# Implicit 'wanting' without implicit 'liking': A test of incentive-sensitization-theory in the context of smoking addiction using the wanting-implicit-association-test (W-IAT)

Laura Anne Grigutsch[*], Gesa Lewe, Klaus Rothermund, Nicolas Koranyi

*Friedrich Schiller University of Jena, Germany*

## ABSTRACT

*Background and objectives:* According to incentive-sensitization theory (IST), addiction is characterized by the decoupling of two subconsciously operating psychological processes 'wanting' (i.e., incentive salience) and 'liking' (i.e., sensory pleasure). The present study set out to test predictions derived from IST in the context of smoking addiction with two variants of the Implicit Association Test (IAT): a Liking-IAT and a Wanting-IAT. In line with IST, we hypothesized that smokers differ from nonsmokers with regard to 'wanting' but not 'liking'.
*Methods:* Smokers (n = 24) and nonsmokers (n = 24) completed a Liking-IAT (L-IAT) and a Wanting-IAT (W-IAT) to assess their implicit 'liking' and 'wanting' for smoking-cues. Smokers completed these measures twice: once immediately after smoking, and once after a 12 h period of abstinence.
*Results:* While nonsmokers exhibited negative scores on both IATs that were highly correlated, smokers' W-IAT scores were significantly more positive than and uncorrelated with their L-IAT scores. In line with the notion of chronically increased 'wanting' in addicted individuals, smokers' W-IAT scores were unaffected by the deprivation manipulation.
*Limitations:* Results were obtained on a non-clinical sample. Compliance with abstinence instructions was assessed solely via self-report.
*Conclusion:* Results obtained in this study support the assumption that nicotine addiction is linked to a dissociation of 'wanting' and 'liking' for smoking as postulated by IST. Furthermore, the study provides further evidence that the newly developed W-IAT-is a valid measure of implicit 'wanting' that can be used to examine the processes that underlie human reward seeking behavior.

## 1. Introduction

When directly questioned, most addicts will state to be well aware of the adverse consequences of their drug-taking behavior, yet to be unable to suppress consumption despite not even liking the drug anymore – a phenomenon sometimes referred to as "the paradox of addiction" (Stacy & Wiers, 2010). Given the enormous costs addiction causes not only to inflicted individuals but to society at large, the search for scientific explanations of the phenomenon is the focus of a considerable branch of research.

The fact that most addictive substances initially produce euphorigenic effects of some kind encouraged the formulation of reinforcement-based explanations (Markou et al., 1993; Solomon & Corbit, 1973). According to those, drugs are initially taken for the pleasure experienced following consumption (positive reinforcement) and later, when the hedonic effects of the drug eventually wear off due to tolerance development, consumption is continued for its alleviating effect on aversive withdrawal symptoms (negative reinforcement). Upon closer examination of the characteristics of addictive behavior, however, several aspects become apparent that are hard to explain exclusively in terms of positive and negative reinforcement. Most importantly, the extremely high risk of relapse which persists even after years of abstinence and cessation of withdrawal symptoms cannot be explained with positive or negative reinforcement. Another strand of theorizing sees addiction as the result of a pathological alteration of processes involved in learning and memory caused by repeated drug-influence (Everitt, Dickonson, & Robbins, 2001; Hyman, 2005). According to this view, repeated exposure to addictive substances leads to an extreme overlearning of reward-directed behaviors and associated cues, thus turning originally instrumental (i.e., reward-oriented) drug-

taking behavior into automatized cue-triggered habits. Again, however, the specific characteristics of addictive behavior such as its compulsivity and flexibility challenge an explanation in terms of habit learning alone (Robinson & Berridge, 2008). First of all, behavioral patterns in general do not become compulsive just by being well-trained habits. Furthermore, addicts in a state of acute craving will do almost anything to procure the desired drug, resorting even to extreme (e.g., criminal) actions they never engaged in before, thus defying an explanation based on simple stimulus-response habits. Yet another line of reasoning links addictive behavior to impaired executive functioning, especially losses in the ability of behavioral inhibition and weighing of different behavioral alternatives caused by repeated drug-intake (Jentsch & Taylor, 1999). However, while drug-induced changes in orbitofrontal brain regions and accompanying losses in executive control do occur in addicts (e.g., Rogers et al., 1999), their role as cause vs. effect of compulsive drug-taking behavior is still debated (Schoenbaum & Shaham, 2008).

In the face of these challenges, Robinson and Berridge formulated their influential *Incentive-Sensitization Theory* (IST, Robinson & Berridge, 1993; see also Berridge & Robinson, 1995; 2003) of addiction, which gives a comprehensive account of addictive behavior. The major claim of IST is that a pattern of repeated and intermittent drug-taking behavior leads to chronic neuroadaptational changes in the mesolimbic dopamine system, rendering this brain system hyper-sensitized to the drug and to drug-associated cues. Critically, IST claims that this brain system is responsible for the attribution of incentive salience, the process by which stimuli become 'wanted' (i.e., highly attention grabbing, attractive, and potent to elicit approach behavior), but is not involved in mediating implicit 'liking' (i.e., hedonic impact). In short, IST proposes that the paradoxical nature of addictive behavior is caused by a decoupling of 'wanting' and 'liking' for the addictive substance, caused by an increase in learned 'wanting' for the drug that is not matched by a concurrent increase in 'liking'. The compulsivity and impulsivity of addictive behavior is, thus, explained by an excessive increase in motivational value attributed to drugs and drug-associated stimuli, subjugating addictive behavior to the external control of certain addictive cues. At the same time, the flexibility of drug-seeking and drug-taking behavior is compatible with IST, since what is learned is the association between certain stimuli and their motivational value, not a specific behavioral response. Furthermore, as drug-induced neuroadaptational changes in the mesolimbic dopamine system are postulated to be extremely long-lasting, IST is able to explain why even after years of abstinence, addicts are still at a high risk of relapse which is often triggered by encounters with previously drug-associated cues or contexts. Lastly, since the brain system affected by repeated drug-intake is assumed to be responsible for the assignment of motivational value but not for the experience of hedonia, IST is able to explain the paradoxical finding that long-time addicts often report the compulsion to pursue and consume the drug despite not experiencing any pleasurable drug-effects anymore.

Owing to its great explanatory power, IST is relevant not just for addiction, but for research on motivated behavior in general. So far, though, most conclusive evidence for IST stems from studies on animals, where 'wanting' is typically operationalized via actual consumption or instrumental behavior directed at obtaining a reward, whereas 'liking' is coded via affective facial expressions in the so-called taste reactivity paradigm (e.g., Berridge & Valenstein, 1991; Pecina, Cagniard, Berridge, Aldridge, & Zhuang, 2003; Wyvell & Berridge, 2000). When it comes to human behavior, however, measurement of the processes in question poses a problem. Due to their controllability, consumption behavior and facial expressions are open to faking and thus are unsuitable indices of 'wanting' and 'liking' (Tibboel, De Houwer, Spruyt, Kemps, & Crombez, 2011; Tibboel, De Houwer, & van Bockstaele, 2015). Explicit self-report also disqualifies as a valid operationalization since both processes are assumed to operate at least partially outside conscious awareness (Anselme & Robinson, 2016;

Berridge & Robinson, 2003). Furthermore, ordinary language use does not clearly distinguish between 'wanting' and 'liking' so that people may confuse these terms or use them idiosyncratically. Therefore, researchers interested in 'wanting' and 'liking', and their respective contributions to motivated human behavior have turned to implicit measures as a promising tool to solve the above mentioned problems.

One of the most widely used implicit measures is the Implicit Association Test (IAT, Greenwald, McGhee, & Schwartz, 1998). It consists in a combination of two binary categorization tasks where stimuli have to be sorted on two different dimensions with the same pair of keys. In the most widely employed version of the IAT, one set of stimuli (the attributes) has to be categorized according to valence (e.g., "positive" vs. "negative"), whereas the other stimuli (the targets) have to be sorted according to their category membership (e.g., "flowers" vs. "insects"). During the critical blocks of the test, attribute and target stimuli are presented in an alternating fashion. In the compatible block, the evaluation of the target and attribute categories that are assigned to the same key matches (e.g., flower/positive vs. insect/negative), whereas it mismatches in the incompatible block (e.g., insect/positive vs. flower/negative). The difference in performance between the compatible and incompatible blocks is interpreted as reflecting the relative implicit likability for one target category over the other.

Based on the idea that the nature of the attribute categorization task determines *what* is being measured in an IAT, several researchers have tried to develop different versions of the IAT for the assessment of 'wanting' and 'liking', respectively, which – if valid – should allow researchers to detect wanting-liking dissociations in the context of addiction as hypothesized by IST. The resulting 'liking'-IATs typically share the above-mentioned set-up of a valence-based discrimination task on the attribute dimension (e.g., 'positive' vs. 'negative'), while different approaches have been taken to develop a corresponding 'wanting'-IAT.

Wiers, van Woerden, Smulders, and de Jong (2002), for example, devised a variant of the IAT with the target categories "alcohol" and "soda" where the attribute task required participants to sort words into categories labelled "active" and "passive". In this task, heavy – but not light –drinkers exhibited significant alcohol-arousal associations, while a standard 'liking'-IAT with the attribute categories "positive" and "negative" revealed both participant groups to hold equally strong alcohol-negative associations. The authors related their findings to IST by pointing out that, as a consequence of incentive salience, 'wanted' stimuli should become arousing, rendering the strength of arousal-associations an index of relative implicit 'wanting' for the employed target concepts. However, this reasoning was later challenged by findings of Wiers, Houben, and Kraker (2007) who demonstrated that cocaine users exhibit strong cocaine-sedation rather than cocaine-arousal associations. These results indicate that the arousal-IAT might be more apt to measure implicit associations reflecting drug-effects instead of actual 'wanting' for the respective substance.

A similar approach was chosen by Palfai and Ostafin (2003) who created an approach/avoidance-IAT by having participants sort synonyms of "approach" and "avoidance" into attribute categories that were marked by these labels. Employing the target concepts "alcohol" vs. "electricity", the authors could demonstrate that alcohol-approach associations were related to problematic drinking behaviors. Since the process of incentive salience attribution is supposed to result in the impulse to approach 'wanted' stimuli, alcohol-approach associations assessed with this variant of the IAT were argued to reflect 'wanting' in the sense of IST. Contrary to the notion that 'wanting' should be more important in explaining addictive behavior than 'liking', however, De Houwer, Custers, and de Clercq (2006) found that the approach/avoidance-IAT did not outperform a valence-based 'liking'-IAT in differentiating smokers from nonsmokers.

By directly employing the category labels "I want" and "I don't want", Tibboel et al. (2011) tried to create a wanting-IAT that directly taps into the concept of 'wanting' without taking a detour via

theoretically associated constructs, like arousal or approach. Again however, this wanting-IAT failed to outperform a valence-based liking-IAT in discriminating smokers from nonsmokers. Furthermore, scores obtained with this wanting-IAT and a valence-based liking-IAT were found to be highly correlated in the smoker sample, where, according to IST, the two processes should be decoupled.

In sum, existing evidence obtained with different 'wanting'-IAT measures in the context of addiction is mixed at best, and the question arises whether this heterogeneity of findings is simply the result of poor validity of the employed IAT-measures, or whether indeed wanting-liking dissociations do not play the central role in addiction postulated by IST. As was meticulously pointed out by Tibboel et al. (2015), there is indeed ample reason to question the validity of the existing wanting-IAT measures. The accuracy of IST-postulates for human behavior thus still has to be tagged with a question mark.

In the present study, we aim to provide evidence for the decoupling of 'wanting' and 'liking' in human addiction by employing a newly developed and validated version of a Wanting-IAT (W-IAT; Koranyi, Grigutsch, Algermissen, & Rothermund, 2017). The W-IAT differs from all aforementioned attempts to develop a 'wanting'-IAT in one central aspect. Specifically, all previous 'wanting'-IATs basically focused on changing the semantic meaning of the attribute category labels into something related to the concept of 'wanting' (e.g., activation, approach, or "wanting"). In contrast, in the W-IAT, the attribute response itself is endowed with consummatory quality, which allows us to asses 'wanting' based on stimulus-response compatibility effects (De Houwer, 2001). This is achieved by making execution of one of the IAT responses (e.g., the right key press) instrumental for the satisfaction of a currently activated physiological need. Specifically, participants are made thirsty before completing the Wanting-IAT by having to eat salty snacks. The attribute task of the following Wanting-IAT then consists in sorting stimuli that are either relevant (images of drinks) or irrelevant (images of neutral non-drink objects) for need-satisfaction into the categories "I want" vs. "I don't want". By correctly sorting drink stimuli into the category "I want", participants can gain water for later consumption. The consummatory character of the "I want"-response is further highlighted by immediate visual and auditory action effects signaling a gain in water following each correct classification of a drink. Evidence for the validity of this Wanting-IAT was already provided by a first study in the mating context, where it was compared to a valence-based Liking-IAT (Koranyi et al., 2017). The target task required classification of male and female facial photographs as "attractive" vs. "not attractive". Heterosexual male participants exhibited a preference for attractive female over attractive male faces in the Wanting-IAT while no such difference was observed in the Liking-IAT. These results are in line with previous findings of a wanting-liking dissociation in the mating context that were obtained with explicit measures (Dai, Brendl, & Ariely, 2010).

### 1.1. The present study

The present study aims at testing predictions derived from IST regarding a dissociation between 'wanting' and 'liking' in addiction by employing the new Wanting-IAT. For this purpose we adopted an experimental set-up that closely matched the design used by Tibboel et al. (2011). Smokers and nonsmokers completed the Wanting-IAT, as well as a standard Liking-IAT. Target categories in both IATs were "smoking" and "tooth brushing". While nonsmokers took the experiment once, smokers completed two experimental sessions, one of which was preceded by a 12 h period of abstinence from smoking while the other session started with the smoking of a cigarette.

Drawing on IST, we expected the following pattern of results: Since addictive behavior should be mainly driven by 'wanting', but not 'liking' for addiction-cues, we expected an interaction between the factors participant group (smokers vs nonsmokers) and IAT-type (Liking-IAT vs Wanting-IAT). This interaction should be driven by the fact that smokers as compared to nonsmokers should obtain

significantly higher scores in the Wanting-IAT, while any difference between both groups observed in Liking-IAT scores should be significantly smaller. As IST further predicts that 'wanting' for addiction-cues should be decoupled from 'liking' in addicts, but not in non-addicts, we expected Wanting- and Liking-IAT scores to be uncorrelated in smokers, but correlated in nonsmokers. Lastly, we expected the smoking of one cigarette immediately before completing the other measures to quench smokers' explicit desire to smoke. With respect to the implicit measures, we expected the deprivation manipulation to have a greater effect – if any – on implicit 'wanting' than on implicit 'liking'.

## 2. Method

### 2.1. Sample

Fifty participants took part in the present study for partial fulfillment of course credits. Twenty-four participants were smokers who smoked a minimum of five cigarettes per day. The other 26 participants were nonsmokers according to self-report. One participant of the smoker sample was excluded from all analyses for indicating a number of cigarettes smoked per day that was lower than the pre-defined minimum number of five. Two more participants were excluded from all analyses due to not meeting our pre-defined requirements for nonsmokers. [Fn: These participants indicated being nonsmokers at the moment, but reported a period of regular smoking in their past. As, according to IST, the neuroadaptational changes leading to a hypersensitization of the 'wanting' system in addicts are assumed to be extremely long-lasting, only participants were included as nonsmokers who indicated never having met our criteria for the smoking sample in their life.] The exclusions left a final sample of 23 smokers (11 females) and 24 nonsmokers (10 females). Mean age was 22.96 (*SD* = 4.90) for smokers and 24.13 (*SD* = 6.45) for nonsmokers. Smokers smoked on average 9.87 (*SD* = 2.30) cigarettes per day and had been smoking on a regular basis for an average of 9.74 (*SD* = 12.04) years. Twelve smokers reported having unsuccessfully tried to quit smoking in the past. Seven smokers indicated a current wish to quit smoking.

While nonsmokers completed a single session of the experiment, smokers were asked to come to the laboratory twice with a mean temporal distance of 5.67 days (*SD* = 2.89 days) between sessions. At one of the sessions (satiation condition), smokers were asked to smoke immediately before completing the experiment, whereas they were requested to abstain from smoking for at least 12 h before the other session (deprivation condition). The order of deprivation and satiation was counterbalanced so that half of the smoker-sample was deprived at the first session and sated at the second, while it was the other way round for the other half of the smoker-sample.

### 2.2. Materials

Target stimuli in both, the W-IAT and the L-IAT were pictures taken from the internet whose content could be unambiguously assigned to the categories "smoking" (e.g., a burning cigarette) or "tooth brushing" (e.g., a toothbrush). Attribute stimuli in the L-IAT were taken from the International Affective Picture System (IAPS; Lang, Bradley, & Cuthbert, 1997) and were selected based on their valence ratings to represent the categories "positive" (e.g., puppies) and "negative" (e.g., cockroaches). In the W-IAT, attribute stimuli were free stock images taken from http://www.freeimages.com which depicted either household objects (e.g., a light bulb) representing the category "I don't want" or drinks (e.g., a glass of water) representing the category "I want". All categories were represented by 8 different pictures.

### 2.3. Procedure

The following is a description of a single session of the current

experiment. For smokers, who were required to come to the laboratory in two separate sessions, the setup of the two sessions was identical with the sole exception that the order of L-IAT and W-IAT was counterbalanced across participants and sessions, and that they were required not to smoke for at least 12 h before one of the sessions, and to smoke immediately before the other.

On arriving at the laboratory, participants were seated at computer work stations in individual cubicles where they completed the different parts of the experiment. At the beginning of a session, participants completed the brief version of the Questionnaire of Smoking Urges (QSU-brief, Cox, Tiffany, & Christen, 2001) and answered questions concerning their smoking behavior. To check for compliance with the deprivation instructions, participants were explicitly asked to indicate when they had smoked their last cigarette before coming to the laboratory. For reasons of comparability, nonsmokers were administered the same questionnaires at the beginning of their experimental session, even though these data were discarded from further analysis due to a lack of variance. After collection of these explicit measures, participants completed the Liking-IAT (L-IAT) and the Wanting-IAT (W-IAT). The order of L- and W-IAT was balanced over participants and (for smokers) experimental sessions.

*L-IAT.* The L-IAT consisted of five blocks that were completed by all participants in a fixed order. The first block comprised 16 practice trials of the attribute categorization task which consisted in sorting pictures appearing on screen into the categories "negative" (left key) and "positive" (right key) according to their valence. Response keys (the 'D' and 'L' key) were marked on a QWERTZ-keyboard. Next, participants completed a block of 16 practice trials of the target categorization task in which the same two response keys were used to sort pictures of either toothbrushes or cigarettes into the categories "tooth brushing" (left key) vs. "smoking" (right key). The third block consisted of 96 critical test trials in which both tasks were combined. Stimuli of the target and attribute dimension appeared on screen in random order and had to be sorted into their respective category. During this block, each of the response keys corresponded to two categories, one of each dimension (left key: "negative/tooth brushing", right key: "positive/smoking"). The fourth block comprised another 16 practice trials of the target task. This time, however, category-to-key assignment for the target task was reversed. In the fifth and final block, participants completed another 96 trials of the combined task with the new category-to-key-assignment for the target dimension (i.e., left key: "negative/smoking", right key: "positive/tooth brushing").

Category labels were presented in the upper left and right corners of the screen throughout the entire test. Stimuli were presented in random order one-at-a-time. They appeared centrally and remained on screen until a response was made. Trials were separated by a 1500 ms inter-trial-interval. If the latency of a response exceeded a pre-defined threshold of 750 ms, participants received a feedback message that their response was "too slow", which was displayed for 750 ms. Erroneous responses were indicated by the appearance of a red "X" below the stimulus which required the delivery of a correct response before the next trial started.

*W-IAT.* Before starting the W-IAT, participants were subjected to a thirst induction procedure, which is critical for assuring the motivational character of the attribute task. For this purpose, participants were handed a plate with salty crackers and were informed that they had to consume at least 8 crackers within 2 min in order to continue with the experiment. Next, participants were informed that they would be able to gain water for later consumption during the next part of the experiment. Also, the consumption of additional crackers at the end of the experiment was announced to further increase the motivation to gain water.

Participants then started the W-IAT, whose structure was identical to that of the L-IAT with the following exceptions: Category labels in the attribute task were "I don't want" (left key) and "I want" (right key) and attribute stimuli consisted in pictures of household objects ("I don't

want") and drinks ("I want"). For correct categorizations of drinks with response times that were faster than 750 ms, participants earned 10 ml of water for later consumption. This gain was signaled to them by the appearance of a small shot glass at the lower part of the screen and the playing of drink-associated sounds (e.g., water gurgling) via headphones.

After completing the two IATs, participants answered questions on their general smoking behavior, their water consumption during the day, and their subjective motivation to earn water during the W-IAT.

## 3. Results

We analyzed the data in two steps. First, we compared smokers with nonsmokers, using only data of the first session in the smoker-sample in order to have measures that are fully comparable for both samples. Second, to test for an effect of the deprivation manipulation, we compared the two measurement sessions (satiation vs. deprivation) within the smoker-sample (for a similar analysis procedure, see Tibboel et al., 2011). Both analyses will be reported in separate sections.

### 3.1. Group comparisons

IAT effect scores were calculated based on the D600 measure defined by Greenwald, Nosek, and Banaji (2003). In the L-IAT, positive values indicated 'liking' for smoking. Similarly, positive scores in the W-IAT indicated 'wanting' for smoking. IAT effects were submitted to a 2 (IAT type: L-IAT/W-IAT) x 2 (group: smokers/nonsmokers) mixed-designs analysis of variance with repeated measurement on the first factor (see Fig. 1). The ANOVA yielded a significant main effect of group, $F(1, 45) = 34.66$, $p < .001$, $\eta^2 = 0.435$, indicating more negative scores for nonsmokers ($M = -0.26$, $SD = 0.28$) than for smokers ($M = 0.11$, $SD = 0.28$), a significant main effect of IAT type, $F(1, 45) = 34.58$, $p < .001$, $\eta^2 = 0.435$, indicating more negative scores in the L-IAT ($M = -0.20$, $SD = 0.29$) as compared to the W-IAT ($M = 0.03$, $SD = 0.35$), and the predicted interaction of group and IAT type, $F(1, 45) = 11.74$, $p = .001$, $\eta^2 = 0.207$. Pairwise comparisons were performed to further explain this interaction. These comparisons revealed that nonsmokers scored more negatively than smokers on both, L-IAT (smokers: $M = -0.08$, $SD = 0.24$; nonsmokers: $M = -0.31$, $SD = 0.29$), $t(45) = -3.03$, $p = .004$, $d = 0.89$, and W-IAT (smokers: $M = 0.29$, $SD = 0.19$; nonsmokers: $M = -0.22$, $SD = 0.27$), $t(45) = -7.23$, $p < .001$, $d = 2.12$, the group difference being larger for W-IAT scores. Furthermore, both smokers and nonsmokers scored more negatively on the L-IAT than on the W-IAT. However, the difference was significant only for smokers, $t(22) = -6.09$, $p < .001$, $d_z = -0.71$, but not for nonsmokers, $t(23) = -1.89$, $p = .072$, $d_z = -0.25$.
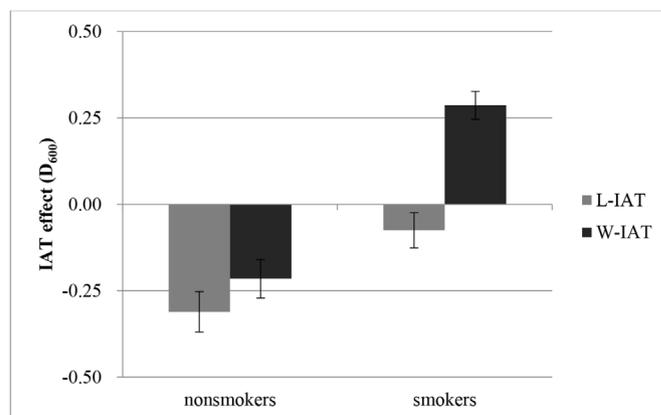
**Fig. 1.** Mean IAT effect scores depending on IAT-type and Group. Error bars depict standard errors of the mean.

*Correlations between W-IAT and L-IAT*. Across groups, there was a significant and substantial correlation of effect scores of W-IAT and L-IAT, $r = 0.58$, $p < .001$. Crucially, a comparison of the separate correlation coefficients for smokers and nonsmokers using Fishers's z-transformation revealed that the W-IAT/L-IAT-correlation differed significantly between groups, $z = 1.71$, $p = .040$. While there was a strong a significant correlation in the nonsmoker sample, r = 0.61, p = .002, the correlation in the smoker sample was only, r = 0.17, p = .430. The nonsignificance of the latter should be interpreted with caution, though, since a Bayesian analysis conducted with JASP using the standard prior (JASP Team, 2018) yielded a Bayes Factor, $BF_{01}$, of 2.88. While this means that the observed data are almost three times more likely under the nullhypothesis (no correlation) than under the alternative hypothesis, a Bayes Factor of this size is conventionally interpreted as only "anecdotal evidence" in favor of the nullhypothesis.

### 3.2. Deprivation vs. satiation

*Effects on IAT effects*. IAT effects were calculated based on the D600 measure (Greenwald et al., 2003) for both IATs and both conditions (deprivation/satiation) separately. IAT-effects were subjected to a 2 (condition: deprived/sated) x 2 (IAT type: L-IAT/W-IAT) analysis of variance with repeated measurement on both factors. The ANOVA yielded a significant main effect of IAT type, $F(1, 22) = 47.96$, $p < .001$, $\eta^2 = 0.686$, indicating more positive scores in the W-IAT ($M = 0.22$, $SD = 0.25$) as compared to the L-IAT ($M = -0.10$, $SD = 0.24$). Neither the main effect of condition, nor the interaction of condition x IAT type reached significance (both $Fs < 1$).

*Explicit Wanting*. Explicit wanting ratings within the group of smokers were obtained by averaging the ratings on all items of the Questionnaire of Smoking Urges. A comparison of these ratings between conditions revealed that smokers reported a significantly stronger desire to smoke ($M = 3.95$, $SD = 1.00$) if the session was preceded by a 12 h period of abstinence, than when they had just smoked prior to taking the tests ($M = 1.72$, $SD = 0.50$), $t(22) = 9.97$, $p < .001$, $d_z = 1.87$.

## 4. Discussion

In the present study we tested predictions derived from incentive-sensitization-theory (Robinson & Berridge, 1993) concerning the interplay of implicit 'wanting' and 'liking' for addiction-relevant cues in addicts and non-addicts. Specifically, we assessed smokers' and non-smokers' implicit 'wanting' and 'liking' for smoking-cues with two variants of the Implicit Association Test: a standard valence-based Liking-IAT (Greenwald et al., 1998) and the newly developed and validated Wanting-IAT (Koranyi et al., 2017). The two IATs differ with respect to the response quality established by the attribute task. In the Liking-IAT, attribute stimuli have to be sorted according to their valence, turning the response into an evaluative 'liking'-response. In the Wanting-IAT, attribute stimuli have to be sorted according to their motivational property of being 'wanted', which is established by making one of the two response keys (labelled "I want") instrumental for the satisfaction of a currently active need (thirst). In line with predictions derived from incentive-sensitization-theory, we found evidence for a decoupling of 'wanting' and 'liking' in the smoker sample. While nonsmokers exhibited negative scores on both types of IATs, smokers' scores on the Wanting-IAT were significantly more positive than their nominally negative respective Liking-IAT scores.

The dissociation of 'wanting' and 'liking' for smoking in smokers was further corroborated by the correlational analyzes between the two IAT-scores. While nonsmokers displayed a strong positive correlation of IAT-scores, indicating correspondence between 'wanting' and 'liking', the relationship between both measures was significantly weaker in smokers. The comparison between smokers and nonsmokers in the present study, thus, fits nicely with predictions of IST in that 'wanting' and 'liking' for smoking go hand in hand in nonsmokers but become uncoupled in smokers.

In the smoker sample, we further tested the influence of a deprivation manipulation on implicit 'wanting', as well as the explicit desire to smoke. For this purpose smokers were required to complete two identical sessions of the experiment that differed only in the instruction to smoke immediately before one and to abstain from smoking for at least 12 h before the other. Consistent with the notion that addictive behavior is characterized by chronically increased implicit 'wanting' for the addictive substance, the deprivation manipulation had no influence on Wanting-IAT scores, even though the explicit desire to smoke was significantly increased in the deprivation condition. This finding suggests that subjects' explicit desire for smoking might be related predominately to consciously accessible factors such as withdrawal-related discomfort. Implicit 'wanting', however, is activated in a bottom-up manner by the mere presence of smoking-cues in the environment irrespective of the presence or absence of withdrawal symptoms (Robinson & Berridge, 1993).

In a recent study, Tibboel, De Houwer, Dirix, and Spruyt (2017) also found that deprivation had no effect on different versions of a smoking-related IAT, but had an influence on another implicit measure - a variant of the Relational Responding Task (RRT; De Houwer, Heider, Spruyt, Roets, & Hughes, 2015) assessing implicit beliefs regarding smoking urges. These findings can be explained by assuming (a) that the IAT variants that were used did not assess 'wanting' proper by not employing a motivationally charged attribute task, and that (b) the RRT functions similar to measures of explicit wanting, capturing beliefs about 'wanting' rather than 'wanting' itself. Findings in the RRT thus paralleled our findings for self-reports of 'wanting'.

Of course, our findings should not be taken to indicate that wanting is independent of deprivation or internal states of need. Instead, we would rather argue that even in a situation of low deprivation, cue-triggered wanting can compensate for a lack of internal drive (Joyner, Kim & Gearhardt, 2017). Especially in addiction, environments containing stimuli that function as strong signals for potential rewards can elicit motivational states of cue-triggered 'wanting' even in the absence of internal need states (Berridge & Aldridge, 2009; Robinson, Robinson & Berridge, 2013), since the neural 'wanting'-system has become hyper-persensitized to attribute incentive salience to drug-related cues. The current version of our W-IAT might already suffice to induce such a need, because we used highly appetitive stimuli that evoke strong associations with nicotine consumption (e.g., burning cigarettes, people smoking cigarettes).

### 4.1. Conclusion

Taken together, the pattern of results captured with the Liking-IAT and the Wanting-IAT is perfectly in line with predictions formulated on the basis of IST. Specifically, the dissociation of implicit 'wanting' and 'liking' for addiction cues seems to be an essential feature of addiction. Furthermore, the findings suggest that the two IATs employed in the present study are able to capture separate implicit processes involved in human reward seeking behavior. These results are highly relevant not only for research on addiction, but also for research on motivated behavior in other content domains, since they highlight the fact that evaluative measures might not always be suitable indices of behavioral tendencies when implicit 'wanting' is not taken into account. The present study can thus be directly linked to previous research puzzling over the fact that humans engage in certain behaviors (e.g., gambling, watching pornography, eating high-calorie food) despite negative (implicit) evaluations of the behavior held by the actors themselves (e.g., Swanson, Swanson, & Greenwald, 2001). The presence of increased 'wanting' might be the often neglected variable driving behavior in such situations.

## Acknowledgements

## Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jbtep.2019.01.002.

## References

Anselme, P., & Robinson, M. J. F. (2016). Wanting," "liking," and their relation to consciousness. *Journal of Experimental Psychology: Animal Learning and Cognition, 42*(2), 123–140. https://doi.org/10.1037/xan0000090.

Berridge, K. C., & Aldridge, J. W. (2009). Decision utility, incentive salience, and cue-triggered "wanting". *Oxford series in social cognition and social neuroscience,* 509–533 2009.

Berridge, K. C., & Robinson, T. E. (1995). The mind of an addicted brain: Neural sensitization of wanting versus liking. *Current Directions in Psychological Science, 4*(3), 71–76.

Berridge, K. C., & Robinson, T. E. (2003). Parsing reward. *Trends in Neurosciences, 26*(9), 507–513. https://doi.org/10.1016/S0166-2236(03)00233-9.

Berridge, K. C., & Valenstein, E. S. (1991). What psychological process mediates feeding evoked by electrical stimulation of the lateral hypothalamus? *Behavioral Neuroscience, 105*(1), 3–14.

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 : Official Journal of the Society for Research on Nicotine and Tobacco, 3*(1), 7–16. https://doi.org/10.1080/14622200020032051.

Dai, X., Brendl, C. M., & Ariely, D. (2010). Wanting, liking, and preference construction. *Emotion, 10*(3), 324–334. https://doi.org/10.1037/a0017987.

De Houwer, J. (2001). A structural and process analysis of the implicit association test. *Journal of Experimental Social Psychology, 37*, 443–451. https://doi.org/10.1006/jesp.2000.1464.

De Houwer, J., Custers, R., & Clercq, A. de (2006). Do smokers have a negative implicit attitude toward smoking? *Cognition & Emotion, 20*(8), 1274–1284. https://doi.org/10.1080/02699930500484506.

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*, 319. https://doi.org/10.3389/fpsyg.2015.00319.

Everitt, B. J., Dickinson, A., & Robbins, T. W. (2001). The neuropsychological basis of addictive behaviour. *Brain Research Reviews, 36*(2–3), 129–138.

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*(6), 1464–1480. https://doi.org/10.1037//0022-3514.74.6.1464.

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*(2), 197–216. https://doi.org/10.1037/0022-3514.85.2.197.

Hyman, S. E. (2005). Addiction: a disease of learning and memory. *American Journal of Psychiatry, 162*(8), 1414–1422.

JASP Team (2018). *JASP.* ([Computer software])Version 0.9.

Jentsch, J. D., & Taylor, J. R. (1999). Impulsivity resulting from frontostriatal dysfunction in drug abuse: implications for the control of behavior by reward-related stimuli. *Psychopharmacology, 146*(4), 373–390.

Joyner, M. A., Kim, S., & Gearhardt, A. N. (2017). Investigating an incentive-sensitization model of eating behavior: impact of a simulated fast-food laboratory. *Clinical Psychological Science, 5*(6), 1014–1026.

Koranyi, N., Grigutsch, L. A., Algermissen, J., & Rothermund, K. (2017). Dissociating implicit wanting from implicit liking: Development and validation of the Wanting Implicit Association Test (W-IAT). *Journal of Behavior Therapy and Experimental Psychiatry, 54*, 165–169. https://doi.org/10.1016/j.jbtep.2016.08.008.

Lang, P. J., Bradley, M. M., & Cuthbert, B. N. (1997). *International affective picture system (IAPS): Technical manual and affective ratings.* NIMH Center for the Study of Emotion and Attention.

Markou, A., Weiss, F., Gold, L. H., Caine, S. B., Schulteis, G., & Koob, G. F. (1993). Animal models of drug craving. *Psychopharmacology, 112*(2–3), 163–182.

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*(10), 1149–1162. https://doi.org/10.1016/S0005-7967(03)00018-4.

Pecina, S., Cagniard, B., Berridge, K. C., Aldridge, J. W., & Zhuang, X. (2003). Hyperdopaminergic mutant mice have higher "wanting" but not "liking" for sweet rewards. *Journal of Neuroscience, 23*(28), 9395–9402.

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

Robinson, T. E., & Berridge, K. C. (2008). The incentive sensitization theory of addiction: some current issues. *Philosophical Transactions of the Royal Society B: Biological Sciences, 363*(1507), 3137–3146.

Robinson, M. J. F., Robonsin, T. E., & Berridge, K. C. (2013). Incentive salience and the transition to addiction. In P. Miller (Ed.). *Biological Research on Addiction* (pp. 391–399). San Diego, USA: Academic Press.

Rogers, R. D., et al. (1999). Dissociable deficits in the decision-making cognition of chronic amphetamine abusers, opiate abusers, patients with focal damage to prefrontal cortex, and tryptophan-depleted normal volunteers: evidence for monoaminergic mechanisms. *Neuropsychopharmacology, 20*(4), 322–339.

Schoenbaum, G., & Shaham, Y. (2008). The role of orbitofrontal cortex in drug addiction: a review of preclinical studies. *Biological Psychiatry, 63*(3), 256–262.

Solomon, R. L., & Corbit, J. D. (1973). An opponent-process theory of motivation: II. Cigarette addiction. *Journal of Abnormal Psychology, 81*(2), 158–171.

Stacy, A. W., & Wiers, R. W. (2010). Implicit cognition and addiction: a tool for explaining paradoxical behavior. *Annual Review of Clinical Psychology, 6*, 551–575.

Swanson, J. E., Swanson, E., & Greenwald, A. G. (2001). Using the Implicit Association Test to investigate attitude-behaviour consistency for stigmatised behaviour. *Cognition & Emotion, 15*(2), 207–230. https://doi.org/10.1080/02699930125706.

Tibboel, H., De Houwer, J., Dirix, N., & Spruyt, A. (2017). Beyond associations: Do implicit beliefs play a role in smoking addiction? *Journal of Psychopharmacology, 31*(1), 43–53. https://doi.org/10.1177/0269881116665327.

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*(3), 284–292. https://doi.org/10.1016/j.jbtep.2011.01.002.

Tibboel, H., De Houwer, J., & van Bockstaele, B. (2015). Implicit measures of "wanting" and "liking" in humans. *Neuroscience & Biobehavioral Reviews, 57*, 350–364. https://doi.org/10.1016/j.neubiorev.2015.09.015.

Wiers, R. W., Houben, K., & Kraker, J. de (2007). Implicit cocaine associations in active cocaine users and controls. *Addictive Behaviors, 32*(6), 1284–1289. https://doi.org/10.1016/j.addbeh.2006.07.009.

Wiers, R. W., van Woerden, N., Smulders, F. T. Y., & Jong, P. J. de (2002). Implicit and explicit alcohol-related cognitions in heavy and light drinkers. *Journal of Abnormal Psychology, 111*(4), 648–658. https://doi.org/10.1037/0021-843X.111.4.648.

Wyvell, C. L., & Berridge, K. C. (2000). Intra-accumbens amphetamine increases the conditioned incentive salience of sucrose reward: Enhancement of reward "wanting" without enhanced "liking" or response reinforcement. *Journal of Neuroscience, 20*(21), 8122–8130.