.
- Barkby, H., Dickson, J. M., Roper, L., and Field, M. (2012). To approach or avoid alcohol? Automatic and self-reported motivational tendencies in alcohol dependence. *Alcohol. Clin. Exp. Res.* 36, 361–368.
- Berridge, K. (2007). The debate over dopamine's role in reward: the case for incentive salience. *Psychopharmacology (Berl.)* 191, 391–431.
- Berridge, K. C., Robinson, T. E., and Aldridge, J. W. (2009). Dissecting components of reward: "liking," "wanting," and learning. *Curr. Opin. Pharmacol.* 9, 65–73.
- Bindra, D. (1974). A motivational view of learning, performance, and behavior modification. *Psychol. Rev.* 81, 199–213.
- Botvinick, M. M., Cohen, J. D., and Carter, C. S. (2004). Conflict monitoring and anterior cingulate cortex: an update. *Trends Cogn. Sci.* 8, 539–546.
- Bradley, B. P., Field, M., Healy, H., and Mogg, K. (2008). Do the affective properties of smoking-related cues influence attentional and approach biases in cigarette smokers? *J. Psychopharmacol.* 22, 737–745.
- Bradley, B., Field, M., Mogg, K., and De Houwer, J. (2004). Attentional and evaluative biases for smoking cues in nicotine dependence: component processes of biases in visual orienting. *Behav. Pharmacol.* 15, 29.
- Bray, S., Rangel, A., Shimojo, S., Balleine, B., and O'Doherty, J. P. (2008). The neural mechanisms underlying the influence of Pavlovian cues on human decision making. *J. Neurosci.* 28, 5861–5866.
- Carey, K. B., and Correia, C. J. (1997). Drinking motives predict alcohol-related problems in college students. *J. Stud. Alcohol* 58, 100–105.
- Chen, M., and Bargh, J. A. (1999). Consequences of automatic evaluation: immediate behavioral predispositions to approach or avoid the stimulus. *Pers. Soc. Psychol. Bull.* 25, 215–224.
- Colwill, R. M., and Rescorla, R. A. (1988). Associations between the discriminative stimulus and the reinforcer in instrumental learning. *J. Exp. Psychol. Anim. Behav. Process.* 14, 155.
- Corbit, L. H., and Balleine, B. W. (2005). Double dissociation of basolateral and central amygdala lesions on the general and outcome-specific forms of Pavlovian-instrumental transfer. *J. Neurosci.* 25, 962–970.
- Cousijn, J., Goudriaan, A. E., Ridderinkhof, K. R., van den Brink, W., Veltman, D. J., and Wiers, R. W. (2012). Approach-bias predicts development of cannabis problem severity in heavy cannabis users: results from a prospective fMRI study. *PLoS ONE* 7, e42394. doi:10.1371/journal.pone.0042394
- Cousijn, J., Goudriaan, A. E., and Wiers, R. W. (2011). Reaching out towards cannabis: approach-bias in heavy cannabis users predicts changes in cannabis use. *Addiction* 106, 1667–1674.
- Daw, N. D., Niv, Y., and Dayan, P. (2005). Uncertainty-based competition between prefrontal and dorsolateral striatal systems for behavioral control. *Nat. Neurosci.* 8, 1704–1711.

- Dawe, S., Gullo, M. J., and Loxton, N. J. (2004). Reward drive and rash impulsiveness as dimensions of impulsivity: implications for substance misuse. *Addict. Behav.* 29, 1389–1405.
- Dayan, P., Niv, Y., Seymour, B., and Daw, N. D. (2006). The misbehavior of will and the discipline of the will. *Neural Netw.* 19, 1153–1160.
- De Houwer, J. (2003). “A structural analysis of indirect measures of attitudes,” in *The Psychology of Evaluation: Affective Processes in Cognition and Emotion*, eds J. Musch and K. C. Klauer (Mahwah, NJ: Lawrence Erlbaum), 219–244.
- De Houwer, J. (2006). “What are implicit measures and why are we using them,” in *The Handbook of Implicit Cognition and Addiction*, eds R. W. Wiers and A. W. Stacy (Thousand Oaks, CA: Sage Publishers), 11–28.
- De Houwer, J., Crombez, G., Baeyens, F., and Hermans, D. (2001). On the generality of the affective Simon effect. *Cogn. Emot.* 15, 189–206.
- De Houwer, J., Teige-Mocigemba, S., Spruyt, A., and Moors, A. (2009). Implicit measures: a normative analysis and review. *Psychol. Bull.* 135, 347.
- de Wit, S., Corlett, P. R., Aitken, M. R., Dickinson, A., and Fletcher, P. C. (2009). Differential engagement of the ventromedial prefrontal cortex by goal-directed and habitual behavior toward food pictures in humans. *J. Neurosci.* 29, 11330–11338.
- de Wit, S., and Dickinson, A. (2009). Associative theories of goal-directed behaviour: a case for animal–human translational models. *Psychol. Res.* 73, 463–476.
- de Wit, S., Standing, H. R., Devito, E. E., Robinson, O. J., Ridderinkhof, K. R., Robbins, T. W., et al. (2012a). Reliance on habits at the expense of goal-directed control following dopamine precursor depletion. *Psychopharmacology (Berl.)* 219, 621–631.
- de Wit, S., Watson, P., Harsay, H. A., Cohen, M. X., van de Vijver, I., and Ridderinkhof, K. R. (2012b). Corticostriatal connectivity underlies individual differences in the balance between habitual and goal-directed action control. *J. Neurosci.* 32, 12066–12075.
- Delamater, A. R., and Oakshott, S. (2007). Learning about multiple attributes of reward in Pavlovian conditioning. *Ann. N. Y. Acad. Sci.* 1104, 1–20.
- Dickinson, A. (1985). Actions and habits: the development of behavioural autonomy. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 308, 67–78.
- Dickinson, A., Campos, J., Varga, Z. I., and Balleine, B. (1996). Bidirectional instrumental conditioning. *Q. J. Exp. Psychol. Sect. B Comp. Physiol. Psychol.* 49, 289–306.
- Dickinson, A., and Dearing, M. F. (1979). “Appetitive-aversive interactions and inhibitory processes,” in *Mechanisms of Learning and Motivation*, eds A. Dickinson and R. A. Boakes (Hillsdale, NJ: Erlbaum), 203–231.
- Eissenberg, T. (2004). Measuring the emergence of tobacco dependence: the contribution of negative reinforcement models. *Addiction* 99(Suppl 1), 5–29.
- Estes, W. K. (1948). Discriminative conditioning. II. Effects of a Pavlovian conditioned stimulus upon a subsequently established operant response. *J. Exp. Psychol.* 38, 173–177.
- Everitt, B. J., Dickinson, A., and Robbins, T. W. (2001). The neuropsychological basis of addictive behaviour. *Brain Res. Rev.* 36, 129–138.
- Everitt, B. J., and Robbins, T. W. (2005). Neural systems of reinforcement for drug addiction: from actions to habits to compulsion. *Nat. Neurosci.* 8, 1481–1489.
- Fazio, R. H. (2001). On the automatic activation of associated evaluations: an overview. *Cogn. Emot.* 15, 115–141.
- Field, M., Caren, R., Fernie, G., and De Houwer, J. (2011). Alcohol approach tendencies in heavy drinkers: comparison of effects in a relevant stimulus-response compatibility task and an approach/avoidance Simon task. *Psychol. Addict. Behav.* 25, 697–701.
- Field, M., Eastwood, B., Bradley, B. P., and Mogg, K. (2006). Selective processing of cannabis cues in regular cannabis users. *Drug Alcohol Depend.* 85, 75–82.
- Field, M., Kiernan, A., Eastwood, B., and Child, R. (2008). Rapid approach responses to alcohol cues in heavy drinkers. *J. Behav. Ther. Exp. Psychiatry* 39, 209–218.
- Field, M., Mogg, K., and Bradley, B. P. (2005a). Alcohol increases cognitive biases for smoking cues in smokers. *Psychopharmacology (Berl.)* 180, 63–72.
- Field, M., Mogg, K., and Bradley, B. P. (2005b). Craving and cognitive biases for alcohol cues in social drinkers. *Alcohol* 40, 504–510.
- Fishbach, A., and Shah, J. Y. (2006). Self-control in action: implicit dispositions toward goals and away from temptations. *J. Pers. Soc. Psychol.* 90, 820–832.
- Franklin, T. R., Harper, D., Kampman, K., Kildea-McCrea, S., Jens, W., Lynch, K. G., et al. (2009). The GABA B agonist baclofen reduces cigarette consumption in a preliminary double-blind placebo-controlled smoking reduction study. *Drug Alcohol Depend.* 103, 30–36.
- Goldman, M. S. (2002). Expectancy and risk for alcoholism: the unfortunate exploitation of a fundamental characteristic of neurobehavioral adaptation. *Alcohol. Clin. Exp. Res.* 26, 737–746.
- Goldman, M. S., Brown, S. A., and Christiansen, B. A. (1987). “Expectancy theory: thinking about drinking,” in *Psychological Theories of Drinking and Alcoholism*, eds H. T. Blane and K. E. Leonard (New York: Guilford Press), 181–226.
- Gorsane, M. A., Kebir, O., Hache, G., Blecha, L., Aubin, H. J., Reynaud, M., et al. (2012). Is baclofen a revolutionary medication in alcohol addiction management? Review and recent updates. *Subst. Abuse.* 33, 336–349.
- Gottfried, J. A., O’Doherty, J., and Dolan, R. J. (2002). Appetitive and aversive olfactory learning in humans studied using event-related functional magnetic resonance imaging. *J. Neurosci.* 22, 10829–10837.
- Gottfried, J. A., O’Doherty, J., and Dolan, R. J. (2003). Encoding predictive reward value in human amygdala and orbitofrontal cortex. *Science* 301, 1104–1107.
- Havermans, R. C., Giesen, J. C. A. H., Houben, K., and Jansen, A. (2011). Weight, gender, and snack appeal. *Eat. Behav.* 12, 126–130.
- Hermans, D., Vansteenwegen, D., Crombez, G., Baeyens, F., and Eelen, P. (2002). Expectancy-learning and evaluative learning in human classical conditioning: affective priming as an indirect and nonobtrusive measure of conditioned stimulus valence. *Behav. Res. Ther.* 40, 217–234.
- Hershberger, W. (1986). An approach through the looking-glass. *Learn. Behav.* 14, 443–451.
- Hofmann, W., De Houwer, J., Perugini, M., Baeyens, F., and Crombez, G. (2010). Evaluative conditioning in humans: a meta-analysis. *Psychol. Bull.* 136, 390–421.
- Hofmann, W., Friese, M., and Strack, F. (2009). Impulse and self-control From a dual-systems perspective. *Perspect. Psychol. Sci.* 4, 162–176.
- Hofmann, W., Gschwendner, T., Friese, M., Wiers, R. W., and Schmitt, M. (2008). Working memory capacity and self-regulatory behavior: toward an individual differences perspective on behavior determination by automatic versus controlled processes. *J. Pers. Soc. Psychol.* 95, 962–977.
- Hogarth, L. (2012). Goal-directed and transfer-cue-elicited drug-seeking are dissociated by pharmacotherapy: evidence for independent additive controllers. *J. Exp. Psychol. Anim. Behav. Process.* 38, 266–278.
- Hogarth, L., and Chase, H. W. (2011). Parallel goal-directed and habitual control of human drug-seeking: implications for dependence vulnerability. *J. Exp. Psychol. Anim. Behav. Process.* 37, 261–276.
- Hogarth, L., Dickinson, A., Wright, A., Kouvaraki, M., and Duka, T. (2007). The role of drug expectancy in the control of human drug seeking. *J. Exp. Psychol. Anim. Behav. Process.* 33, 484–496.
- Hogarth, L., and Duka, T. (2006). Human nicotine conditioning requires explicit contingency knowledge: is addictive behaviour cognitively mediated? *Psychopharmacology (Berl.)* 184, 553–566.
- Holland, P. C. (2004). Relations between Pavlovian-instrumental transfer and reinforcer devaluation. *J. Exp. Psychol. Anim. Behav. Process.* 30, 104–117.
- Howard, M. O. (2001). Pharmacological aversion treatment of alcohol dependence. I. Production and prediction of conditioned alcohol aversion. *Am. J. Drug Alcohol Abuse* 27, 561–585.
- Huys, Q. J. M., Cools, R., Gölzer, M., Friedel, E., Heinz, A., Dolan, R. J., et al. (2011). Disentangling the roles of approach, activation and valence in instrumental and pavlovian responding. *PLoS Comput. Biol.* 7, e1002028. doi:10.1371/journal.pcbi.1002028
- Hyman, S. E. (2005). Addiction: a disease of learning and memory. *Am. J. Psychiatry* 162, 1414–1422.
- Kawakami, K., Phillips, C. E., Steele, J. R., and Dovidio, J. F. (2007). (Close) distance makes the heart grow fonder: improving implicit racial attitudes and interracial interactions through approach behaviors. *J. Pers. Soc. Psychol.* 92, 957–971.
- Kawakami, K., Steele, J. R., Cifa, C., Phillips, C. E., and Dovidio, J. F. (2008). Approaching math increases math = me and math = pleasant. *J. Exp. Soc. Psychol.* 44, 818–825.

- Konorski, J. (1967). *Integrative Activity of the Brain*. Chicago: University of Chicago Press.
- Koob, G. F., and Le Moal, M. (1997). Drug abuse: hedonic homeostatic dysregulation. *Science* 278, 52–58.
- Koob, G. F., and Le Moal, M. (2005). Plasticity of reward neurocircuitry and the “dark side” of drug addiction. *Nat. Neurosci.* 8, 1442–1444.
- Koob, G. F., and Volkow, N. D. (2010). Neurocircuitry of addiction. *Neuropharmacology* 35, 217–238.
- Krieglmeier, R., De Houwer, J., and Deutsch, R. (2011). How farsighted are behavioral tendencies of approach and avoidance? The effect of stimulus valence on immediate vs. ultimate distance change. *J. Exp. Soc. Psychol.* 47, 622–627.
- Krieglmeier, R., and Deutsch, R. (2010). Comparing measures of approach–avoidance behaviour: the manikin task vs. two versions of the joystick task. *Cogn. Emot.* 24, 810–828.
- Krieglmeier, R., Deutsch, R., De Houwer, J., and De Raedt, R. (2010). Being moved: valence activates approach-avoidance behavior independently of evaluation and approach-avoidance intentions. *Psychol. Sci.* 21, 607–613.
- Kuntsche, E., Knibbe, R., Gmel, G., and Engels, R. (2005). Why do young people drink? A review of drinking motives. *Clin. Psychol. Rev.* 25, 841–861.
- Lavender, T., and Hommel, B. (2007). Affect and action: towards an event-coding account. *Cogn. Emot.* 21, 1270–1296.
- Mitchell, C. J., De Houwer, J., and Lovibond, P. F. (2009). The propositional nature of human associative learning. *Behav. Brain Sci.* 32, 183–198; discussion 198–246.
- Mogg, K., Bradley, B. P., Field, M., and 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, 825–836.
- Mogg, K., Field, M., and Bradley, B. P. (2005). Attentional and approach biases for smoking cues in smokers: an investigation of competing theoretical views of addiction. *Psychopharmacology* 180, 333–341.
- Nadler, N., Delgado, M. R., and Delamater, A. R. (2011). Pavlovian to instrumental transfer of control in a human learning task. *Emotion* 11, 1112–1123.
- Neal, D. T., Wood, W., Labrecque, J. S., and Lally, P. (2012). How do habits guide behavior? Perceived and actual triggers of habits in daily life. *J. Exp. Soc. Psychol.* 48, 492–498.
- Ostafin, B. D., and Palfai, T. P. (2006). Compelled to consume: the implicit association test and automatic alcohol motivation. *Psychol. Addict. Behav.* 20, 322–327.
- Palfai, T. P., and Ostafin, B. D. (2003). Alcohol-related motivational tendencies in hazardous drinkers: assessing implicit response tendencies using the modified-IAT. *Behav. Res. Ther.* 41, 1149–1162.
- Pavlov, I. P. (1927). *Conditioned Reflexes*. London: Oxford University Press.
- Peeters, M., Wiers, R. W., Monshouwer, K., van de Schoot, R., Janssen, T., and Vollebergh, W. A. M. (2012). Automatic processes in at-risk adolescents: the role of alcohol-approach tendencies and response inhibition in drinking behavior. *Addiction* 107, 1939–1946.
- Porrino, L. J., Lyons, D., Smith, H. R., Daunais, J. B., and Nader, M. A. (2004). Cocaine self-administration produces a progressive involvement of limbic, association, and sensorimotor striatal domains. *J. Neurosci.* 24, 3554–3562.
- Prévost, C., Liljeholm, M., Tyszka, J. M., and O’Doherty, J. P. (2012). Neural correlates of specific and general Pavlovian-to-instrumental transfer within human amygdalar subregions: a high-resolution fMRI study. *J. Neurosci.* 32, 8383–8390.
- Rescorla, R. (1994). Transfer of instrumental control mediated by a devalued outcome. *Learn. Behav.* 22, 27–33.
- Rescorla, R. A., and Solomon, R. L. (1967). Two-process learning theory: relationships between Pavlovian conditioning and instrumental learning. *Psychol. Rev.* 74, 151–182.
- Ridderinkhof, K. R., Ullsperger, M., Crone, E. A., and Nieuwenhuis, S. (2004). The role of the medial frontal cortex in cognitive control. *Science* 306, 443–447.
- Rinck, M., and Becker, E. S. (2007). Approach and avoidance in fear of spiders. *J. Behav. Ther. Exp. Psychiatry* 38, 105–120.
- Robinson, T. E., and Berridge, K. C. (1993). The neural basis of drug craving: an incentive-sensitization theory of addiction. *Brain Res. Rev.* 18, 247–291.
- Robinson, T. E., and Berridge, K. C. (2000). The psychology and neurobiology of addiction: an incentive-sensitization view. *Addiction* 95(Suppl 2), S91–117.
- Robinson, T. E., and Berridge, K. C. (2001). Incentive-sensitization and addiction. *Addiction* 96, 103–114.
- Roefs, A., Huijding, J., Smulders, F. T. Y., MacLeod, C. M., de Jong, P. J., Wiers, R. W., et al. (2011). Implicit measures of association in psychopathology research. *Psychol. Bull.* 137, 149–193.
- Roelofs, K., Minelli, A., Mars, R. B., van Peer, J., and Toni, I. (2009). On the neural control of social emotional behavior. *Soc. Cogn. Affect. Neurosci.* 4, 50–58.
- Schoenmakers, T., Wiers, R. W., and Field, M. (2008). Effects of a low dose of alcohol on cognitive biases and craving in heavy drinkers. *Psychopharmacology* 197, 169–178.
- Schooler, J. W., and Mauss, I. B. (2009). “To be happy and to know it: the experience and meta-awareness of pleasure,” in *Pleasures of the Brain*, eds M. L. Kringelbach and K. C. Berridge (New York: Oxford University Press), 244–254.
- Seibt, B., Häfner, M., and Deutsch, R. (2007). Prepared to eat: how immediate affective and motivational responses to food cues are influenced by food deprivation. *Eur. J. Soc. Psychol.* 37, 359–379.
- Seibt, B., Neumann, R., Nussinson, R., and Strack, F. (2008). Movement direction or change in distance? Self- and object-related approach–avoidance motions. *J. Exp. Soc. Psychol.* 44, 713–720.
- Sharbanee, J. M., Stritzke, W. G., Wiers, R. W., Young, P., Rinck, M., and MacLeod, C. M. (in press). The interaction of approach-alcohol action tendencies and working memory capacity predicts the inability to regulate drinking behaviour. *Psychol. Addict. Behav.*
- Sheffield, F. D. (1965). “Relation between classical and instrumental conditioning,” in *Classical Conditioning*, ed. W. F. Prokasy (New York: Appleton-Century-Crofts), 302–322.
- Shiffman, S., and Waters, A. J. (2004). Negative affect and smoking lapses: a prospective analysis. *J. Consult. Clin. Psychol.* 72, 192–201.
- Skinner, B. F. (1938). *The Behavior of Organisms*. New York: Appleton-Century-Crofts.
- Spruyt, A., De Houwer, J., Tibboel, H., Verschuere, B., Crombez, G., Verbanck, P., et al. (in press). On the predictive validity of automatically activated approach/avoidance tendencies in abstaining alcohol-dependent patients. *Drug Alcohol Depend.*
- Stacy, A. W., and Wiers, R. W. (2010). Implicit cognition and addiction: a tool for explaining paradoxical behavior. *Annu. Rev. Clin. Psychol.* 6, 551–575.
- Strack, F., and Deutsch, R. (2004). Reflective and impulsive determinants of social behavior. *Pers. Soc. Psychol. Rev.* 8, 220–247.
- Talmi, D., Seymour, B., Dayan, P., and Dolan, R. J. (2008). Human Pavlovian-instrumental transfer. *J. Neurosci.* 28, 360–368.
- Thewissen, R., Havermans, R., Geschwind, N., van den Hout, M., and Jansen, A. (2007). Pavlovian conditioning of an approach bias in low-dependent smokers. *Psychopharmacology* 194, 33–39.
- Tiffany, S. T. (1990). A cognitive model of drug urges and drug-use behavior: role of automatic and nonautomatic processes. *Psychol. Rev.* 97, 147–168.
- Tiffany, S. T., and Conklin, C. A. (2000). A cognitive processing model of alcohol craving and compulsive alcohol use. *Addiction* 95, 145–153.
- Tricomi, E., Balleine, B. W., and O’Doherty, J. P. (2009). A specific role for posterior dorsolateral striatum in human habit learning. *Eur. J. Neurosci.* 29, 2225–2232.
- Valentin, V. V., Dickinson, A., and O’Doherty, J. P. (2007). Determining the neural substrates of goal-directed learning in the human brain. *J. Neurosci.* 27, 4019–4026.
- van Dantzig, S., Pecher, D., and Zwaan, R. A. (2008). Approach and avoidance as action effects. *Q. J. Exp. Psychol.* 61, 1298–1306.
- van Gucht, D., Baeyens, F., Vansteenwegen, D., Hermans, D., and Beckers, T. (2010). Counterconditioning reduces cue-induced craving and actual cue-elicited consumption. *Emotion* 10, 688–695.
- van Gucht, D., Vansteenwegen, D., Van den Bergh, O., and Beckers, T. (2008). Conditioned craving cues elicit an automatic approach tendency. *Behav. Res. Ther.* 46, 1160–1169.
- van Hemel-Ruiter, M. E., de Jong, P. J., and Wiers, R. W. (2011). Appetitive and regulatory processes in young adolescent drinkers. *Addict. Behav.* 36, 18–26.
- Veenstra, E. M., and de Jong, P. J. (2010). Restrained eaters show enhanced automatic approach tendencies

- towards food. *Appetite* 55, 30–36.
- Wiers, R. W., Bartholow, B. D., van Den Wildenberg, E., Thush, C., Engels, R. C. M. E., Sher, K. J., et al. (2007). Automatic and controlled processes and the development of addictive behaviors in adolescents: a review and a model. *Pharmacol. Biochem. Behav.* 86, 263–283.
- Wiers, R. W., Eberl, C., Rinck, M., Becker, E. S., and Lindenmeyer, J. (2011). Retraining automatic action tendencies changes alcoholic patients' approach bias for alcohol and improves treatment outcome. *Psychol. Sci.* 22, 490.
- Wiers, R. W., Rinck, M., Dictus, M., and Van den Wildenberg, E. (2009). Relatively strong automatic appetitive action-tendencies in male carriers of the OPRM1 G-allele. *Genes Brain. Behav.* 8, 101–106.
- Wiers, R. W., Rinck, M., Kordts, R., Houben, K., and Strack, F. (2010). Retraining automatic action-tendencies to approach alcohol in hazardous drinkers. *Addiction* 105, 279–287.
- Wood, W., and Neal, D. T. (2009). The habitual consumer. *J. Consum. Psychol.* 19, 579–592.
- Woud, M. L., Anschutz, D. J., Van Strien, T., and Becker, E. S. (2011). Measuring thin-spiration and fear of fat indirectly. A matter of approach and avoidance. *Appetite* 56, 451–455.

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

Received: 16 June 2012; accepted: 03 October 2012; published online: 22 October 2012.

Citation: Watson P, de Wit S, Hommel B and Wiers RW (2012) Motivational

mechanisms and outcome expectancies underlying the approach bias toward addictive substances. *Front. Psychology* 3:440. doi: 10.3389/fpsyg.2012.00440  
This article was submitted to *Frontiers in Cognition*, a specialty of *Frontiers in Psychology*.

Copyright © 2012 Watson, de Wit, Hommel and Wiers. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in other forums, provided the original authors and source are credited and subject to any copyright notices concerning any third-party graphics etc.