Learning in alcohol dependence

Pavlovian-to-instrumental transfer in alcohol dependence and relapse risk

Zur Erlangung des akademischen Grades Doctor rerum naturalium (Dr. rer. nat.)
   im Fach (Promotionsfach) Psychologie

eingereicht
an der Lebenswissenschaftlichen Fakultät

von Dipl.-Psych. Maria Garbusow

Präsidien der Humboldt-Universität zu Berlin:
Prof. Dr.-Ing. Dr. Sabine Kunst

Dekan der Lebenswissenschaftlichen Fakultät:
Prof. Dr. Bernhard Grimm

GutachterIn:
1. Prof. Dr. Norbert Kathmann
2. Prof. Dr. Dr. Andreas Heinz
3. Dr. Annette Horstmann

Tag der Verteidigung: 11.07.2017
EIDESSTATTLICHE ERKLÄRUNG

Hiermit erkläre ich,

- dass ich die vorliegende Arbeit selbstständig und ohne unerlaubte Hilfe verfasst habe,

- dass ich die Dissertation an keiner anderen Universität eingereicht habe und keinen Doktorgrad in dem Promotionsfach Psychologie besitze und


Berlin, 08.02.2017, Maria Garbusow
I am very grateful to my supervisor Prof. Andreas Heinz from Charité Universitätsmedizin Berlin, who always supported and challenged me and thus motivated me to maximize the quality of this dissertation. I thank Prof. Andreas Heinz for the opportunity to be a member of his research group and for the time he invested in my supervision. I am also thankful to my second supervisor Prof. Norbert Kathmann from Humboldt-Universität zu Berlin, who kindly supported my dissertation and was always available for all my scientific and organizational questions. I also would like to thank Dr. Annette Horstmann from Max Planck Institute for Human Cognitive and Brain Sciences Leipzig, who agreed to join my dissertation committee. I am very thankful for the support from all members of LeAD research group in Berlin and Dresden. I thank Dr. Florian Schlagenauf, Dr. Quentin Huys, Dr. Daniel Schad and Prof. Michael Rapp for all the support regarding data analyses and all the fruitful methodological and scientific discussions. Stephan Nebe, Christian Sommer, Sören Kuitunen-Paul and Elisabeth Jünger, I thank you for all your support in data acquisition and quality control. Dr. Eva Friedel and Dr. Claudia Hägele, thank you for always being there for psychiatric questions. I would like to express my special gratitude to Miriam Sebold and Alexander Genauck. You not only became scientific role models but also very good friends. Alexander, your scientific enthusiasm is impressive and motivating, thank you for all the supportive discussions about data analyses. Miriam, you always motivated me to think outside the box and to develop an own research profile. Thank you for all your support and inspiring discussions. You are the best colleague one can ever wish for! I would also like to thank all the subjects participating in this study. Without their time and motivation to help us study the fundamentals of alcohol dependence this study could obviously never have been conducted. Finally, I would like to express my most sincere gratitude to my family and friends first of all to my loving and wonderful partner Phillip, for their support and understanding.
**TABLE OF CONTENTS**

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>ABSTRACT</td>
<td>VII</td>
</tr>
<tr>
<td>ZUSAMMENFASSUNG</td>
<td>VIII</td>
</tr>
<tr>
<td>LIST OF PAPERS</td>
<td>IX</td>
</tr>
<tr>
<td>LIST OF ABBREVIATIONS</td>
<td>X</td>
</tr>
<tr>
<td>LIST OF FIGURES</td>
<td>XI</td>
</tr>
<tr>
<td>1. INTRODUCTION</td>
<td>1</td>
</tr>
<tr>
<td>2. THEORETICAL BACKGROUND</td>
<td>2</td>
</tr>
<tr>
<td>2.1 Value-based learning in addiction</td>
<td>2</td>
</tr>
<tr>
<td>2.2 Pavlovian-to-instrumental transfer</td>
<td>4</td>
</tr>
<tr>
<td>2.2.1 Nondrug-related PIT effects</td>
<td>5</td>
</tr>
<tr>
<td>2.2.2 Drug-related PIT effects</td>
<td>7</td>
</tr>
<tr>
<td>3. RESEARCH QUESTIONS</td>
<td>8</td>
</tr>
<tr>
<td>4. SUMMARY OF RELATED PAPERS</td>
<td>9</td>
</tr>
<tr>
<td>4.1 Paper I</td>
<td>12</td>
</tr>
<tr>
<td>4.2 Paper II</td>
<td>14</td>
</tr>
<tr>
<td>4.3 Paper III</td>
<td>16</td>
</tr>
<tr>
<td>4.4 Paper IV</td>
<td>18</td>
</tr>
<tr>
<td>5. DISCUSSION</td>
<td>21</td>
</tr>
<tr>
<td>5.1 Behavioral PIT effects</td>
<td>21</td>
</tr>
<tr>
<td>5.2 Neural PIT effects</td>
<td>23</td>
</tr>
<tr>
<td>5.3 Limitations</td>
<td>25</td>
</tr>
<tr>
<td>5.4 Outlook and final conclusion</td>
<td>26</td>
</tr>
<tr>
<td>5.4.1 Clinical implications</td>
<td>26</td>
</tr>
<tr>
<td>5.4.2 Directions for future studies</td>
<td>27</td>
</tr>
<tr>
<td>5.4.3 Final summary and conclusion</td>
<td>27</td>
</tr>
<tr>
<td>REFERENCES</td>
<td>29</td>
</tr>
</tbody>
</table>
ABSTRACT

**Importance.** This thesis summarizes the first studies investigating Pavlovian-to-instrumental transfer in alcohol-dependent (AD) patients and how this relates to functional activation, relapse risk and choice impulsivity. Hence it adds to animal and preclinical human addiction literature with implications for clinical interventions.

**Background.** It has long been hypothesized that contextual stimuli influence our behavior. This mechanism can be operationalized using a Pavlovian-to-instrumental transfer (PIT) task. Here animal and human studies have shown that positive Pavlovian stimuli enhance and negative Pavlovian stimuli reduce instrumental behavior (PIT effect). Regarding to substance use dependence, this mechanism might be relevant for relapse risk, as drug-associated stimuli have shown to enhance e.g. craving and functional activation in reward-related brain areas in patients compared to controls. In animal addiction models and preclinical human studies enhanced PIT effects have been described with functional activation particularly in the nucleus accumbens (NAcc). Moreover, control subjects with stronger PIT effects and AD patients have been shown to be more impulsive on different facets of impulsivity.

**Methods.** The PIT task consists of four parts: i) instrumental conditioning, ii) Pavlovian conditioning, iii) transfer with Pavlovian background stimuli and instrumental task in the foreground (nondrug-related PIT: Pavlovian contextual cues; drug-related PIT: alcohol-related contextual cues), and iv) query trials of the Pavlovian stimuli. Choice impulsivity was measured by choosing either a small immediate or a larger delayed reward. To investigate PIT-related activation we focused on bilateral NAcc region of interest analyses.

**Results.** We observed significantly enhanced nondrug-related PIT effects in AD patients compared to controls with a neuro-functional activation in the NAcc being predictive for relapse. Regarding drug-related PIT effects, we observed significantly reduced instrumental behavior during alcohol-related backgrounds with neural correlates in the NAcc in abstainers compared to relapsers. Choice impulsivity was positively related to PIT in AD patients only.

**Conclusion.** Our data suggest that PIT is a mechanism contributing to relapse in AD patients with neuro-functional correlations within the NAcc, which based on our data is involved in motivational properties as well as the attribution of salience. Moreover, the subgroup of high impulsive patients is particularly susceptible for PIT effects, thus should be main target for intervention programs.
ZUSAMMENFASSUNG


Schlussfolgerung. Die Studien lassen darauf schließen, dass PIT ein für Rückfall wichtiger Mechanismus ist mit neurofunktionalem Korrelat im NAcc, der sich in den Studien sowohl als für motivationale Prozesse als auch als Salienzsignal relevant gezeigt hat. Da die Subgruppe von hoch impulsiven Patienten im Besonderen durch Kontextreize im instrumentellen Antwortverhalten beeinflussbar ist, sollte ihr besondere Aufmerksamkeit bei Interventionen zukommen.
LIST OF PAPERS

This thesis is based on the following original papers:

**Paper I**

S. Karger AG kindly permits the use of the paper for this thesis.

**Paper II**

Wiley kindly permits the use of the paper for this thesis.

**Paper III**

Springer kindly permits the use of the paper for this thesis.

**Paper IV**

Springer Nature kindly permits the use of the paper for this thesis.

*equal contribution
<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Full Form</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>AD</td>
<td>alcohol-dependent</td>
<td></td>
</tr>
<tr>
<td>CIDI</td>
<td>Composite International Diagnostic Instrument</td>
<td></td>
</tr>
<tr>
<td>CS</td>
<td>conditioned stimulus</td>
<td></td>
</tr>
<tr>
<td>DA</td>
<td>dopaminergic</td>
<td></td>
</tr>
<tr>
<td>EEG</td>
<td>electro encephalography</td>
<td></td>
</tr>
<tr>
<td>EPI</td>
<td>echo planar imaging</td>
<td></td>
</tr>
<tr>
<td>fMRI</td>
<td>functional magnetic resonance imaging</td>
<td></td>
</tr>
<tr>
<td>LeAD</td>
<td>Learning in alcohol dependence</td>
<td></td>
</tr>
<tr>
<td>NAcc</td>
<td>nucleus accumbens</td>
<td></td>
</tr>
<tr>
<td>PET</td>
<td>positron emission tomography</td>
<td></td>
</tr>
<tr>
<td>PIT</td>
<td>Pavlovian-to-instrumental Transfer</td>
<td></td>
</tr>
<tr>
<td>ROI</td>
<td>region of interest</td>
<td></td>
</tr>
<tr>
<td>S-R</td>
<td>stimulus-response</td>
<td></td>
</tr>
<tr>
<td>US</td>
<td>Unconditioned stimulus</td>
<td></td>
</tr>
<tr>
<td>VS</td>
<td>Ventral striatum</td>
<td></td>
</tr>
<tr>
<td>VTA</td>
<td>ventral tegmental area</td>
<td></td>
</tr>
</tbody>
</table>
LIST OF FIGURES

Figure 1  Anatomical bilateral ROI of the NAcc derived from the WFU PickAtlas
Figure 2  LeAD cohort - schematic overview of sample sizes
Figure 3  Pavlovian-to-instrumental transfer paradigm
Figure 4  Summary of main results
1. INTRODUCTION

Detoxified alcohol-dependent (AD) patients often report that environments formerly associated with alcohol intake evoke craving and thus the likelihood of relapse. Relapse rates up to 80 percent are reported in the literature (Boothby & Doering, 2005) and point to the importance of better understanding mechanisms leading to relapse. How contextual cues influence the motivation to conduct instrumental behavior is operationalized in a Pavlovian-to-instrumental transfer (PIT) paradigm. For example, it has been shown that subjects prefer to choose a branded cup of lemonade compared to the same drink without label (McClure et al., 2004), chocolate in the presence of chocolate pictures (Seabrooke, Hogarth, & Mitchell, 2015) and unhealthy snacks in the presence of unhealthy compared to healthy snack pictures (Watson, Wiers, Hommel, Ridderinkhof, & de Wit, 2015). Moreover, food-associated cues enhance approach to food independent of hunger (Colagiuri & Lovibond, 2015; Watson, Wiers, Hommel, & de Wit, 2014), although PIT effects can be attenuated under satiety (Aitken, Greenfield, & Wassum, 2016).

These examples illustrate how the influence of contextual cues can lead to maladaptive behavior. From an evolutionary perspective, contextual cues biasing our value-based motivation can also be highly beneficial, especially in environments where rewards are uncertain (Cartoni, Moretta, Puglisi-Allegra, Cabib, & Baldassarre, 2015). For instance, when food sources are rare eating when food is available (signaled by food-related cues) not only when being hungry can save your life. Crucially, when transferring this phenomenon to an environment with plenty of food sources, it becomes maladaptive and thus diseases, in this example obesity, may arise (Reddish, 2013), possibly reflecting automatic overeating due to changes in motivational value of food in obese individuals (Horstmann et al., 2015).

From a theoretical perspective, it has long been hypothesized that PIT plays an important role in alcohol relapse (e.g. Di Chiara et al., 1999; Robbins & Everitt, 1999). However, how PIT effects and the neural correlates relate to relapse behavior in human alcohol dependence has not been investigated this far. Here, four original papers will be summarized that shed light on behavioral and neural PIT effects as well as its relation to relapse behavior in AD patients.
2. THEORETICAL BACKGROUND

2.1 Value-based learning in addiction

To understand how contextual cues play a role in relapse behavior, it is important to consider Pavlovian conditioning. During Pavlovian conditioning initially neutral cues are associated with reward or punishment (unconditioned stimulus, US). Thereby the US value is transferred to the initially neutral cue, acquiring rewarding properties themselves (conditioned stimulus; CS). This effect is measurable behaviorally by approach behavior towards the CS after conditioning with a reward (e.g. Fitzpatrick & Morrow, 2016) and neurobiologically by a shift of dopaminergic (DA) neurotransmission in the ventral tegmental area (VTA) from the US to the CS after conditioning (Schultz, 1998). However, there are individual differences: Flagel and colleagues showed that the shift in DA response from US to CS occurs only in so-called sign-tracking animals. After Pavlovian conditioning, sign-trackers approach the CS, i.e. the reward-predicting cue itself becomes attractive, while so-called goal-trackers still engage the US (Flagel, Akil, & Robinson, 2009). Crucially, sign-tracking is also associated with an addictive phenotype including greater behavioral disinhibition and reduced impulse control (Flagel et al., 2011). Impulsivity is not only a trait associated with dependence but also known to be a risk factor for poor treatment outcome in substance dependence (Loree, Lundahl, & Ledgerwood, 2015) and appears to play a role in the neural and behavioral consequences of value-based learning.

On a neural level, drugs act like reward, i.e. they directly influence DA neurotransmission (for review see Hyman, Malenka, & Nestler, 2006) and thus become motivationally relevant. Studies showed that drug consumption leads to increased DA release in the ventral striatum (VS), more specifically the nucleus accumbens (NAcc; Clarke, Söderpalm, Lotfi, Ericson, & Adermark, 2015; Kienast & Heinz, 2006; Robinson & Berridge, 1993). According to the law of effect postulated by Thorndike (1927), stimulus-response (S-R) associations are strengthened when followed by reward and weakened when followed by non-reward or punishment. Due to this neurobiological based rewarding effect of drugs, S-R learning is strengthened resulting in a self-reinforcing effect of the drug. Crucially, drug-related S-R associations are largely resistant to extinction (e.g. de Wit & Stewart, 1981), leading to the reinstatement model of drug relapse (Shaham, Shalev, Lu, de Wit, & Stewart, 2003). There are two potential factors of reinstatement. First, the inactivation of DA neurons in the VTA has been shown to block drug-induced reinstatement (McFarland & Kalivas, 2001), pointing to a relevant neurobiological mechanism. Secondly, renewed drug-seeking in a drug context
after extinction has been observed (Crombag & Shaham, 2002), reflecting the relevance of cue-induced motivational processes.

Consequently, drug consumption affects Pavlovian conditioning by promoting DA release for drug-associated stimuli; thus they become motivationally relevant and salient. In AD patients, salience attribution relates to drug cues, a phenomenon postulated by the *incentive salience theory* (Di Chiara & Bassareo, 2007; Di Chiara et al., 1999; Robinson & Berridge, 1993, 2001). On a psychological level, incentive motivation can be subdivided into at least two components: *wanting* and *liking*. Whereas *liking* is the hedonic affect or subjective pleasure of a drug associated with consumption, *wanting* reflects a motivational property in drug dependence that leads to enhanced motivation of drug consumption independent of hedonia (Robinson & Berridge, 2001). These two psychological processes are supposed to be mediated by different neural systems, with dopamine function in the VS as main target for investigating incentive salience attribution and hence *wanting* (Heinz et al., 2005). Initial voluntary and hedonic drug consumption may transform into loss of control over drinking with a shift from prefrontal (controlled behavior) to striatal (habitual behavior) control (Everitt et al., 2008; Robbins & Everitt, 1999). Evidences for DA dysfunctions due to chronic alcohol consumption are reported by studies using positron emission tomography (PET). In AD patients, low dopamine synthesis capacity (Heinz et al., 2005) and dopamine receptor availability (Heinz et al., 2004) have been observed in the VS, associated with higher levels of craving. In animals consuming alcohol daily for one year, acute ethanol consumption increases dopamine uptake in the VS possibly explaining alcohol intake as an attempt to restore dopamine levels (Siciliano et al., 2016). These might be mechanisms that clarify, why AD patients relapse despite their intention to remain abstinent (for review see Garbusow, Sebold, Beck, & Heinz, 2014).

Hypothetically, in some individuals repeated use of drugs can produce neuroadaptations in the DA reward system leading to a hypersensitive neural response to drug-associated stimuli (Robinson & Berridge, 1993, 2001, 2008; Volkow, Fowler, Wang, Baler, & Telang, 2009); the drug literally *hijacks* the reward system (Wrase et al., 2007). While nondrug-related reward anticipation in patients suffering from addiction is associated with reduced activity in the VS (Beck et al., 2009), a stronger activation of the VS for drug-related cues has indeed been observed (Beck, Wüstenberg, et al., 2012; Grüsser et al., 2004; A. Heinz et al., 2004; Wrase et al., 2007). *Cue-reactivity*, a phenomenon of stronger neural, physiological and subjective reaction for drug-related, cues has been evidenced in patients suffering from alcohol dependence by a large number of studies (e.g. Beck, Wüstenberg, et al., 2012;
Garland, Franken, & Howard, 2012; Garland, Franken, Sheetz, & Howard, 2012; Kirsch, Gruber, Ruf, Kiefer, & Kirsch, 2016; Mainz et al., 2012; Schacht, Anton, & Myrick, 2013; Sjoerds, van den Brink, Beekman, Penninx, & Veltman, 2014; Wietschorke, Lippold, Jacob, Polak, & Herrmann, 2016; Witteman et al., 2015). **Cue-reactivity** may reflect the final consequence of Pavlovian learning that occurs during chronic drug consumption, however, whether cues (e.g. drug-related stimuli) become motivationally and thus behaviorally relevant in AD patients is not shown so far in an experimental design. If demonstrated, this is an important mechanism to understand relapse behavior in addiction (for review see Belin, Jonkman, Dickinson, Robbins, & Everitt, 2009; Everitt & Robbins, 2005). This thesis is dedicated to elucidate such mechanisms.

### 2.2 Pavlovian-to-instrumental transfer

Pavlovian-to-instrumental transfer is an effect whereby a Pavlovian CS influences the motivation for ongoing instrumental responding, a mechanism that has been implicated in drug relapse (Di Chiara et al., 1999; Robbins & Everitt, 1999). Thereby, positively valued CS can enhance approach tendencies in ongoing instrumental behavior, while negatively valued CS can promote inhibition or withdrawal actions (Huys et al., 2011). PIT effects have been investigated in animals (first study by Estes, 1943), representing a long research tradition. However, the task is rather complex, especially when translating it to a human design. Therefore, several aspects need to be considered, i.a.:

First, one can distinguish between specific and general PIT effects (e.g. Corbit & Balleine, 2011). Specific PIT tasks measure the same reward during instrumental and Pavlovian conditioning, whereas during general PIT tasks, different rewards can be used during Pavlovian and instrumental conditioning. Thus a specific PIT effect occurs when Pavlovian conditioned cues influence instrumental behavior that has been associated with the same reward but not instrumental behavior that has been associated with another reward. On the contrary, a general PIT effect occurs when a Pavlovian conditioned stimulus influences instrumental behavior independent of the reward associated during conditioning.

Secondly, the reward used during Pavlovian and instrumental conditioning can differ, e.g. it can be drug- or nondrug-related. In most animal studies, for nondrug-related rewards food and in human studies food or money has been used. For drug-related PIT designs, self-administration of the drug (instrumental conditioning) or drug-associated cues (Pavlovian conditioning) have been used. A recent human study showed that stimuli of different reward types (here: money, food and social rewards) are able to influence instrumental responding and thus produce PIT effects (Lehner, Balsters, Herger, Hare, & Wenderoth, 2016).
Thirdly, appetitive (for review see Cartoni, Balleine, & Baldassarre, 2016) and aversive (e.g. Geurts, Huys, den Ouden, & Cools, 2013a) PIT effects can be investigated. While during appetitive PIT instrumental approach is enhanced by positively valued Pavlovian cues, aversive PIT is the phenomenon of inhibited instrumental approach by negatively valued Pavlovian cues.

2.2.1 Nondrug-related PIT effects

Animal addiction models using a nondrug-related PIT design showed enhanced PIT effects in addicted animals: For example, in cocaine- compared to saline-treated rats and in amphetamine-treated rats, enhanced food-related PIT effects have been observed (LeBlanc, Maidment, & Ostlund, 2013a; Ostlund, LeBlanc, Kosheleff, Wassum, & Maidment, 2014; Shiflett, Riccie, & Dimatteo, 2013). Moreover, in cocaine-treated rats this effect was insensitive to devaluation (LeBlanc, Maidment, & Ostlund, 2013b). On a neural level, the authors observed in cocaine-treated compared to drug-naive rats augmented DA release in the NAcc during a food-related PIT task (Ostlund et al., 2014). These studies suggest that repeated drug exposure can increase cue-induced control over behavior in general (not only for drug-related cues). Thus adopting the incentive salience theory, motivational processes in general may be altered during chronic drug consumption, potentially via alterations in the neural systems that mediate reward processing, motivation, and behavioral control. Moreover, drug exposure appears to bias action selection towards the automatic execution of habits and to reduce more deliberate goal-directed control (LeBlanc et al., 2013b).

On a neurobiological level, it has been shown in animals that the medial prefrontal cortex (Homayoun & Moghaddam, 2009; Keistler, Barker, & Taylor, 2015), VTA (Corbit, Janak, & Balleine, 2007; Mahler & Aston-Jones, 2012; Murschall & Hauber, 2006) and dorsolateral striatum (Corbit & Janak, 2007b) are recruited during non-drug related PIT. Moreover, lesions of the amygdala abolished PIT effects (Campese, Gonzaga, Moscarello, & LeDoux, 2015; Campese et al., 2014; Corbit & Balleine, 2005; Holland & Gallagher, 2003; Holland & Hsu, 2014; McCue, LeDoux, & Cain, 2014). According to the incentive salience theory, the NAcc has been the main target of animal PIT studies. In rats with NAcc lesions, PIT effects have been abolished (Chang, Wheeler, & Holland, 2012; Corbit & Balleine, 2011; Corbit, Muir, & Balleine, 2001; Hall, Parkinson, Connor, Dickinson, & Everitt, 2001). In line with that, DA antagonists in the NAcc reduced nondrug-related PIT effects in rats (Corbit & Balleine, 2011; Dickinson, Smith, & Mirenowicz, 2000; Laurent, Bertran-Gonzalez, Chieng, & Balleine, 2014; Lex & Hauber, 2008; Shiflett et al., 2008; Wassum, Ostlund, Balleine, & Maidment, 2011). Stimulation of DA neurotransmission in the NAcc enhanced PIT effects, suggesting
that NAcc dopamine specifically mediates the ability of reward cues to trigger wanting for their associated rewards (Everitt, Dickinson, & Robbins, 2001; Pecina & Berridge, 2013; Pecina, Schulkin, & Berridge, 2006; Wyvell & Berridge, 2000, 2001). Moreover, Wassum and colleagues observed a direct relationship between PIT and the DA release in the NAcc, which was positively correlated with the strength of the behavioral PIT effect (Wassum, Ostlund, Loewinger, & Maidment, 2013). Interestingly, the authors also found that dopamine release in the NAcc encodes a need-based motivational value, as the behavioral and neural PIT effect was reduced under satiety (Aitken et al., 2016).

It has been shown that animal PIT designs can be successfully translated to research questions involving human subjects (e.g. Huys et al., 2011; Lovibond & Colagiuri, 2013; Nadler, Delgado, & Delamater, 2011; Trick, Hogarth, & Duka, 2011). First, when simulating animal studies, comparable behavioral PIT effects were measured in humans. Food-related PIT effects persisted independent of satiation in humans (Colagiuri & Lovibond, 2015; Watson et al., 2014) and were reduced but not eliminated after extinction (Lovibond, Satkunarajah, & Colagiuri, 2015). Secondly, on a neural level, human PIT studies were able to replicate findings from animal studies. For example, dopamine depletion (Hebart & Gläscher, 2015) and dopamine antagonists (Weber et al., 2016) reduced the influence of appetitive Pavlovian cues on instrumental responses. Imaging studies using functional magnetic resonance imaging (fMRI) in healthy subjects showed increased activation during PIT in the ventrolateral putamen (Bray, Rangel, Shimojo, Balleine, & O'Doherty, 2008; Prevost, Liljeholm, Tyszka, & O'Doherty, 2012), putamen and insula (Lewis, Niznikiewicz, Delamater, & Delgado, 2013), the NAcc (Geurts et al., 2013a; Mendelsohn, Pine, & Schiller, 2014; Talmi, Seymour, Dayan, & Dolan, 2008) and amygdala (Geurts et al., 2013a; Prevost et al., 2012; Talmi et al., 2008). Thirdly, individual differences in the strength of the PIT effect were made by investigating sign- and goal tracking as well as impulsivity in humans. While it was found in animals, that sign tracking individuals are more vulnerable to addiction and relapse (Flagel, Watson, Akil, & Robinson, 2008; Tomie, Grimes, & Pohorecky, 2008) as well as being more impulsive (Flagel et al., 2009), in humans this finding was replicated by Garofalo and di Pellegrino (2015), who showed that sign-tracking human subjects were more impulsive. Secondly, they observed a stronger PIT effect in sign-tracking compared to goal-tracking subjects. As impulsivity has also been shown to be an important risk factor for initiating drug and alcohol use as well as addiction (Loree et al., 2015), this data gives rise to the question of whether impulsive choice behavior explains differences in PIT and how this may relate to addiction-associated behavioral pattern.
2.2.2 Drug-related PIT effects

In animals, it has been demonstrated that cocaine-paired cues can provoke the seeking and taking of a drug through a Pavlovian motivational process (LeBlanc, Ostlund, & Maidment, 2012). Similarly, it has been shown that ethanol-related cues facilitate instrumental performance for ethanol in animals (e.g. Krank, 2003). Moreover, this effect was generalizable to nondrug-related instrumental performance (Corbit & Janak, 2007a; Glasner, Overmier, & Balleine, 2005). These data suggest a specific (for ethanol) and general motivational property of ethanol-paired cues to induce enhanced instrumental responding (for review on drug-related PIT effects in animals see Lamb, Schindler, & Pinkston, 2016). However it is not clear whether this is due to specific design setups as Lamb, Ginsburg, and Schindler (2016) observed an ethanol-specific PIT effect only and did not replicate a general ethanol-related PIT effect. On a neural level, in line with studies on nondrug-related PIT effects, the NAcc is in focus of investigations. For example, it was found that DA antagonists in the NAcc reduced drug-related PIT effects in rats (Chaudhri, Sahuque, & Janak, 2009; Di Ciano & Everitt, 2004). Further, in cocaine addicted rats the NAcc was activated during cocaine-related PIT (Hollander & Carelli, 2007) after abstinence. These studies suggest that the DA neurotransmission within the NAcc is critical for stimulus-controlled drug seeking.

As far as we know, the only drug-related studies in humans were conducted by Hogarth and colleagues. They showed in current smokers tobacco-related specific PIT effects (Hogarth, Dickinson, Wright, Kouvaraki, & Duka, 2007; Hogarth, Maynard, & Munafo, 2015) even after devaluing nicotine. This data suggest that drug-seeking is a habitual process that is independent of the expected value of the drug (Hogarth, 2012). Moreover, the same group reported in a study with social drinkers that beer-related specific PIT effects are not affected by Pavlovian extinction, however PIT effects were abolished after instrumental extinction (Hogarth et al., 2014). The only neural study investigating drug consuming participants was conducted in social drinkers using electro encephalography (EEG; Martinovic et al., 2014). However due to methodical restraints of EEG studies, no firm conclusions can be drawn with respect to subcortical structures that are important to understand PIT effects; however, this is possible using fMRI.

To date, no study investigated neural PIT effects in patients suffering from substance dependence. Thus, to the best of our knowledge, we here describe the first studies investigating behavioral and neural PIT effects in patients suffering from alcohol dependence. The presented work shows individual differences in the strength of PIT and discusses its relevance for relapse prediction.
3. RESEARCH QUESTIONS

This thesis aims to investigate drug- and nondrug-related PIT effects in alcohol dependence. As PIT studies have so far never been conducted in an AD patient cohort, the first step was to modify and pilot a PIT paradigm to be suitable for patients suffering from alcohol dependence. Secondly, we investigated neural and behavioral PIT effects in a cohort of patients and controls. Moreover, we assessed the relevance of neural PIT correlates in the NAcc for the prediction of relapse behavior. Motivated by the large body of animal literature, we conducted a priori region of interest (ROI) analyses focusing on the NAcc (see figure 1). Finally, we were interested in how far PIT effects in AD patients can be explained by preexisting differences in traits that are associated with risk for alcohol dependence including impulsivity.

According to the overall aim of my dissertation, the following research questions arise:

1) Do patients suffering from alcohol dependence and controls differ in behavioral PIT effects regarding the influence of drug- and nondrug-related Pavlovian cues?
2) Is the NAcc functional activity during nondrug-related PIT associated with relapse status in patients suffering from alcohol dependence?
3) Is the NAcc functional activity during drug-related PIT associated with relapse status in patients suffering from alcohol dependence?
4) Does impulsive choice behavior influence the degree to which drug- and nondrug-related Pavlovian cues affect instrumental performance?

The present dissertation is publication oriented and the above specified questions are addressed in different papers. Question 1 is addressed in paper I, II, III and IV; question 2 in paper II; paper III speaks to question 3 and question 4 is addressed in paper IV.
4. SUMMARY OF RELATED PAPERS

All papers resulted from the research group investigating learning in alcohol dependence (LeAD, www.lead-studie.de, DFG FOR 1617) headed by Prof. Dr. Dr. Heinz in Berlin and Prof. Dr. Wittchen in Dresden. We here assessed AD patients and controls matched for age, gender, education and smoking status. Data was collected in Berlin, Charité Universitätsmedizin, Campus Mitte, Department of Psychiatry and Psychotherapie and in Dresden, Technical University, Carl Gustav Carus Clinic and Neuromodeling Unit from 02/2013 till 03/2015. The study was performed in accordance with the 1964 Declaration of Helsinki and approved by the Ethics Committees of Charité - Universitätsmedizin Berlin (EA1/157/11) and Technische Universität Dresden (EK 228072012). Participants received monetary compensation for study participation (10 €/h) plus a performance-dependent compensation. AD patients were recruited in in- and outpatient departments in several clinics in Berlin and Dresden and controls via online advertisement matched to patients with respect to age, gender, smoking status and education (degree).

All participants gave written informed consent and underwent two appointments: i) assessment of questionnaires, clinical diagnostic according to DSM-IV-TR (investigated using a computer-based clinical interview: Composite International Diagnostic Instrument, CIDI; Jacobi et al., 2013; Wittchen & Pfister, 1997), neuropsychological tests and impulsivity tasks, including delay discounting to measure impulsive choice behavior, and ii) a scanning session, assessing subsequently two tasks in a fixed order during echo planar imaging (EPI) sequences including the PIT paradigm (see figure 3). Moreover, all patients were followed up for one year to assess their drinking behavior for relapse prediction (with relapse defined as consumption of at least 60/48 [male/female] gram of pure alcohol per occasion). For this thesis, the behavioral PIT data (all papers) and imaging PIT data (paper II and III) as well as data of the delay discounting task (paper IV) have been used. Sample sizes and follow-up time points are differed between the analyses for this thesis, as they were conducted during the ongoing study (see figure 2).
**Pavlovian-to-Instrumental transfer paradigm.** The PIT paradigm (see figure 3) consists of four parts: A) instrumental conditioning, B) Pavlovian conditioning, C) the transfer part and D) query trials/rating. During instrumental conditioning, subjects were asked to collect ‘good’ shells by repeated button presses and to not collect ‘bad’ shells by not pressing the button to receive reward and avoid punishment. Instrumental conditioning ended after a minimum of 60 trials and 80% correct responses in the last 12 trials or, if this learning criterion was not reached, after a maximum of 120 trials. Pavlovian conditioning consisted of 80 trials and included a set of 5 fractal-like visual stimuli and auditory stimuli deterministically followed by an picture of a coin (reward: 1€, 2€), a crossfaded coin (punishment: -1€, -2€) or zero (neutral condition). During the transfer part, Pavlovian CSs (nondrug-related PIT, 90 trials) or drink stimuli (water or the favorite alcoholic drink; drug-related PIT, 72 trials) were shown in a randomized order in the background while subjects were required to perform the instrumental response in the foreground. The transfer part was conducted under nominal extinction. After that, subjects were asked to choose: i) the better out of two Pavlovian stimuli (query trials, 30 trials), and ii) the favorite stimulus out of a Pavlovian stimulus and a drink stimulus (rating, 78 trials). The assignment of stimulus and reward in part A and B was randomized between subjects, part B and C were performed inside the scanner. The dependent variable for behavioral analyses was the mean number of button presses as a measure of the strength of instrumental response behavior during the transfer part.
**Figure 3.** Pavlovian-to-instrumental transfer paradigm. A) Instrumental training: collecting a ‘good’ shell was rewarded in 80% and not collecting a ‘good’ shell punished in 80% (vice versa for ‘bad’ shells). Red arrows indicate repeated number of button presses for instrumental response (five or more button presses resulted in collecting a shell). Via trial and error, subjects learned to collect or not collect certain shells (out of a set of 6 shells). B) Pavlovian conditioning: subjects passively viewed a stimulus reward association. Stimuli comprised a tone and a fractal-like picture. USs were a picture of a coin (-2€,-1€,0€,+1€,+2€). C) Transfer: subjects were asked for the instrumental response while the background was tiled with the CS (C1) or with a drink picture (the favorite alcoholic drink or water; C2). D) Subjects were asked to choose the better out of two CS (D1) or what they liked more out of a CS and a drink stimulus (D2).

**Delay discounting task.** The delay discounting task measured impulsive decision making, one facet of impulsivity (Heinz, Beck, Meyer-Lindenberg, Sterzer, & Heinz, 2011). During each of 30 trials, subjects had to choose either a small immediate or a larger delayed reward (randomly assigned to the left or to the right of the screen). Delays were set to 3 - 365 days. Monetary rewards varied between 0.30€ - 10€. At the end of the experiment one trial was selected randomly by the computer and credited to the subjects’ compensation.
4.1 Paper I


S. Karger AG kindly permits the use of the paper for this thesis.

*equal contribution

We aimed to pilot an established human PIT task (Geurts et al., 2013a; Huys et al., 2011) to ensure its economic efficiency and practicability for AD patients, as this task was never assessed in this population. Piloting included adaption of length, instructions and complexity of the task and, in a second step, adaptations for measurement in the MRI scanner including timing and feasibility. Moreover, reliability and validity was investigated and the task was adapted to measure drug- and nondrug-related PIT effects. Therefore, during the transfer part we arranged conditions with monetarily associated CS (nondrug-related PIT, figure 3-C1) and with alcohol- versus water-related stimuli (drug-related PIT, figure 3-C2). We investigated \( n = 64 \) subjects (\( n = 32 \) recently detoxified patients suffering from alcohol dependence and \( n = 32 \) age- and gender matched controls).

**Hypotheses.** We hypothesized stronger PIT effects (i.e. a higher number of button presses in conditions with higher valued background pictures and vice versa for aversive Pavlovian cues) in patients for both, drug- and nondrug-related Pavlovian stimuli during the PIT part.

**Methods:** Repeated measures ANOVAs were computed in MATLAB (dependent variable: number of button presses; 2 factors: group [between subject] and Pavlovian background condition [within subject]) separately for nondrug- and drug-related PIT effects and for positive and negative valued background conditions. Moreover, individual PIT effects were quantified by regression slopes of number of button presses as a function of background value. For split half reliability this was repeated for the first and second half of the transfer part, respectively. Group comparisons were performed by Wilcoxon rank sum \( t \)-tests and \( \chi^2 \)-test (comparing the amount of individuals with a significant individual PIT slope between groups). For reliability, individual slopes of first and second half and for validity slopes of nondrug- and drug-related PIT parts were correlated.
Main findings. We observed significantly stronger nondrug-related PIT effects in AD patients compared to controls for negative Pavlovian backgrounds only (conditioned suppression), and a significantly higher number of patients who showed a drug-related PIT effect (with more button presses in water compared to alcohol conditions). The PIT part showed a moderate to high split half reliability of the nondrug-related PIT and a low significant split half reliability of the drug-related PIT. We observed a high construct validity reflected by a significantly high correlation between the strength of the nondrug- and drug-related PIT effect.

Conclusion. The first paper indicated that our modified PIT task is suitable to measure nondrug- and drug-related PIT effects in a cohort of AD patients and age and gender matched controls. We observed stronger PIT effects in AD patients and a high interindividual variability, which made us confident to further investigate the association of PIT and relapse behavior.
4.2 Paper II


Wiley kindly permits the use of the paper for this thesis.

*equal contribution

After successfully establishing a modified PIT paradigm to investigate nondrug- and drug-related PIT effects in patients suffering from alcohol dependence, we wanted to shed light to the neural basis for nondrug-related PIT effects and the relation to relapse behavior. We therefore assessed \( n = 31 \) AD patients and \( n = 24 \) controls matched for age, gender and social economic status using fMRI during the nondrug-related PIT part. Patients were followed up over a three month period assessing drinking behavior.

**Hypotheses.** We hypothesized stronger behavioral nondrug-related PIT effects in patients suffering from alcohol dependence compared to controls. We moreover expected a PIT-related activation within the NAcc (a priori defined ROI, see figure 1) with predictive power for relapse and drinking amount during relapse.

**Methods:** Data was analyzed in MATLAB and R Studio. The behavioral PIT effect was calculated as described in *paper I*. Imaging data were preprocessed and analyzed using general linear model approach (SMP8). On a first level, Pavlovian background conditions were modeled as separate events (+2€,+1€,0€,-1€,-2€) each parametrically modeled by number of button presses trial-by-trial for each subject. Next, we constructed a linear contrast weighting the parametric modulator by the value of each condition. For second-level analyses of the neural PIT effect we used a priori defined ROIs of the NAcc (see figure 1). Moreover, using extracted beta values of the NAcc effect, we conducted leave-one-out cross-validation to predict relapse status and drinking amount by neural activation.
Main findings. On a behavioral level, we observed a stronger nondrug-related PIT effect in AD patients compared to controls across all Pavlovian background conditions (see figure 4-A). On a neural level, we observed a stronger nondrug-related PIT NAcc activation in the patient group only. Group comparisons revealed that subsequent relapsers activated the NAcc compared to abstainers after a follow up period of three month and compared to controls (see figure 4-C). Using the leave-one-out cross validation, we showed that nondrug-related PIT NAcc activation was predictive of relapse status (via a support vector machine approach) and drinking amount (via Poisson regression) among the patients. This prediction was statistically controlled for the pure cue-reactivity effect of the CS.

Conclusion. Our data suggest that nondrug-related PIT effects are a mechanism contributing to relapse behavior. Thus, the extent to which Pavlovian cues induce motivation to conduct a behavior in general appears to be a risk factor for subsequent relapse. The motivational properties transferred to the CS were reflected in a neural response within the NAcc.
4.3 Paper III


Springer kindly permits the use of the paper for this thesis.

*equal contribution

After having investigated nondrug-related PIT effects in patients suffering from alcohol dependence in paper II, we aimed to investigate drug-related PIT effects in the same cohort.

**Hypotheses.** We hypothesized that: i) drug-related cues act as Pavlovian stimuli and thus elicit PIT effects in the NAcc (a priori defined ROI see figure 1), ii) these neural drug-related PIT effects are stronger in patients suffering from alcohol dependence compared to controls and iii) drug-related PIT effects are related to higher severity of alcohol dependence and relapse.

**Methods:** Data was analyzed in MATLAB and R Studio. For behavioral analyses, we performed linear mixed-effects analyses regressing on drink (alcohol versus water, within subject factor) and several between subject factors as group (patients versus controls; abstainers versus relapsers) and severity (low versus high) measure by a questionnaire about alcohol dependence severity. For imaging analyses, we used the general linear model approach (SMP8) after preprocessing of the functional data. On a first level, drink-related background conditions were modeled as separate events (water versus alcohol) each parametrically modeled by number of button presses trial-by-trial for each subject. Next, we constructed a linear contrast weighting the parametric modulator with +1 for alcohol and -1 for water conditions (alcohol PIT > water PIT). For second-level analyses of the neural drug-related PIT effect we again used the a priori defined NAcc ROI (see figure 1). To test for behavioral and neural group differences, we used Welch’s $t$-test.
**Main findings.** On a behavioral level, patients showed a trend-wise stronger drug-related PIT effect compared to healthy controls. Interestingly, we observed a significantly reduced approach behavior during alcohol- versus water-related conditions in subsequent abstainers, while behavior of relapsers did not differ between conditions (see figure 4-B). On a neural level, patients showed a higher activation in the NAcc compared to controls (see figure 4-D). This neural activation predicted relapse status: increased activation was observed in subsequent abstainers compared to relapsers. Behavioral and imaging results were confirmed when comparing patients with low and high dependence severity.

**Conclusion.** In subsequent abstainers and patients with lower dependence severity drug-related cues acquire inhibitory behavioral features typical of aversive stimuli. This PIT effect was associated with increased NAcc activation for the contrast alcohol PIT > water PIT in abstainers compared to relapsers. That both (behavioral and neural) effects were absent in subsequent relapsers and more severely dependent patients points to potential resilience factors regarding the attribution of salience and the behavioral impact of alcohol cues in abstainers only.
In a fourth step, we wanted to replicate our previous finding of enhanced PIT effects in AD patients compared to controls by analyzing the full sample (n = 116 patients and n = 91 controls) and a ‘true’ replication sample excluding subjects analyzed in paper I, II and III (n = 72 patients and n = 58 controls). We extended our analysis by investigating how impulsivity might influence both nondrug- and drug-related PIT effects. We therefore additionally conducted a delay discounting task to measure impulsive choice behavior, which is considered to be one facet of the multidimensional construct of impulsivity (Dick et al., 2010). Regarding delay discounting, there are numerous studies reporting steeper discounting in alcohol-dependent patients compared to healthy controls (for a meta-analysis see MacKillop et al., 2011), and within alcohol-dependent patients, relapsers tend to show higher discounting rates than abstainers (Petry, 2001).

**Hypotheses.** We expected stronger PIT effects and impulsive choice behavior in patients compared to controls. Regarding the interaction of the two constructs, we expect stronger PIT effects in high impulsive compared to low impulsive subjects.

**Methods:** Data analysis was conducted in R Studio. To analyze differences in accuracy (of the instrumental response) between patients and controls and high versus low impulsive subjects (as indicated by $k$ parameter extracted from the delay discounting task), we used binomial mixed-effects models.

**Main findings.** First, we were able to replicate that patients stronger devaluate reward over time and thus show higher impulsive choice behavior. Moreover, nondrug and drug-related PIT effects were stronger in patients compared to controls (replication sample), replicating our results from paper II and III. Next, we extracted the discount rate $k$ from the delay discounting task as a measure for choice impulsivity. We observed stronger nondrug-related and drug-related PIT effects in high impulsive patients compared to controls. In sub analyses
separately for the patient and controls, we observed stronger PIT effects in high compared to low impulsive patients whereas we found no differences regarding the PIT effect in low compared to high impulsive controls.

**Conclusion.** This data suggest that choice impulsivity is one factor that explains PIT effects, particularly in AD patients with the strongest PIT effects in high impulsive patients. Therefore, this subgroup of patients might be particularly susceptible for the influence of contextual cues in motivating behavior possibly contrary to their original explicit behavioral intention (abstinence) which explains possible mechanisms leading to relapse.
Figure 4. Summary of main results (paper II and III).

A) + C) Nondrug-related PIT effects in relapers and abstainers at FU week 12 and controls.

B) + D) Drug-related PIT effects in relapers and abstainers at FU week 6 and controls.

A) Slopes represent the strength of the group PIT effect (number of button presses as a function of the value of Pavlovian background stimuli) for the respective group. Alcohol-dependent patients (relapers and abstainers collapsed) showed a steeper slope compared to controls.

B) Abstainers compared to relapers showed a stronger reduction in instrumental behavior during alcohol compared to water conditions.

C) Relapers showed an enhanced nondrug-related PIT activation in the left NAcc compared to abstainers and controls. D) Alcohol-dependent patients (relapers and abstainers collapsed) showed an enhanced drug-related PIT activation in the right NAcc compared to controls. Crucially, the NAcc was activated in abstainers.

All results are significant with \( p < 0.05 \). Error bars represent standard errors of the mean. Note: Result plots may differ from the published versions as they were adapted for this synopsis by reason of coherence.
5. DISCUSSION

The here summarized papers show - for the first time - behavioral and neural PIT effects in patients suffering from alcohol dependence and complement a large body of animal and preclinical human literature (paper I-IV). We observed that behavioral and neural PIT effects help to predict relapse after detoxification (paper II-III). Furthermore, paper IV investigated for the first time how individual differences in impulsive choice behavior are related to PIT effects in AD patients.

5.1 Behavioral PIT effects

First, we observed stronger behavioral nondrug-related PIT effects in patients compared to controls. In paper I, we observed, a significant difference between AD patients and controls with stronger aversive PIT effects (conditioned suppression) in patients. Studies with serotonin depletion showed a selective decrease of aversive PIT effects in controls (den Ouden et al., 2015; Geurts, Huys, den Ouden, & Cools, 2013b) and our results of conditioned suppression in AD patients are in line with emerging evidence for serotonergic involvement in alcohol dependence (Heinz et al., 2011). Against our expectations, we did not observe a significant PIT effect for conditions with appetitive background stimuli in paper I. Moreover, we observed a high variability between subjects with individually significant PIT effects in 61% of the patients and 37% of the controls. A high individual variability was important for our subsequent research questions investigating relapse in a longitudinal design (paper II and III). This high individual variability helps to explain the rather small between group effects in paper I.

Our original hypothesis of stronger nondrug-related PIT effects (across negative and positive valued background CS) in AD patients was supported in paper II and IV. The PIT effect is a rather implicit phenomenon, e.g. it is not explicitly instructed to adapt instrumental behavior dependent on the value of the background stimulus. Thus, Pavlovian cues may influence instrumental responding directly (e.g. Guitart-Masip et al., 2012), or via more complex cognition-based control decision mechanisms (Huys et al., 2012). Moreover, it has been shown that the subjective value influences the strength of the PIT effect across reward types so that individuals who rated the value of reward (US) higher, showed stronger PIT effects (Lehner et al., 2016). Regarding alcohol dependence, salient stimuli previously associated with reward might thus enhance instrumental approach and avoidance behavior (Grüsser et al., 2002; Robinson & Berridge, 1993).
Secondly, we observed stronger behavioral drug-related PIT effects in patients compared to controls in *paper I, III* and *IV*. We measured reduced instrumental responding during alcohol conditions in AD patients. Crucially, when including follow up information to the analysis, this effect emerged in prospective abstainers only pointing to possible resilience factors. This suggests that after detoxification, alcohol stimuli may become aversive in patients with a prospectively positive treatment outcome. This effect mirrors previous studies showing an approach bias towards non-alcoholic pictures in patients after treatment (Spruyt et al., 2013; Wiers, Eberl, Rinck, Becker, & Lindenmeyer, 2011) particularly in early stages of abstinence (Townshend & Duka, 2007; Vollstädt-Klein, Loeber, von der Goltz, Mann, & Kiefer, 2009). Moreover, it is in line with a study in social drinkers showing abolished drug-related PIT effects after discriminative extinction (Hogarth et al., 2014). The aversive PIT effect for alcohol-related stimuli in our study may reflect an aversive conditioning process during detoxification. Here, alcohol-related thoughts may be repeatedly paired with subjectively negative states, thus creating aversive implicit alcohol-related associations (Houben, Havermans, & Wiers, 2010), possibly leading to avoidance of alcohol-related stimuli during abstinence (Vollstädt-Klein et al., 2009). The advice to avoid alcoholic drinks and to drink a lot of water during and after detoxification may lead to explicit avoidance of alcoholic drinks (Grüsser et al., 2002), resulting in a reduced approach behavior during alcohol conditions in our task. As PIT effects reflect a rather implicit measurement of the motivational strength of stimuli on ongoing behavior, this effect might be interpreted as a motivation to abstain in patients recently after detoxification by ‘upvalue’ non-alcohol or ‘devaluate’ alcohol-related stimuli. This is also in line with our data acquired at the end of the paradigm (rating, see figure 2D): patients chose neutral fractal CSs and water stimuli over alcoholic pictures. This effect differs from studies observing that in dependent patients, drug cues are appetitive (Mucha, Geier, Stuhlunger, & Mundle, 2000). However, temporal stability of this effect is not clear and how this effect might change during and after detoxification and in abstinence needs further clarification.

Thirdly, we were interested in quality criteria of our modified PIT paradigm. We observed moderate to high split half reliability for nondrug-related PIT and low for drug-related PIT (*paper I*), indicating temporally stable PIT effects conducted in nominal extinction. This suggest that PIT effects are stable and rather insensitive to extinction, as it is also reported in a nicotine replacement study by Hogarth (2012), where tobacco-related specific PIT effects have been insensitive to extinction. Moreover, we observed a significant correlation between
behavioral nondrug- and drug-related PIT effects, pointing to internal consistency of the paradigm (*paper I*). This also suggests that the nondrug-related PIT effect we measure in our paradigm is rather general, because it correlates with the general drug-related PIT effect. This is important to consider when interpreting the PIT effects: while specific PIT is sensitive to devaluation and hypothesized to be goal-directed, general PIT is rather insensitive to devaluation and thus habitual and may be relevant for learning processes in addiction (Huys, Tobler, Hasler, & Flagel, 2014).

Fourthly, we were interested in how individual differences influence the strength of PIT. Since impulsivity has been implicated as a risk factor for the development and maintenance of alcohol dependence (for review see Dick et al., 2010), we investigated if subjects showing high impulsive choice behavior exhibit stronger PIT effects compared to less impulsive subjects. Therefore, we assessed participants’ delay discounting parameter k and further included k into our analysis of PIT (*paper IV*). We found a more pronounced PIT effect in high impulsive compared to low impulsive participants. We observed the strongest PIT effects in high impulsive patients (compared to low impulsive patients and healthy controls), which was found for both nondrug-related and drug-related PIT. These data suggest that high impulsive patients tend to be influenced in their motivation to conduct a behavior by unrelated background environmental cues. This draws parallels to evidences by research on goal- and sign-tracking individuals and how this may explain individual variance in the motivational approach of Pavlovian stimuli (Flagel et al., 2009; Garofalo & di Pellegrino, 2015; Tomie et al., 2008). Thus our data point to individual differences in choice impulsivity that are linked to the degree to which incentive stimuli can evoke motivational states and drive behavior.

### 5.2 Neural PIT effects

The neural PIT studies (*paper II and III*) aimed to follow up on the neurobiological correlates of the behavioral results and to relate the behavioral and neural effects to relapse risk. The association between neural PIT effects and relapse were controlled for neural effects of the cues themselves (cue-reactivity).

First of all, relapers compared to abstainers and to healthy controls showed a stronger activity in the NAcc during nondrug-related PIT. The nondrug-related PIT NAcc effect was positively associated with relapse status and drinking amount during relapse and significantly predicted relapse (*paper II*). This data suggest that the nondrug-related PIT effect in the NAcc is a bio marker for relapse in alcohol dependence.
This finding mirrors results from animal studies investigating the NAcc and DA during PIT: individual difference in how Pavlovian cues trigger the motivation to conduct a behavior may be influenced by the level of incentive salience (wanting), which is dynamically generated by mesocorticolimbic brain systems, and influenced especially by dopaminergic neurotransmission in the NAcc (Flagel & Robinson, 2017; Pecina & Berridge, 2013; Robinson & Berridge, 1993; Wassum et al., 2013). In animals, PIT was reduced by microinjections of a dopamine D1 and D2 receptor antagonist into the NAcc, suggesting that dopamine D1 and D2 receptors in the NAcc mediate the general activating effects of Pavlovian stimuli on instrumental behavior (Lex & Hauber, 2008). Further, a NAcc lesion study in rats abolished PIT effects (Parkinson, Olmstead, Burns, Robbins, & Everitt, 1999). Moreover, in drug-addicted compared to naive- and saline-treated rats, enhanced non-drug related PIT effects (LeBlanc et al., 2013b) have been shown to be associated with enhanced NAcc-specific task encoding (Ostlund et al., 2014; Saddoris, Stamatakis, & Carelli, 2011), supporting our results of enhanced behavioral nondrug-related PIT effects in AD patients and enhanced NAcc nondrug-related PIT effects in relapsers. Our results complement human studies suggesting an important role of NAcc dysfunction for cue reactivity, craving and relapse (Heinz et al., 2004; Volkow et al., 1996). Functional alterations in subsequent relapsers thus may be triggered by dopaminergic effects in the striatum, which can help to explain associated alterations in the BOLD signal (Knutson & Gibbs, 2007).

Secondly, the drug-related PIT effect was enhanced in AD patients and related to relapse behavior: abstainers but not relapers showed an increased activation in the NAcc for alcohol-compared to water-related PIT (paper III). These results suggest that the impact of drug-related cues on instrumental behavior is encoded in the ventral striatum. Since this effect was significant in abstainers only, thus the NAcc appears to encode a salience signal (Zink, Pagnoni, Martin-Skurski, Chappelow, & Berns, 2004) relevant for PIT effects. Hence, NAcc activation could index salience and the approach driven by the salience of aversively valued alcoholic stimuli. This is conforming to aversive PIT studies in humans showing an involvement of NAcc, with stronger activation in subjects displaying stronger behavioral inhibition during aversive contextual stimuli (Geurts et al., 2013a; Lewis et al., 2013). This drug-related neural PIT effect contrasts with the result of the nondrug-related neural PIT effect (paper II), where relapers showed a stronger activation in the NAcc. Studies with AD patients, cannabis-abusers and methamphetamine-users suggest that the salience attribution effects on behavior might be abolished in severe dependence, as these studies showed a
decreased activation in NAcc in relapsers and decreased dopaminergic functioning in severe addiction (Beck, Wüstenberg, et al., 2012; Boileau et al., 2012; Volkow et al., 2014).

Prefrontal functioning may also contribute to the suppression of approach towards drug-related stimuli, interacting with NAcc-mediated promotion of approach in abstainers. However, possibly due to a small sample size, we observed no prefrontal activation during PIT in our sample of abstainers, thus further studies are needed to investigate this hypothesis.

5.3 Limitations

There are several limitations that need to be addressed: First, the models to compute behavioral PIT effects differ between the analyses. This limits comparisons of behavioral PIT effects across analyses. Moreover, cohorts of the paper overlap (e.g. behavioral data for 15 of the patients and 18 of the controls have been reported in paper I and II/III), which limits the replication of the behavioral results. However, paper II, III did not aim to replicate the behavioral PIT effects of paper I but rather to follow up the behavioral results on a neural level, to assess relapse behavior and to further explore individual differences in PIT by one measure of impulsivity. For replication purposes we conducted our analyses for behavioral nondrug- and drug-related PIT effects in AD patients compared to controls in a ‘true’ replication sample in paper IV. Secondly, the reliability of the PIT paradigm is encouraging (paper I), which, however does not replace the need for test-retest reliability. Next, neural analysis are not controlled for voxel-wise structural differences between groups. Therefore, further analyses should include regressors of no interest in general linear models for imaging analyses addressing structural differences between groups in ROIs. Fourthly, our paradigm was not able to disentangle general and specific PIT effects as we use monetary reward for instrumental conditioning only. However, results point to the nondrug-related PIT effects being general, as it was highly correlated with the general drug-related PIT effect (paper I). This might be crucial to interpret the data, as neural correlates of specific versus general PIT effects differ (Corbit & Balleine, 2005, 2011; Pielock, Lex, & Hauber, 2011; Prevost et al., 2012). Further studies should investigate specific PIT effects in patients using drug stimuli (i.e. drug reward for instrumental and Pavlovian conditioning) to complement our results on general versus specific drug-related PIT effects and to replicate animal studies in humans (LeBlanc et al., 2012). To overcome ethical concern about administering alcohol to AD patients, cover stories could be implemented (Martinovic et al., 2014). Finally, as analyses have been conducted during the ongoing study, samples of relapsers and abstainers are rather small. This rises the need of replicating our results. Moreover, we focused our imaging
analyses on the NAcc as this is the main target in animal PIT studies. However, this strategy avoids the investigation of further relevant brain areas for PIT effects like amygdala, prefrontal cortex, ventral tegmental area and orbitofrontal cortex. This should be done in future studies with bigger sample sizes.

5.4 Outlook and final conclusion

5.4.1 Clinical implications
When conducting fundamental clinical studies, it is important to consider implications for the clinical practice. Our results implicate the importance of contextual stimuli for relapse risk and prevention. Regarding our studies on nondrug-related PIT effects, the importance of contextual stimuli in motivating behavior as one mechanism relevant for alcohol dependence is emphasized. Therefore, clinical interventions should focus on reversing those Pavlovian associations, e.g. in by cognitive behavioral therapy (McHugh, Hearon, & Otto, 2010) including cognitive bias modification and training of self-awareness (for review see Copersino, 2017) or by pharmacological treatments disrupting memories associated with drug consumption (for review see Torregrossa & Taylor, 2016). Here, this rather implicit and automatic mechanism (Sebold et al., 2016) should get into the patients consciousness. Thereby, the patient learns how contextual stimuli influence the decisions, emotions and cognitions. This is the basic to establish in a next step alternative environments that enhances the probability of behavior the patient desires. Moreover, reconsolidation methods could be of interest, as we here showed, that contextual cues motivate to conduct a behavior. Thereby, blocking the reconsolidation of Pavlovian cue memories could reduce relapse in drug addiction (for review see Vousden & Milton, 2017).

Our results on drug-related PIT effects point to the importance of relearning the incentive value of alcohol-associated contextual stimuli as this seems to be a protective factor for early relapse. An intervention by Wiers et al. (2011) addresses the retraining of automatic approach tendencies towards alcohol-related cues. The authors showed, that this training improves treatment outcome in recently detoxified AD patients. The effectiveness of the training may be explained by PIT, as the trained patients relearn the value of alcohol-related cues with behavioral consequences, in this case remaining abstinent.

Further, fundamental clinical research on neural correlates helps to understand how therapeutical interventions target altered neural functions and behavior. In AD patients it has been shown that cognitive regulation of cue-induced craving, i.e. thinking of long-term negative consequences of alcohol consumption, is disrupted (Naqvi et al., 2015) and in obese
subjects cognitive regulation strategies has been shown to be associated with activation of top-down control and inhibitory network areas (Hollmann et al., 2012). Interestingly, it has been shown that obesity and substance addiction share similar alterations in neural processing of cue-reactivity (Garcia-Garcia et al., 2014), therefore similar clinical inventions are plausible. As we observed PIT-related activation in the NAcc, clinical interventions targeting top-down regulation and behavioral inhibition are needed. Taken together, fundamental clinical studies improve our understanding of underlying relapse mechanisms that help to identifying clinical meaningful subgroups for targeted interventions and implies new intervention targets (Heinz et al., 2016).

5.4.2 Directions for future studies
It would be interesting to consider mechanisms leading to the development of AD. Therefore, PIT effects could be assessed in a sample of adolescence before developing addiction and then followed up in a longitudinal design regarding to drinking behavior and PIT effects over several years. This would be important to give insights into learning processes that lead to addiction. Moreover, PIT might help to further elucidate our understanding of behavioral and neural mechanisms of action of established therapeutical interventions (e.g. attentional bias training). Moreover, combined pharmaco-fMRI studies are of interest, as cue-reactivity to subliminal cues in cocaine dependent individuals with baclofen treatment revealed reduced neural response to drug cues possibly via GABA–B induced inhibition of mesolimbic dopaminergic activation (Young et al., 2014). On this basis, it would be interesting to investigate a neural PIT task with a pharmacological challenge in AD patients.

5.4.3 Final summary and conclusion
To sum up, the here presented studies show nondrug- and drug-related PIT effects on a behavioral and neural level as well as associations with impulsive choice behavior in alcohol-dependent patients and its relation to relapse. We observed enhanced nondrug-related PIT effects in the group of alcohol-dependent patients with a neural correlate in the NAcc in relapsers compared to abstainers and healthy controls. Regarding drug-related PIT effects, we observed reduced instrumental responses during alcohol-related background with neural correlates in the NAcc in abstainers compared to relapsers. The neural nondrug-related PIT effects observed in our study are in line with the incentive sensitization theory of addiction (Robinson & Berridge, 1993). Our results suggest that chronic alcohol exposure alters dopaminergic function associated with cue effects on instrumental behavior by increasing
sensitivity to environmental stimuli and thus modulating the risk for relapse. Furthermore, our results on drug-related PIT effects demonstrate possible resilience factors, as aversive effects of alcoholic stimuli (possibly due to detoxification) seem to be an important factor to abstain. Our results give implications for therapeutical interventions and relapse prevention strategies such as alcohol approach avoidance training modulating the alcohol approach bias (Wiers et al., 2011).


Wiers, R. W., Eberl, C., Rinck, M., Becker, E. S., & Lindenmeyer, J. (2011). Retraining automatic action tendencies changes alcoholic patients' approach bias for alcohol and improves treatment outcome. *Psychological Science*, 22(4), 490-497.


