Beyond associations:

Do implicit beliefs play a role in smoking addiction?

Helen Tibboel

Jan De Houwer

Nicolas Dirix

Adriaan Spruyt

Department of Experimental-Clinical and Health Psychology, Ghent University, Belgium

Author for correspondence:

Helen Tibboel

Department of Psychology, Ghent University, Henri Dunantlaan 2, B-9000 Ghent, Belgium

Email: hftibboel@gmail.com

Phone: 0032 9 264 9143

Fax : 0032 9 264 64 89

Abstract

Background and objectives: Influential dual system models of addiction suggest that an automatic system that is associative and habitual promotes drug use, whereas a controlled system that is propositional and rational inhibits drug use. It is assumed that effects on the Implicit Association Test (IAT) reflect the automatic processes that guide drug seeking. However, results have been inconsistent, challenging 1) the validity of addiction IATs and 2) the assumption that the automatic system contains only simple associative information. We aimed to further test the validity of IATs that are used within this field of research using an experimental design. Second, we introduced a new procedure aimed at examining the automatic activation of complex propositional knowledge, the Relational Responding Task (RRT) and examine the validity of RRT effects in the context of smoking. *Methods:* In two experiments, smokers performed two different tasks: an approach/avoid IAT and a liking IAT in Experiment 1, and a smoking urges RRT and a valence IAT in Experiment 2. Smokers were tested once immediately after smoking, and once after 10 hours of nicotine-deprivation. *Results:* None of the IAT scores were affected by the deprivation manipulation. RRT scores revealed a stronger implicit desire for smoking in the deprivation condition compared to the satiation condition. *Conclusions:* IATs that are currently used to assess automatic processes in addiction have serious drawbacks. Furthermore, the automatic system may contain not only associations but complex drug-related beliefs as well. The RRT may be a useful and valid tool to examine these beliefs.

Keywords: Implicit association test, automatic beliefs, relational responding task, smoking, addiction

1. Introduction

Dual system models of addiction (e.g., Wiers et al., 2007) suggest that there are two different systems that control addictive behaviors: an automatic (or impulsive) system and a controlled (or reflective) system. The automatic system is fast, associative, and irrational and promotes drug use in addicts (e.g., the automatic activation of the association “cigarette” – “pleasant” triggers smoking). It is assumed that automatic processes play an important role in the development and maintenance of addiction (e.g., Robinson & Berridge, 2003; Wiers, Rinck, Kordts, Houben, & Strack, 2010). Furthermore, a dual system view explains why addiction is persistent and why relapse rates are high: automatic processes are notoriously difficult to control (e.g., Logan, 1988; Shiffrin & Schneider, 1977). The controlled system, in contrast, is slower, rational and involves more complex propositional knowledge. The latter system is assumed to prevent drug use in addicts (e.g., rational thoughts about the damaging long-term effects of smoking will stop smokers from lighting another cigarette). When the automatic and controlled system compete to govern behavior, the controlled system will win (and prevent drug pursuit) but only when an individual has the motivation and the cognitive resources available to do so. In other cases, the automatic system will win (and promote drug pursuit; e.g., Wiers et al., 2010). Hence, in order to understand drug pursuit, it is vital to understand also the automatic processes that contribute to drug pursuit.

Recently, alternative models have been proposed, in which there is no longer a strict separation between automatic and controlled systems. Within iterative reprocessing models (see Cunningham & Zelazo, 2007, for such a model in the context of evaluation in general, and Gladwin, Figner, Crone, & Wiers, 2011, a model in the context of addiction), the assumption is made that evaluations and subsequent behaviors are the result of evaluative processes that are based on representations in memory. These processes follow a specific hierarchy. Whereas lower-order evaluative processes provide basic information about valence and arousal, higher-order processes contain more complex evaluative information. The low-order processes are activated quickly, whereas the high-order processes are slower. Thus, when an individual has sufficient time, cognitive resources, and motivation, the system will reprocess the information, giving higher-order or "top-down biasing" (Gladwin et al., 2011) processes a chance to affect behavior. In contrast, when time, resources, and motivation are limited, lower-order processes will control behavior. Assuming that relational information exerts an impact already at early levels of reiteration, such models could explain that implicit beliefs are formed and influence (addictive) behavior.

It is argued that we can use so-called implicit measures to examine the automatic or "low-order" processes that drive addictive behavior. For instance, effects on implicit attitude measures are considered to be a proxy of automatic attitudes toward smoking (e.g., De Houwer, Teige-Mocigemba, Spruyt, & Moors, 2009). The most commonly used implicit measure in this context is the Implicit Association Test (IAT; Greenwald, McGhee, & Schwartz, 1998). IAT effects are assumed to be driven by the strength of mental associations in memory. Participants are required to classify target stimuli in one of two categories (e.g., "flowers" or "insects"). Furthermore, they are asked to categorize attribute stimuli as referring to one of two attributes (e.g., "pleasant" or "unpleasant"). Importantly, the same two keys are used for both tasks. In one test block, stimulus response-assignments are compatible with associations in memory (e.g., "flowers" and "pleasant" require the same response and "insects" and "unpleasant" require the other response), whereas in another test block, response-assignments are reversed (e.g., "flowers" are combined with "unpleasant" and "insects" are combined with "pleasant"). Commonly, participants are faster to respond in the compatible block compared to the incompatible block.

Different versions of the IAT have been developed to assess different types of drug associations. In this case, addiction-related targets (e.g., "smoking" or "alcohol") are compared with control targets (e.g., "non-smoking" or "soda") and different types of attributes are used to assess different types of associations. For example, IATs with attribute labels "active" and "passive" are assumed to asses implicit arousal associations, whereas attributes "pleasant" and "unpleasant" can be used to asses implicit valence (e.g., Wiers, Van Woerden, Smulders, & De Jong, 2002); the labels "I like" and "I dislike" are used to assess implicit liking in a way that minimizes effects of extra-personal associations (e.g., Houben & Wiers, 2006; Olson & Fazio, 2004); “approach” and “avoid” labels have been used to asses automatic approach/avoidance tendencies (e.g., Ostafin, Palfai, & Wechsler, 2003, Palfai & Ostafin, 2003); and “I want” and “I do not want” have been used to assess implicit wanting (e.g., Tibboel et al., 2011).

In spite of the increasing number of studies concerning the role of automatic processes in addiction, results have been mixed. For instance, whereas some research with different versions of the valence IAT has suggested that there is a relationship between drug use and positive implicit associations (e.g., De Houwer, Custers, & De Clercq, 2006; De Houwer & De Bruycker, 2007; Jajodia & Earlywine, 2003; McCarthy & Thompsen, 2006), other research found no relation or even suggested that there is in fact a relationship between drug use and *negative* implicit associations (e.g., Tibboel, De Houwer, Spruyt, et al., 2015; Wiers et al., 2002). Similarly, some findings suggested that there is a relationship between drug (ab)use and arousal associations (e.g., Houben & Wiers, 2006; Wiers et al., 2002), whereas other studies have revealed a relationship between drug use and *sedation* associations (e.g., Wiers, Houben, & De Kraker, 2007). Whereas the approach-avoidance IAT is not as widely used, results seem to consistently show a relationship between drug use and automatic approach associations (e.g., Ostafin & Palfai, 2006; Ostafin, Palfai, & Wechsler, 2003, Palfai & Ostafin, 2003; Wiers et al., 2010). The wanting IAT, on the other hand (Tibboel et al., 2011; Tibboel, De Houwer, Spruyt, et al., 2015) seems to consistently show no relationship between drug use and automatic drug wanting.

The mixed findings with different types of IATs shed doubts on the validity of these measures. In other words, it is unclear whether the outcomes of these different valence, arousal, approach/avoid, and wanting IATs actually reflect the attributes they aim to measure. It is important to note that most studies that examine automatic processes in addiction use correlational or quasi-experimental designs. As was noted by De Houwer et al. (2009), a measure can be considered valid only if variations in the to-be-measured attribute cause variations in the measure (see also Borsboom, Mellenbergh, & Van Heerden, 2004). In other words, in order to know whether a measure is a valid, we need to manipulate the drug-related attributes we aim to measure (i.e., implicit valence, wanting, arousal, or approach-tendencies) in an experimental manner and examine whether changes in these attributes cause changes in the measurement outcomes. Even though the existing results suggests that at least some of these measures seem to be very promising (e.g., the approach-avoid IAT), experimental research could give a more definitive answer to this question.

Recent studies in the context of Cognitive Bias Modification (CBM) have included different types of IATs to examine changes in implicit associations as a result of alcohol-avoidance training. For instance, in a seminal study, Wiers, Eberl, Rinck, Becker, and Lindenmeyer (2011), aimed to strengthen alcohol-avoidance associations in alcohol-dependent patients. An experimental group was trained to avoid alcohol-related pictures by pushing a joystick in response to alcohol-related pictures, whereas a control group received a sham training (i.e., the alcohol-related pictures were not predictive of whether participants should make an approach or avoidance response). Not only did alcohol avoidance-training decrease participants' relapse rates, it also resulted in a reduction of alcohol-approach associations as captured by the approach-avoidance IAT.

However, CBM procedures used by Wiers et al. (2011) are very demanding in that they involve multiple training sessions, and might lead not only to a change in implicit associations, but also to general changes in cognitive functioning (e.g., Beard, 2011). For instance, the relationship between effects on approach-avoidance tasks and relapse is moderated by participants' level of attentional control (e.g., Spruyt et al., 2013). It is possible that CBM improves does not directly change approach tendencies, but rather that it improves attentional control, which subsequently leads to a change in approach tendencies. We therefore opted to use a different type of manipulation to affect short-term changes in drug-related affective processes. The most efficient way to do this is by performing a deprivation manipulation; a manipulation that, surprisingly, is not widely used in research on implicit attitudes. Importantly, the severity of withdrawal symptoms is a strong predictor of relapse in smokers (e.g., Kenford et al., 2002; Piasecki et al., 2003). Furthermore, such a manipulation is assumed to increase implicit liking (e.g., Baker, Morse, & Sherman, 1987; Cox et al., 2006; Sherman et al., 2003), to increase arousal in response to drug cues (Parrot & Joyce, 1993), to strengthen approach bias (Watson, De Wit, Cousijn, Hommel, & Wiers, 2013), and increase craving and wanting (e.g., Field, Mogg, & Bradley, 2004; Tibboel et al., 2011) and should thus have an effect on a variety of addiction IATs.The first aim of our current research was to validate different types of IATs using a deprivation manipulation. In Experiment 1, we used an approach/avoid IAT and a liking IAT. The approach/avoid IAT has not yet been used in an experimental validation study, but correlational evidence suggests that this task taps into the incentive motivational properties of drug use. We expected that our deprivation manipulation would increase approach tendencies, resulting in higher approach/avoid IAT scores when participants were deprived compared to when they were sated. For the liking IAT, predictions are less clear. According to some researchers, this IAT measures implicit associations that play an important role in guiding drug seeking (e.g., Houben & Wiers, 2007). According to others (e.g., Wiers et al., 2002) this IAT measures “liking”, the hedonic component of drug use that actually does not play a causal role in drug seeking (e.g., Robinson & Berridge, 2003). However, experimental research did show that the liking IAT is sensitive to craving manipulations (Tibboel et al., 2011). The aim of Experiment 1 was therefore to examine whether previous findings using the liking IAT can be replicated, to examine the validity of the approach/avoid IAT, and to make a comparison between the two IATs.

In Experiment 2, we used a similar design to test the validity of two additional measures. First, we used the valence IAT. The popularity of the valence IAT to examine implicit attitudes towards drugs has decreased because, as mentioned above, results obtained with this task seem to be inconsistent in correlational and quasi-experimental studies. Furthermore, experimental studies using this IAT are rare and provide inconclusive evidence for its validity. To our knowledge, two experimental studies have been performed so far. First, Sherman, Rose, Koch, Presson, and Chassin (2003) compared two groups of smokers (one deprived group and one sated group) on a valence IAT. Results revealed no effect of this manipulation. However, inclusion criteria for this study were minimal: participants who smoked only one cigarette per day were eligible to participate. Within the literature on nicotine dependence, such smokers are at best considered “very light” smokers (Husten, 2009). Furthermore, deprivation lasted for only four hours and was manipulated in a between-subjects design. Possibly, with a more extreme manipulation and with a sample consisting of more seasoned smokers, there might be effects on a valence IAT. Second, Tibboel et al. (2011) used a valence IAT with a deprivation similar to that of the current study and did report increased scores when smokers were deprived (i.e., indicating more positive implicit attitudes). By including the valence IAT in our study, we hoped to increase our knowledge concerning effects of participants’ motivational state (i.e., craving caused by deprivation) on valence IAT scores.

A second aim of our study was to explore and validate a new type of implicit measure that was recently developed by De Houwer, Heider, Spruyt, Roets, and Hughes (2015): the Relational Responding Task (RRT). This measure was designed to assess the automatic activation of propositional beliefs. Like the IAT, the RRT is a categorization task. However, whereas the IAT is designed to examine automatic associations (e.g., cigarette – pleasant), the RRT is designed to examine automatic beliefs (e.g., “A cigarette would make me feel relaxed”). In other words, whereas the IAT focuses on an unqualified relationship between two concepts, the RRT focuses on more complex, relational, knowledge. Whereas dual systems models suggest that such knowledge is not part of the automatic system, and thus does not play a role in automatic drug pursuit, recent research suggests that such complex knowledge can, in fact, be activated automatically (e.g., Remue, De Houwer, Barnes-Holmes, Vanderhasselt & De Raedt, 2013; Remue, Hughes, De Houwer, & De Raedt, 2014; Wiswede, Koranyi, Müller, Langner, & Rothermund, 2013). In contrast, the single system view of iterative reprocessing models does leave room for the possibility that more complex information is activated at relatively early iteration stages. Nevertheless, within the field of addiction, the role of automatic beliefs remains unexplored.

Within the RRT, participants are required to classify inducer words (that are presented in a specific color) on the basis of whether they are synonyms of “true” or “false”. Using the same response keys, participants are required to also classify target statements (that are presented in another color) as being either “true” or “false” according to one of two opposing rules. In one critical test block, participants are asked to respond in line with the first rule; in another critical test block, they are asked to respond in line with the second rule. It is assumed that participants are faster to respond in blocks in which they are required to respond in line with the rule that matches their personal beliefs. For instance, in the original study of De Houwer et al. (2015), Flemish participants saw target statements that referred to the intelligence of Flemish people relative to immigrants (e.g., “Flemish people are more intelligent than immigrants”, “Immigrants are smarter than Flemish people”). The first rule was to respond as if Flemish people are more intelligent than immigrants (e.g., classify the statement “Flemish people are more intelligent than immigrants” as being true), whereas the second rule was to respond as if immigrants are more intelligent than Flemish people (e.g., classify the statement “Flemish people are more intelligent than immigrants” as being false). Results showed a positive correlation between the RRT score and scores on racism scales.

We aimed to measure specific implicit beliefs concerning smoking urges with a new version of the RRT. We used similar inducer words as De Houwer et al., and based the target statements on items of the brief version of the Questionnaire for Smoking Urges (QSU; Cox, Tiffany, & Christen, 2001), that was designed to explicitly examine participants’ urge to smoke and has been shown to be sensitive to deprivation manipulations (e.g., Tibboel et al., 2011). In one block, participants were instructed to respond as if they felt the urge to smoke a cigarette at that moment, whereas in another block they were instructed to respond as if they did not feel the urge to smoke a cigarette. We expected that smokers would find it easier to respond as if they have the urge to smoke when they had been deprived of cigarettes compared to when they had just smoked, resulting in higher RRT scores in the deprivation compared to the satiation condition.

2. Experiment 1

2.1. Method

2.1.1. Participants.

We tested 51 smokers who were recruited via advertisements on social media, via flyers handed out at different university buildings, and via the Ghent University website for experiment scheduling. Participants who indicated to be smokers on an online pre-screening website also received a personal email inviting them to participate. The sole criterion was that participants on average smoked ten or more cigarettes per day. Participants took part in two identical sessions and were paid twelve euro after the second session. They were required not to smoke during ten hours preceding one session, and to smoke just before the other session. The order of these sessions was counterbalanced.

2.1.2. Materials and procedure.

Participants were seated in a pleasantly lit testing room. They were seated in front of a 19-inch CRT-monitor with a refresh rate of 85 Hz, at a distance of approximately 60 cm. Because of a technical problem, the final three participants were tested on a laptop computer with a 17 inch screen at a distance of approximately 50 cm. Excluding these participants did not affect the pattern of results. We used a standard AZERTY keyboard to record responses. After giving informed consent, participants' breath carbon monoxide (CO) levels were measured using the Bedfont Smokerlyzer (Bedfont Instruments, Kent, UK) to check whether they had followed the instruction to abstain from smoking (in the deprivation condition) or to smoke just before the session (in the satiation condition). After this, participants filled in the Questionnaire for Smoking Urges (QSU). Subsequently, they performed the IATs and explicit ratings[1]. In Session 1 they then filled in the Fagerstrom

[1] We do not report the results of these ratings in detail because they are not relevant for our current discussion. Participants were asked to rate the neutral and smoking-related pictures on the extent to which they liked the pictures and the extent to which they would want to approach the pictures. We analyzed these ratings using a repeated measures ANOVA with factors condition (deprivation, satiation), rating type (liking, approach), and stimulus type (neutral, smoking-related). Most importantly, this yielded a significant three-way interaction, $F(43) = 12.49$, $p < .002$. When participants were sated, they gave higher liking scores than approach scores to the smoking-related stimuli, $t(44) = 5.00$, $p < .001$. Furthermore, participants liked smoking-related pictures better than neutral pictures, $t(44) = 2.64$, $p < .05$. However, they did not experience stronger approach tendencies toward smoking-related pictures, $t < 1$. When participants were deprived, there was no difference in liking and approach scores for the smoking-related pictures, $t < 1$. However, they liked smoking-related pictures better than neutral pictures, $t(45) = 4.84$, $p < .001$. Furthermore, they now also experienced stronger approach tendencies toward smoking-related pictures, $t(45) = 4.90$, $p < .001$. All other effects also reached significance, $7.94 < Fs < 29.05$.

questionnaire, whereas in Session 2 they filled in a brief questionnaire containing five smoking-related questions discussed below.

2.1.2.1. Questionnaire for Smoking Urges (QSU).

Participants completed the brief version of the Questionnaire for Smoking Urges (QSU; Cox, et al., 2001). This questionnaire measures two different aspects of smoking urges: the desire to smoke and the extent to which smoking is considered to be rewarding (Factor 1), and the anticipation of relief from negative affect (Factor 2).

2.1.2.2. IATs.

For the liking IAT, the labels were “I like” and “I do not like” in Dutch (“vind ik aangenaam” and “vind ik niet aangenaam”). We used eight exemplars of generic positive words (Dutch translations of “love”, “sunshine”, “warmth”, “peace”, “gift”, “vacation”, “freedom” and “friendship”) and negative words (Dutch translations of “war”, “sadness”, “depression”, “pain”, “sickness”, “fight”, “accident”, and “death”) as attribute stimuli. The target category labels were Dutch translations of “smoking” and “non-smoking” (“roken” and “niet roken”). Target stimuli were eight smoking-related pictures and eight neutral pictures that were matched for visual similarity (e.g., a cigarette vs. a pencil).

In the approach/avoid IAT, the labels were “approach” and “avoid” (“toenaderen” and “vermijden” in Dutch) and attributes were approach-related words (Dutch translations of “desire”, “closeby”, “go nearer”, “proximity”, “approximation”, “draw near”, “want”, “come closer”) and avoidance-related words (“flee”, “distance”, “leave”, “avert”, “retreat”, “escape”, “dodge”, “far away”). Target labels and stimuli were the same as in the liking IAT. In both IATs, all words were presented in 25 point Courier New font, picture sizes were 331 by 229 pixels. For 23 participants positive attributes were first paired with smoking targets, whereas for the remaining 25 participants positive attributes were first paired with non-smoking targets.

Participants responded by pressing either the left (“a”) or right (“p”) key. The IAT consisted of 7 blocks. The first block consisted of 24 trials on which only attribute words were presented. In a second block that also consisted of 24 trials, only target pictures were presented. In the third and fourth block, each consisting of 48 trials, both attribute words and target pictures were presented. All stimuli were presented in a random order. Smoking pictures and positive words required the same response, whereas non-smoking pictures and negative words required the other response. The fifth, sixth and seventh block were identical to the second, third and fourth block, respectively, except that the assignment for the target categories was reversed. If target pictures were classified incorrectly, a red X appeared for 400 ms. In line with Olson and Fazio (2004), error feedback was not given for responses to the positive and negative words. The intertrial interval (ITI) was 380 ms. All stimuli were presented randomly, without replacement.

2.1.2.3. Fagerstrom and additional questions

In the first session, participants filled in the Fagerström Test for Nicotine Dependence, which measures the degree of physical smoking dependence (Heatherton, Kozlowski, Frecker, & Fagerstrom, 1991). In the second session, participants filled in a brief questionnaire with 5 smoking related questions: “how long have you smoked”, “how many cigarettes do you smoke each day”, “do you have plans to quit smoking”, “how much would you like to quit smoking” and “how much time passes between the hour of waking and lighting your first cigarette”.

2.2. Results.

Of the 51 participants, three participants did not return for the second testing session. Furthermore, we excluded the data of participants who had higher CO levels in the deprivation condition compared to the satiation condition, suggesting they did not adhere to the instructions to refrain from smoking in one session and to smoke just before the other.[2] This led to the exclusion of one additional participant in our final analyses. Our final sample thus consisted of 47 participants. On average, participants smoked 12.50 cigarettes per day (*SD* = 4.80), and they had smoked for 4.55 years (*SD* = 2.36). They would smoke their first cigarette of the day within 69 minutes after waking

[2] We also analyzed the data with a stricter criterion: a difference of at least 4 ppm between the two conditions. This yielded similar results.

(*SD* = 57). Nine participants indicated to have plans to stop smoking. The average desire to quit smoking (on a 4 point scale) was 2.46 (*SD* = .60). The mean score on the Fagerstrom questionnaire was *M* = 3.09 (*SD* = 2.08).

2.2.1. Manipulation check.

CO levels differed significantly across sessions, $t(46) = 10.60$, $p < .001$. In the deprivation condition, average CO levels were *M* = 4.43 parts per million (*SD* = 3.57), and *M* = 13.87 (*SD* = 7.24) in the satiation condition.

Furthermore, participants had higher QSU scores in the deprivation condition compared to the satiation condition, *M* = 44.19, *SD* = 10.69, and *M* = 21.15, *SD* = 8.38, respectively. We also performed a repeated measures ANOVA with two factors: condition (deprivation vs. satiation) and QSU factor (Factor 1, Factor 2). This revealed a significant interaction effect, $F(46) = 135.38$, $p < .001$. When participants were deprived, their scores were higher than when they were sated. This was the case for both Factor 1 (reflecting the motivational/rewarding aspects of smoking), *M* = 28.62, *SD* = 6.08, and *M* = 12.72, *SD* = 5.88, respectively, as well as for Factor 2 (reflecting the relief of negative affect), *M* = 8.13, *SD* = 3.48, and *M* = 4.87, *SD* = 2.25, respectively. The interaction shows that the manipulation mainly affected Factor 1. The main effects for condition and for QSU factor both reached significance as well, $F(46) = 230.46$, $p < .001$, and $F(46) = 607.04$, $p < .001$.

2.2.2. IATs

We computed IAT scores on the basis of the D600 measure (Greenwald, Nosek, & Banaji, 2003). We analyzed these scores using a repeated measures ANOVA with factors condition (deprivation, satiation) and IAT (liking, approach/avoid). This yielded a significant main effect of IAT, $F(46) = 9.76$, $p < .005$, reflecting higher scores on the approach/avoid IAT compared to the liking IAT. Other effects did not reach significance, *F*s < 1.15.[3]

---

[3] We reanalyzed our data with a repeated measures ANOVA that included the number of cigarettes participants smoked each day as a covariate, to examine if effects were stronger for heavier smokers. This yielded no significant main effect of the number of cigarettes, nor any significant interactions, *F*s

Because our main aim was to examine whether our manipulation had an effect on each IAT separately, we performed t-tests to examine effects of our manipulation for the approach/avoid IAT and for the liking IAT. Results showed that there were no differences between the deprivation and the satiation conditions on the approach/avoid IAT, $t < 1$, $M = .32$, $SD = .40$, and $M = .31$, $SD = .45$, respectively. Nor was there a significant difference between the deprivation and satiation condition on the liking IAT, $t < 1.45$, $M = .22$, $SD = .42$, and $M = .12$, $SD = .57$, respectively.

We performed additional Bayesian analyses to investigate how strongly the data support the null hypothesis (i.e., no difference in IAT scores between the conditions) using the guidelines of Rouder, Speckman, Sun, Morey, and Iverson (2009). The Bayes factor (BF) indicates the extent to which the data support the absence (BF_0) or the presence (BF_1) of a significant effect. A BF smaller than 1 suggest there is "no evidence" for an effect, whereas a BF between 3 and 10 suggests that there is "anecdotal evidence", whereas a BF larger than 10 suggests "substantial evidence". For both the approach/avoid IAT, $BF_0 > 8$, and for the liking IAT, $BF_0 > 3$, there was substantial evidence in favor of the null hypothesis (i.e., no effect of the deprivation manipulation on IAT scores).

Finally, to examine test-retest validity of both IATs, we calculated the correlation between the two approach/avoid IAT scores (in the deprivation condition and the satiation condition), and the two liking IAT scores (for both conditions. Both for the approach/avoid IAT scores and the liking IAT scores these correlations were significant, $r(47) = .31$, $p < .05$, and, $r(46) = .59$, $p < .001$, respectively.

2.3. Discussion.

Our craving manipulation was successful in many respects: when participants were deprived, their explicit craving, CO-levels, and QSU-scores were higher than when they were sated. In contrast, our manipulation did not affect IAT scores. This sheds serious doubt on the validity of these IATs. In contrast, we found significant correlations between the two conditions for both IATs. This adds more

---

< .44. Second, we used Fagerstrom scores as a covariate, which also did not produce a significant main effect of the Fagerstrom score, nor any significant interaction effects, *F*s < .51.

evidence for the reliability of IATs over time (e.g., Lane, Banaji, Nosek, & Greenwald, 2007). However, this does not imply that IATs are valid measures of addiction-related implicit affect. The main effect for type of IAT did reveal that IAT scores for the approach/avoid IAT were overall higher than those for the liking IAT. However, the idea that absolute IAT scores are meaningful is disputed (e.g., Blanton & Jaccard, 2006).

3. Experiment 2

3.1. Method

3.1.1. Participants.

We tested 50 smokers who were recruited via flyers, social media and through word of mouth advertising. Payment and session orders were the same as in Experiment 1.

3.1.2. Materials and procedure.

Participants were seated in a pleasantly lit testing room in front of laptop with a 17-inch screen at a distance of approximately 50 cm from the screen of the laptop. After giving informed consent, participants' breath carbon monoxide (CO) levels were measured using the Bedfont Smokerlyzer (Bedfont Instruments, Kent, UK). Subsequently, they performed the RRT, followed by a questionnaire (the Fagerstrom questionnaire in the first session; the five smoking-related questions used in Experiment 1 in the second session), and finally they performed the valence IAT. Because different implicit measures might affect each other, it is common to counterbalance the order of the tasks. However, in this experiment we instead reduced such contamination effects by asking our participants to fill in questionnaires in between the two tasks. At the end of the session, participants filled in the QSU.

3.1.2.1. RRT.

We implemented the RRT using E-Prime (Schneider, Eschman, & Zuccolotto, 2002a, 2002b). We used ten different words for the inducer trials. Five words were synonyms for "true" (Dutch translations of "good", "accurate", "correct", "exact", "okay") and five words were synonyms for "false" (Dutch translations of "wrong", "untrue", "incorrect", "faulty", "mistaken"). We included these inducer statements to make sure that participants encoded the two responses as “true” and “false” (De Houwer et al., 2015). Furthermore, there were 24 target statements (see Appendix A). Half of these statements referred to the desire to smoke, whereas others referred to the desire not to smoke. Ten of these statements were adapted from the QSU. Inducer statements were always presented in yellow, whereas target statements were always presented in blue. All stimuli were presented in 18 point Verdana font.

Participants were instructed to categorize words and statements presented on the screen as either “true” or “not true” by pressing the A or P keys of the keyboard, respectively. In a first practice block, each of the 10 inducer words was presented 4 times, resulting in 40 trials. Participants were required to categorize words as being synonyms for “true” by pressing the “A” key or as synonyms for “false” by pressing the “P” key. In the second practice block of 48 trials, each of the 24 target statements were presented twice. This time, participants were instructed to respond as if they felt the urge to smoke a cigarette now. If a statement was true according to this rule, they were required to press the “A”. If it was false, they were required to press “P”. In the third block, a critical test block, 96 target and inducer trials were presented intermixed. In each block, each inducer statement was presented four times, and each target statement was presented twice. Participants were required to respond in the same manner as they did in the practice blocks.

The fourth block was another practice block in which the 24 target statements were again each presented twice. However, now the response rule was different: participants were instructed to respond as if they did not feel the urge to smoke a cigarette now. If a statement was true according to this rule, they were required to press the “A”. If it was false, they were required to press “P”. In the fifth block, another critical test block, 96 target and inducer trials were presented intermixed. In each block, each inducer statement was presented four times, and each target statement was presented twice.

Inducer words and target statements were presented in the center of the screen. They remained on the screen until a response was given. If a response was incorrect, a red X appeared in the center of the screen. It remained there until participants gave the correct response. The intertrial interval (ITI) was 750 ms.

3.1.2.2. Valence IAT.

The IAT was the same as the liking IAT in Experiment 1, except for the attribute labels: the labels we used were "positive" and "negative".

3.2. Results.

We used the same exclusion criteria as in Experiment 1, and excluded the data of four participants because their CO levels were higher in the deprivation condition compared to the satiation condition, indicating that they did not follow the instructions.[4] Our final sample thus consisted of 46 participants. Their mean age was 26.57 (SD = 5.96).

Participants on average smoked 17.98 cigarettes per day (SD = 9.99), had their first cigarette within 78 minutes after waking up in the morning (SD =100), and had smoked for 10.53 years (SD = 6.86). Eighteen participants had plans to stop smoking. On a 4-point scale, participants' wish to stop smoking was M = 2.84 (SD = .65). Participants' average Fagerstrom score was 3.93 (SD = 2.22), suggesting a tendency toward a moderate addiction.

3.2.1. Manipulation check.

CO levels were higher in the satiation condition compared to the deprivation condition, $t(45)$ = 9.45, p < .001. In the deprivation condition, average CO levels were M = 7.65 parts per million (SD = 5.56), whereas they were M = 16.54 (SD = 7.94) in the satiation condition.

Furthermore, participants' QSU scores were higher in the deprivation condition, M = 50.54, SD = 12.32, compared to the satiation condition, M = 33.24, SD = 11.55, $t(45)$ = 8.24, p < .001. A repeated measures ANOVA two factors: condition (deprivation vs. satiation) and QSU factor (Factor 1, Factor 2) did not reveal a significant interaction effect, F < 1.76. However, there was a main effect for factor, showing higher scores on Factor 1 compared to Factor 2 in both the deprivation condition, M =

[4] We also analyzed the data with a stricter criterion: a difference of at least 4 ppm between the two conditions. This yielded similar results. However, the effect of our manipulation on the RRT now approached significance, $t(33)$ = 1.97, p = .06, which is likely due to a decrease in power.

28.67, *SD* = 6.15, and *M* = 21.20, *SD* = 7.68, respectively, and in the satiation condition, *M* = 19.98, *SD* = 6.80, and *M* = 13.74, *SD* = 6.01, respectively. Furthermore, there was a main effect of condition, showing higher scores when participants were deprived.

3.2.2. RRT.

In order to analyze our RRT data we excluded inducer trials because they were only used to prevent response recoding (cfr. De Houwer et al., 2015). However, including these trials yielded a similar pattern of results. Second, in line with De Houwer et al., we transformed the data into D_{RRT} scores using the D-600 algorithm of Greenwald et al. (2003) with higher scores reflecting a stronger urge to smoke. A paired-samples t-test yielded a significant effect of deprivation condition, $t(45) = 2.18$, $p < .05$. RRT scores were higher in the deprivation condition, *M* = .16, *SD* = .33, compared to the satiation condition, *M* = .04, *SD* = .34.[5]

Because 10 of the 24 target items reflected the items of the QSU, we calculated separate RRT scores for these items as well. We performed these exploratory analyses in order to examine whether our manipulation differentially affected the different types of items. However, it must be noted that these analyses were thus based on only a relatively small subset of trials. A repeated measures ANOVA with factors condition (deprivation vs. satiation) and QSU factor (Factor 1, Factor 2) revealed only a main effect for session, $F(1, 45) = 6.79$, $p < .05$, reflecting higher scores in the deprivation condition compared to the satiation condition. Other effects did not reach significance, *F*s < 1.13. Planned comparisons did reveal a difference between the deprivation condition (*M* = .21, *SD* = .41) and the satiation condition (*M* = .01, *SD* = .38) for Factor 2 items, $t(45) = 2.71$, $p < .01$, but not for the Factor 1 items, $t < 1.27$, (*M* = .21, *SD* = .38 for the deprivation condition and *M* = .11, *SD* = .37 for the satiation condition). However, the lack of a significant interaction between condition and

[5] In order to check whether effects were stronger for heavier smokers, we performed a repeated measures ANOVA with session as a within-subjects variable and the number of cigarettes participants smoked as a covariate. This did not yield a main effect for the number of cigarettes, nor did it interact with the deprivation manipulation, *Fs* < .27. Again, we also used the Fagerstrom score as a covariate. This did not produce significant effects either, *Fs* < 1.23.

factor suggests that it is important to regard these results with some caution.We again calculated the BF. BF_1 was between 1 and 3, suggesting that there is "anecdotal evidence" for a significant difference between sessions.

Finally, the RRT score for the deprivation condition was significantly correlated with the RRT score for the satiation condition, $r(46) = .38$, $p < .01$, suggesting that the RRT is a reliable measure over time.**3.2.3. IAT.**

We computed IAT scores on the basis of the D600 measure (Greenwald et al., 2003). We performed a paired-samples t-test to compare these scores in the two conditions. However, IAT scores in the deprivation condition, $M = .22$, $SD = .36$, did not differ from those in the satiation condition, $M = .24$, $SD = .41$, $t < .27$. [6]

Our Bayesian analyses showed that there was substantial evidence in favor of the null hypothesis, $BF_0 > 6$ (i.e., no effect of the deprivation manipulation on the IAT).

Finally, there was no significant correlation between the two IAT scores in the two sessions, $r(46) = .19$, $p = .21$. This sheds some doubt regarding the reliability of this IAT.

3.3. Discussion.

Our craving manipulation can be considered successful on the basis of participants' CO-levels and QSU-scores.[7] However, this manipulation did not affect the IAT. These findings are in line with previous studies that show that the standard valence IAT does not seem to be a valid measure of the implicit processes that are assumed to play a role in addiction. In contrast, the newly developed RRT was in fact sensitive to our manipulation and seems to be a promising tool for future research on

---

[6] Again, we analysed the data in order to check whether effects might be stronger for heavier smokers. When we included the number of cigarettes participants smoked per day as a covariate, this did not yield a significant effect of the number of cigarettes, nor did it interact with the deprivation manipulation, *Fs* < .39. The same was the case when we used Fagerstrom scores as a covariate, *F*s < 1.82.

[7] It must be noted, however, that the manipulation did not differentially affect the two QSU factors, whereas it did in Experiment 1. One explanation for this is that in Experiment 1, participants filled in the QSU at the beginning of the session, whereas participants in Experiment 2 were asked to fill in the QSU at the end of the session. Thus, in Experiment 2, even in the deprivation condition, a considerable time had passed since participants had last smoked a cigarette. This might have led to higher scores on the reward or motivation factor (Factor 1) compared to the negative affect factor (Factor 2).

implicit cognition in the context of addiction. Furthermore, when we analysed only a subset of trials that were based on the QSU items, we still found a significant effect of our manipulation: RRT scores were higher when participants were deprived compared to when they were sated. Finally, whereas the two RRT scores were significantly correlated with each other, the IAT scores were not. However, we should be careful in making claims concerning the superiority of the RRT compared to the IAT, because 1) we have only used one version of the IAT and 2) the IAT always followed the RRT. Moreover, Bayesian analyses indicated that the evidence for an effect of deprivation on RRT scores was only anecdotal. On the basis of the evidence described above, the valence IAT that we used is arguably a task that has so far been less successful as a measure of automatic drug-related attitudes than other versions of the IAT. Moreover, even though we attempted to prevent contamination effects by inserting a delay between the two implicit measurement tasks, it is still possible that performing the RRT reduced the probability of finding the expected effects on the IAT (e.g., Lane, Banaji, Nosek, & Greenwald, 2007). This shortcoming should be addressed in future research. Finally, because the IAT was performed at the end of the session, it is possible that even participants who had smoked just before the session were already craving for their next cigarette, decreasing the effectiveness of our manipulation.

4. General discussion

In two experiments, we successfully manipulated smoking deprivation in a sample of smokers as revealed by their CO-levels and their scores on the QSU. However, our manipulation did not have the desired effect on the liking IAT, the approach/avoid IAT, and the valence IAT. In contrast, we did find effects on a new smoking urges RRT: deprivation lead to increased RRT scores indicating a stronger implicit urge to smoke. This effect was weak but still present when we analyzed only the subset of the target items that reflected QSU items. We will discuss our findings in relation to 1) our concerns regarding the validity of implicit measures and 2) implications for dual process models of addiction. Furthermore, we will discuss the limitations of our study, and we will suggest avenues for future research.

Regarding the validity of the different IATs that have been used in the context of addiction, our data paint a grim picture. The liking IAT is suggested to tap into one's personal implicit preference for specific targets (e.g., Houben & Wiers, 2007), whereas valence IAT scores might tap into more general valence associations (e.g., Wiers et al., 2002), and the approach/avoid IAT was designed to examine automatic approach associations (e.g., Palfai & Ostafin, 2003). Even though our manipulation should have affected such processes (e.g., Baker et al., 1987; Cox et al., 2006; Field et al, 2004; Parrot, & Joyce, 1993; Watson, et al., 2013), our experimental study provided no evidence for a causal relationship between one crucial aspect of addictive behavior (craving) and scores on these IATs. Even though manipulations are crucial to draw conclusions regarding the validity of (implicit) measures (Borsboom et al., 2004; De Houwer et al., 2009) experimental studies within this area of research are scarce. Our study thus makes an important contribution for research on implicit processes in addiction.

This raises the question why IATs seem to lack validity. One important problem is that the attribute labels used in the different IATs (i.e., I like/I do not like; approach/avoid; positive/negative) are open to interpretation (e.g., Tibboel, De Houwer, & Van Bockstaele, 2015). How a participant interprets these labels depends on a number of factors. So far, research has shown that IAT effects are influenced by the type of attribute-label that is used (e.g., Olson & Fazio, 2004; Han, Czellar, Olson, & Fazio, 2010), the type of target-label that is used (e.g., Mitchell, Nosek, & Banaji, 2003) as well as the type of exemplars in the target category (e.g., Dasgupta & Greenwald, 2001; Govan & Williams, 2004). Because we do not know how participants relate the target and attribute concepts, it is unclear what IAT scores mean. Some novel findings can illustrate this: recently, we compared a group of alcohol-dependent patients, a group of heavy social drinkers and a group of light social drinkers on a wanting IAT and a liking IAT (Tibboel, De Houwer, Spruyt, et al., 2015). Surprisingly, we found more negative IAT scores for both the patients and the light social drinkers compared to the heavy social drinkers, whereas there were no significant differences between the light drinkers and the patients. Even though IAT effects for light drinkers and patients did not differ, it is very likely that the automatic processes that underlie these effects *did* differ. For instance, light drinkers might have more negative automatically activated attitudes than heavy drinkers because the former group has fewer (positive) experiences with alcohol. Patients might have more negative attitudes compared to heavy drinkers because even though both groups have much experience with alcohol, the former has more negative experiences and expectancies regarding alcohol use. Thus, the fact that two strikingly *different* groups (i.e., light drinkers and alcohol-dependent patients) show strikingly *similar* IAT effects suggests that the IAT falls short as a measure that is assumed to be related to drug use and drug abuse.

A possible way to assess associative ambiguity is by using IATs that are unipolar (e.g., instead of examining the relationship between “smoking” and “positive” and “negative”, unipolar IATs use attribute labels “positive” and “neutral” or “negative” and “neutral”; e.g., Houben & Wiers, 2008). However, an important shortcoming to such IATs is that they are not able to distinguish between the *type* of relationship between target and attribute. It is clear that addiction can be driven by a range of different motivational processes. Two factors that are, for instance, suggested to play a role are the type of addiction (e.g., smoking, alcohol) and the type of drug user (e.g., casual users, frequent users, drug-dependent individuals). For example, whereas the relief of negative affect is assumed to be an important reason to continue smoking, intoxication is an important reason to continue drinking (e.g., Koob & Volkow. 2010). Furthermore, a casual drug user is motivated to use drugs because of social factors (e.g., having a drink with friends) whereas an addict has very different goals (e.g., drinking to avoid withdrawal symptoms). For IATs it is impossible to differentiate between these different motivations, like “smoking makes me feel relaxed” or “I like the taste of cigarettes”. However, this information is of crucial importance in order to predict this smoker’s behavior. For instance, in this example, the former belief is likely to play a far more significant role in the maintenance of addiction than the latter (e.g., Koob & Volkow, 2010). This issue has often been overlooked (e.g., Tibboel et al., 2015) but might be of importance in explaining some of the discrepant findings within this area of research.

This brings us to the second, more theoretical issue regarding the validity of dual system models and iterative reprocessing models of addiction. Whereas dual system models suggest that the automatic system contains simple associations, iterative reprocessing models suggest that early, low-order evaluations are driven merely by valence and arousal. However, the latter models allow for the possibility that implicit beliefs affect behavior at relatively early iterations. Our study provides preliminary evidence for the hypothesis that automatic evaluations can be based on implicit beliefs about the outcomes and effects of drug use. Thus, rather than a general automatic association between “cigarette” and “positive”, or “cigarette” and “arousing” we suggest that specific automatic beliefs (e.g., “smoking would relax me now”; “I need to smoke to avoid withdrawal”) drive drug pursuit for different types of drug users. We suggest that it is more fruitful to examine the automatic activation of more specific drug expectancies and motivations instead of a general association. This is in line with a recent model of De Houwer (2014) who suggested that implicit evaluation (and implicit cognition in general) can be based on propositional knowledge. Whereas an association is an unqualified relationship between two concepts (e.g., “cigarettes” – “positive”), a proposition does contain relational information (e.g., “cigarettes have a positive effect on my mood”; “cigarettes do not have positive effects on my health”). Our current RRT study is the first to suggest that implicit evaluation can indeed be influenced by the automatic retrieval of more complex propositional information. This might be useful in explaining discrepancies in the context of the findings reported above. Our results even tentatively suggest that different types of beliefs can be examined separately within one RRT, although more research is needed to verify this claim.

The idea that drug-related expectancies can be activated automatically was already proposed by Palfai and Wood (2001), who presented participants with a list of different statements regarding alcohol outcomes and asked them to list the first and second outcomes that came to mind when thinking about these statements. They found that these types of expectancies could reliably predict alcohol use. However, it is difficult to decide to what extent these beliefs can be considered “automatic”, as participants had ample time to reflect on these expectancies. Moreover, this single study has not been replicated or followed up in more recent research. Of course, it is possible to also question to what extent the RRT can be considered “implicit” or “automatic”. Generally, reaction times in the RRT are longer (for the target statements in our study, the average reaction time in milliseconds was M = 2167.69, SD = 794.62) than those in the IAT (in our study, the average reaction time in the valence IAT of Experiment 2 was M = 1081.02, SD = 345.16), suggesting that the underlying processes might not be considered fully automatic (e.g., Moors & De Houwer, 2006). However, it must be noted that slower responses are likely due to the complexity of the inducer statements. Sentences (in the RRT) take more time to process than single words (in the IAT). When we take this into account, reaction times can still be considered as fast (De Houwer et al, 2015). Furthermore, it must be noted that the RRT definitely an indirect measurement task: participants are not asked to indicate their smoking expectancies in a direct manner (“Do you feel the urge to smoke right now?”) but instead we draw conclusions regarding their urge to smoke on the basis of reaction times when they judge the veracity of specific claims according to specific rules. Whereas direct questions are generally affected by the desire to respond in a socially desirable manner (e.g., Greenwald et al., 1998), indirect measures such as RRT effects might be more difficult to fake. This issue is of great importance in the context of addiction, which can be considered a socially sensitive topic.

A second question regards the added value of the RRT. In our current study, we focused only on the question whether manipulations of craving affected RRT scores, but we do not know whether RRT scores can predict addictive behaviors. Furthermore, we need to examine to what extent the RRT adds value: does it predict behavior better than other implicit measures or explicit measures? It is clear that more research is needed to shed light not only on the extent to which different types of implicit beliefs affect addictive behaviors, but also under which circumstances. According to our hypothesis, it is possible that both positive and negative beliefs can affect behavior in an automatic manner (e.g., “A cigarette would taste nice” vs. “Cigarettes cause cancer”). In order to predict and eventually change addictive behavior, an important first step is to examine under which conditions different type of conflicting beliefs gain control over behavior.

Research on implicit beliefs is highly relevant for therapeutic interventions. For instance, CBM procedures have become an increasingly popular tool to change addictive behavior (e.g., Wiers, Gladwin, Hofmann, Salemink, & Ridderinkhof, 2013). These procedures are based on the main premises of both dual systems models and iterative reprocessing models of addiction (e.g., Wiers et al., 2007; Gladwin et al., 2011): the automatic or “low-level” associative processes drive addictive behavior, thus, by changing these processes, we can change addictive behavior. Our research sheds more light on the content of automatic processes, which should inform us about how to optimize CBM procedures. For instance, in the study of Wiers et al. (2011), CBM training improved treatment outcomes: whereas 59% of patients in a control group relapsed one year later, only 46% of participants who received alcohol-avoidance training relapsed. Whereas these findings are extremely promising, one might wonder what types of implicit cognitions are changed by this CBM task: simple alcohol-avoidance associations or more complex beliefs (e.g., “I should avoid alcohol”). Future research should examine how CBM affects different types of implicit beliefs, and how these beliefs predict addictive behavior. In a subsequent step, CBM procedures can be improved by targeting the relevant implicit beliefs. For instance, a CBM task that, instead of general avoidance tendencies, targets specific implicit beliefs regarding negative effects of drug use might have even stronger therapeutic effects in the long run.

4.1. Limitations

There are several limitations to our study that should not be overlooked. The effect of our manipulation on the RRT was relatively small (i.e., “anecdotal” according to the Bayesian analyses). In this context, it is important to note that tobacco deprivation manipulations do not consistently produce significant effects on tasks that capture cognitive processes (e.g., Heinz et al., 2007; Snyder & Henningfield, 1989). Even though manipulations are crucial to draw conclusions regarding the validity of (implicit) measures (Borsboom et al., 2004; De Houwer et al., 2009) studies in which affective processes are manipulated by depriving participants of cigarettes are scarce. Whereas withdrawal is an important predictor of drug-seeking (Piasecki et al., 2003), and models of addiction suggest that withdrawal should affect automatic processes (e.g., Cox, Fadardi & Pothos, 2006; Wiers et al., 2007), it is unclear to what extent withdrawal can be expected to have strong effects on implicit measures. Future research should use manipulations other than deprivation (e.g., cue exposure) to provide more evidence for the validity of the RRT as a measure of implicit beliefs concerning smoking.

A second, related, limitation is that smokers in our sample might not be considered heavy smokers. As mentioned in the introduction, there is no clear consensus about what makes a smoker a “light” smoker (Husten, 2009), but according to the norms of Fagerstrom (Heatherton et al., 1991), the smokers in our sample had a low to moderate dependence on cigarettes. It is possible that the deprivation manipulation failed to affect IAT scores in these samples, but that such a manipulation would be more successful in highly dependent smokers. However, we performed extra analyses to examine whether smoking severity (i.e., Fagerstrom scores and the number of cigarettes participants smoked on an average day) has an impact on the effect of deprivation on implicit associations and implicit beliefs. These analyses did not reveal effects of smoking severity, which provides some evidence that our manipulation was not less successful in lighter smokers. Finally, the fact that the manipulation did affect RRT scores suggests that the RRT has more potential to assess implicit processes in smoking addiction than the IATs we used in this study.

Finally, one concern that could be raised about Experiment 2 is that the order of the implicit measurement tasks (RRT and IAT) was not counterbalanced. Even though we aimed to avoid contamination effects by inserting a delay between both tasks, we cannot rule out the possibility that this increased chances of finding a significant effect of our manipulation on the RRT relative to the IAT. More research is needed to rule out these potential order effects.

5. Conclusion

To sum up, our research has important implications: we suggest that implicit beliefs play a role in addictive behavior, and that a more nuanced view of the role that automatic processes play in drug-related behaviors is needed. Furthermore, our findings open new avenues for future researchs: we need to examine the predictive validity of measures of implicit beliefs in order to shed more light on how and whether these beliefs play a (causal) role in addiction. Finally, this should aid in improving existing procedures that aim to modify automatic processes in addiction.

Declaration of Interest

Correspondence concerning this article should be addressed to Helen Tibboel, Department of Psychology, Ghent University, H. Dunantlaan 2, B-9000 Ghent, Belgium. Electronic mail may be sent to helen.tibboel@ugent.be. Preparation of this paper was supported by the Special Research Fund (BOF) and Methusalem Grant BOF09/01M00209 of Ghent University. The study was conducted in accordance with the principles expressed in the Declaration of Helsinki.

References

Baker, T.B., Morse, E., & Sherman, J.E. (1987). The motivation to use drugs: A psychobiological analysis of urges. In C. Rivers (Ed.), The Nebraska Symposium on Motivation: Vol. 34. Alcohol use and abuse (pp. 257–323). Lincoln: University of Nebraska Press.

Beard C. 2011. Cognitive bias modification for anxiety: current evidence and future directions. *Expert Review of Neurotherapeutics, 11,* 299–311.

Borsboom, D., Mellenbergh, G. J., & van Heerden, J. (2004). The concept of validity. *Psychological Review, 111*, 1061-1071.

Cox, L. S., Tiffany, S. T., & Christen, A. G. (2001). Evaluation of the brief questionnaire of smoking urges (QSU-brief) in laboratory and clinical settings. *Nicotine & Tobacco Research, 3*, 7-16.

Cox, W. M., Fadardi, J. S., & Pothos, E. M. (2006). The Addiction-Stroop Test: Theoretical considerations and procedural recommendations. *Psychological Bulletin, 132*, 443-476.

Cunningham, W. A., & Zelazo, P. D. (2007). Attitudes and evaluations: a social cognitive neuroscience perspective. *TRENDS in Cognitive Sciences, 11,* 97 – 104.

Dasgupta, N., & Greenwald, A. G. (2001). On the malleability of automatic attitudes: Combating automatic prejudice with images of admired and disliked individuals. *Journal of Personality and Social Psychology, 81,* 800–814.

De Houwer, J. (2014). A propositional perspective on context effects in human associative learning. *Behavioral Processes, 104*, 20-25.

De Houwer, J., Custers, R., & De Clercq, A. (2006). Do smokers have a negative implicit attitude toward smoking? *Cognition and Emotion, 20*, 1274-1284.

De Houwer, J., & De Bruycker, E. (2007). The identification-EAST as a valid measure of implicit attitudes toward alcohol-related stimuli. *Journal of Behavior Therapy and Experimental Psychiatry, 38*, 133-143.

De Houwer, J., Heider, N., Spruyt, A., Roets, A., & Hughes, S. (2015). The Relational Responding Task: Toward a New Implicit Measure of Beliefs. *Frontiers in Psychology, 6.*

De Houwer, J., Teige-Mocigemba, S., Spruyt, A., & Moors, A. (2009). Implicit Measures: A Normative Analysis and Review. *Psychological Bulletin, 135*, 347-368.

Gladwin, T. E., Figner, B., Crone, E. A., & Wiers, R. W. (2011). Addiction, adolescence, and the integration of control and motivation. *Developmental Cognitive Neuroscience, 1,* 364 – 376.

Govan, C. L., & Williams, K. D. (2004). Reversing or eliminating IAT effects by changing the affective valence of the stimulus items. *Journal of Experimental Social Psychology, 40,* 357–365.

Greenwald, A. G., McGhee, D. E., & Schwartz, J. L. K. (1998). Measuring individual differences in implicit cognition: The implicit association test. *Journal of Personality and Social Psychology, 74*, 1464-1480.

Greenwald, A. G., Nosek, B. A., & Banaji, M. R. (2003). Understanding and using the Implicit Association Test: I. An improved scoring algorithm. *Journal of Personality and Social Psychology, 85*, 197-216.

Han, H. A., Czellar, S., Olson, M. A., & Fazio, R. H. (2010). Malleability of attitudes or malleability of the IAT? *Journal of Experimental Social Psychology*, *46*, 286-298.

Heatherton, T. F., Kozlowski, L. T., Frecker, R. C., & Fagerstrom, K. O. (1991). The fagerstrom test for nicotine dependence: A revision of the fagerstrom tolerance questionnaire. *British Journal of Addiction, 86*, 1119-1127

Houben, K., & Wiers, R. W. (2006). Assessing Implicit Alcohol Associations with the Implicit Association Test: Fact or Artifact? *Addictive Behaviors, 31*, 1346-1362.

Houben, K., & Wiers, R. W. (2007). Personalizing the alcohol-IAT with individualized stimuli: Relationship with drinking behavior and drinking-related problems. *Addictive Behaviors, 32*, 2852-2864.

Houben, K., & Wiers, R. W. (2008). Implicitly positive about alcohol? Implicit positive associations predict drinking behavior. *Addictive Behaviors, 33*, 979-986.

Jajodia, A., & Earlywine, M. (2003). Measuring alcohol expectancies with the implicit association test. *Psychology of Addictive Behaviors, 17*, 126-133.

Kenford, S. L., Smith, S. S., Welter, D. W., Jorenby, D. E., Fiore, M. C., & Baker, T. B. (2002). Predicting relapse back to smoking: contrasting affective and physical models of dependence. *Journal of Consulting and Clinical Psychology, 70,* 216 – 227.

McCarthy, D. M., & Thompsen, D. M. (2006). Implicit and explicit measures of alcohol and smoking cognitions. *Psychology of Addictive Behaviors, 20*, 436-444.

Mitchell, J. A., Nosek, B. A., & Banaji, M. R. (2003). Contextual variations in implicit evaluation. Journal of Experimental Psychology: General, 132(3), 455–469.

Moors, A., & De Houwer, J. (2006). Automaticity: A conceptual and theoretical analysis. *Psychological Bulletin, 132*, 297-326.

Olson, M. A., & Fazio, R. H. (2004). Reducing the influence of extrapersonal associations on the implicit association test: Personalizing the IAT. *Journal of Personality and Social Psychology, 86*, 653-667.

Ostafin, B. D., & Palfai, T. P. (2006). Compelled to consume: The Implicit Association Test and automatic alcohol motivation.*Psychology of Addictive Behaviors*, *20*(3), 322-327.

Ostafin, B. D., Palfai, T. P., & Wechsler, C. E. (2003). The accessibility of motivational tendencies toward alcohol: Approach, avoidance, and disinhibited drinking. *Experimental and Clinical Psychopharmacology, 11,* 294-301.

Palfai, T. P., & Ostafin, B. D. (2003). Alcohol-related motivational tendencies in hazardous drinkers: Assessing implicit response tendencies using the modified-IAT. *Behaviour Research and Therapy, 41*, 1149-1162.

Parrott AC, Joyce C (1993). Diurnal patterns of stress and arousal in cigarette smokers, deprived smokers, and non-smokers. *Human Psychopharmacology, 8,* 21-28

Piasecki, T. M., Joernby, D. E., Smith, S. S., Fiore, M. C., & Baker, T. B. (2003). Smoking withdrawal dynamics: II. Improved tests of withdrawal-relapse relations. *Journal of Abnormal Psychology, 112,* 14 – 27.

Remue, J., De Houwer, J., Barnes-Holmes, D., Vanderhasselt, M.-A., & De Raedt, R. (2013).

Self-esteem revisited: Performance on the implicit relational assessment procedure as a measure of self- versus ideal self-related cognitions in dysphoria. *Cognition & Emotion, 27*, 1441-1449.

Remue, J., Hughes, S., De Houwer, J., & De Raedt, R. (2014). To be or want to be: Disentangling the role of actual versus ideal self in implicit self-esteem. *PLoS ONE* 9(9): e108837. Robinson, T. E. & K. C. Berridge (2003). Addiction. *Annual Review of Psychology, 54,* 1-29.

Rouder, J. N., Speckman, P. L., Sun, D., Morey, R. D., & Iverson, G. (2009). Bayesian t tests for accepting and rejecting the null hypothesis. *Psychonomic Bulletin & Review, 16*, 225–237. doi: 10.3758/PBR.16.2.225

Schneider, W., Eschman, A., & Zuccolotto, A. (2002a). *E-Prime User's Guide*. Pittsburgh: Psychology Software Tools Inc.

Schneider, W., Eschman, A., & Zuccolotto, A. (2002b). *E-Prime Reference Guide*. Pittsburgh: Psychology Software Tools Inc.

Sherman, S.J., Presson, C.C., Chassin, L., Rose, J. & Koch, K. (2003) Implicit and explicit attitudes toward cigarette smoking: the effects of context and motivation. *Journal of Social and Clinical Psychology*, **22**, 13–39.

Spruyt, A., De Houwer, J., Tibboel, H., Verschuere, B., Crombez, G., Verbanck, P., Hanak, C., & Noël, X. (2013). On the predictive validity of automatically activated approach/avoidance tendencies in abstaining alcohol-dependent patients. Drug and Alcohol Dependence, 127, 81 - 86.

Tibboel, H., De Houwer, J., Spruyt, A., Brevers, D., Roy, E., & Noël, X. (2015). Heavy social drinkers score higher on implicit wanting and liking for alcohol than alcohol-dependent patients and light social drinkers. *Journal of Behavior Therapy and Experimental Psychiatry, 48,* 185 – 191.

Tibboel, H., De Houwer, J., Spruyt, A., Field, M., Kemps, E., & Crombez, G. (2011). Testing the validity of implicit measures of wanting and liking. *Journal of Behavior Therapy and Experimental Psychiatry, 42*, 284-292.

Watson, P., De Wit, S., Cousijn, J., Hommel, B. & Wiers, R. W. (2013). Motivational Mechanisms

Underlying the Approach Bias to Cigarettes. *Journal of Experimental Psychopathology, 4,* 250 – 262.

Wiers, R. W., Bartholow, B. D., Van den Wildenberg, E., Thush, C., Engels, R. C. M. E., Sher, K. J., Grenard, J., Ames, S. L., & Stacy, A. W. (2007). Automatic and controlled processes and the development of addictive behaviors in adolescents: A review and a model. *Pharmacology, Biochemistry, and Behavior, 86,* 263 – 283.

Wiers, R. W., Houben, K., & De Kraker, J. (2007). Implicit cocaine associations in active cocaine users and controls. *Addictive Behaviors, 32,* 1284-1289.

Wiers, R. W., Rinck, M. Kordts, R., Houben, K., & Strack, F. (2010). Re-training Automatic Action-Tendencies to Approach Alcohol in Hazardous Drinkers. *Addiction, 105,* 279-287.

Wiers, R. W., Van Woerden, N., Smulders, F. T. Y., & De Jong, P. J. (2002). Implicit and explicit alcohol-related cognitions in heavy and light drinkers. *Journal of Abnormal Psychology, 111*, 648-658.