



This postprint was originally published by Elsevier as:  
Wirtz, J., Moritz, S., Gehlenborg, J., Ascone, L., & Kühn, S. (2022). **Is less more? Dismantling imaginal retraining and examining the effects of psychoeducation and embodied cognition on craving and alcohol consumption in problem drinkers. A randomized controlled trial.** *Addictive Behaviors*, 135, Article 107429.  
<https://doi.org/10.1016/j.addbeh.2022.107429>

### **Nutzungsbedingungen:**

Dieser Text wird unter einer Deposit-Lizenz (Keine Weiterverbreitung - keine Bearbeitung) zur Verfügung gestellt. Gewährt wird ein nicht exklusives, nicht übertragbares, persönliches und beschränktes Recht auf Nutzung dieses Dokuments. Dieses Dokument ist ausschließlich für den persönlichen, nicht-kommerziellen Gebrauch bestimmt. Auf sämtlichen Kopien dieses Dokuments müssen alle Urheberrechtshinweise und sonstigen Hinweise auf gesetzlichen Schutz beibehalten werden. Sie dürfen dieses Dokument nicht in irgendeiner Weise abändern, noch dürfen Sie dieses Dokument für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, aufführen, vertreiben oder anderweitig nutzen. Mit der Verwendung dieses Dokuments erkennen Sie die Nutzungsbedingungen an.

### **Terms of use:**

This document is made available under Deposit Licence (No Redistribution - no modifications). We grant a non-exclusive, nontransferable, individual and limited right to using this document. This document is solely intended for your personal, non-commercial use. All of the copies of this documents must retain all copyright information and other information regarding legal protection. You are not allowed to alter this document in any way, to copy it for public or commercial purposes, to exhibit the document in public, to perform, distribute or otherwise use the document in public. By using this particular document, you accept the above-stated conditions of use.

### **Provided by:**

Max Planck Institute for Human Development  
Library and Research Information  
[library@mpib-berlin.mpg.de](mailto:library@mpib-berlin.mpg.de)

**Is less more? Dismantling imaginal retraining and examining the effects of psychoeducation and embodied cognition on craving and alcohol consumption in problem drinkers. A randomized controlled trial.**

Janina Wirtz<sup>1</sup>, Steffen Moritz<sup>1</sup>, Josefine Gehlenborg<sup>1</sup>, Leonie Ascone<sup>1,±</sup> & Simone Kühn<sup>1,2,±</sup>

± LA and SK share senior authorship

<sup>1</sup>University Medical Center Hamburg-Eppendorf, Department of Psychiatry and Psychotherapy, Neuronal Plasticity Working Group, Martinistr. 52, 20246, Hamburg, Germany.

<sup>2</sup>Max Planck Institute for Human Development, Lise Meitner Group for Environmental Neuroscience, Lentzeallee 94, 14195 Berlin, Germany

Corresponding author: Janina Wirtz,  
Martinistr. 52, 20251 Hamburg, Germany;  
J.wirtz@uke.de, +49 40 7410 24115

Steffen Moritz: Moritz@uke.de

Josefine Gehlenborg: J.gehlenborg@uke.de

Leonie Ascone: L.ascone-michelis@uke.de

Simone Kühn: Kuehn@mpib-berlin.mpg.de

**Running head:** Dismantling imaginal retraining

**Number of Tables:** 6 (See supplementary materials)

**Number of Figures:** 4 (See supplementary materials)

**Word count:** 5107 words

**Clinical trial registration:** AsPredicted #37851: 'Segregating the Mechanisms of Imaginal Retraining'

### **Declaration of conflicting interests and funding**

We declare no conflicting interest.

J.W. has no conflict of interest to declare.

L.A. has no conflict of interest to declare.

J.G. received research grants from the German statutory accident insurance (DGUV).

S.M. received research grants from the DGUV, the Gauselmann Group (vending machine industry), and from the German social organization Aktion Mensch e.V. in the last three years. S.M. developed the imaginal retraining technique together with S.K.

S.K. received research grants from the European Union (ERC-2016-StG-Self-Control-677804), the Max Planck Society and the German Science Foundation (TRR 169/C8, SFB 936/C7) S.K. developed the imaginal retraining technique together with S.M.

The study was funded by the European Union (ERC-2016-StG-Self-Control-677804) as well as overhead sources.

### **Data availability statement**

We hereby declare that our data and syntaxes (including a documentation of all analyses that were undertaken) are available upon request.

### **Acknowledgments**

### **Ethics and participant consent**

The study was approved by a local ethics consortium (LPEK-0019) prior to study onset. The study adhered to the declaration of Helsinki. All participants consented to participate in the study.

### **Clinical trial registration**

The trial was pre-registered at AsPredicted #37851: 'Segregating the Mechanisms of Imaginal Retraining' prior to any analyses being undertaken <http://aspredicted.org/blind.php?x=my6ee4>

## **Abstract**

### **Background**

Imaginal retraining, a type of approach-bias-modification performed in one's own imagination, is effective in reducing alcohol craving in problem drinkers. This study aims to segregate three main components of the original technique and compare their effectiveness in reducing craving and alcohol consumption.

### **Methods**

The study was conducted online. Interested individuals wishing to reduce their alcohol consumption underwent a baseline assessment (T0), whereby a total of  $n = 200$  included participants were randomly assigned to one of four study arms. There were three active groups receiving different training manuals: psychoeducation + advice only (PE), PE + a simplified motor in-sensu AAT instruction (AAT0), and the 'enriched' original imaginal retraining technique (including the components PE + enhanced movements + an embodied cognition/emotion induction instruction) (AAT+), vs. a wait-list control (WLC). Participants in the three active arms were instructed to train twice a day (10 mins. each) over the course of four weeks, whereafter a post-test was conducted (T1). A follow-up took place one month later (T2). Repeated measures ANOVAs were computed to test for differential changes of the four groups over time (group\*time interaction effects) concerning craving and alcohol consumption. Both per-protocol and intention-to-treat analyses (PPA, ITTA) were conducted.

### **Results**

In both PPA and ITTA, significant group\*time interaction effects were identified, whereby false-discovery-rate corrected post-hoc tests revealed that AAT0 most consistently reduced craving with stability towards follow-up. Results concerning alcohol consumption (g ethanol) were inconsistent.

### **Conclusion**

## DISMANTLING IMAGINAL RETRAINING

This study demonstrates that a simplified version of the imaginal retraining technique yields consistent results concerning craving reduction. This may hint towards the central importance of the core element of retraining: a movement of pushing (imagined) alcoholic beverages away. Possibly, more complex forms of imaginal retraining need further instruction than a written manual, such as instructive videos.

**Keywords:** *imaginal retraining, in-sensu, training effectiveness, craving, alcohol consumption, addiction.*

## INTRODUCTION

Globally, the consumption of alcohol is related to 3 million deaths per year, accounts for 5.1% of the global burden of disease and is the leading risk factor for premature mortality and disability in people between 15 to 49 years of age (World Health Organization; 2018). Even in sub-clinical cases, alcohol consumption rates represent a serious health risk with over 3 billion people worldwide recorded drinking alcohol in the year 2016 (World Health Organization; 2018). Although people are usually aware of the negative consequences and might even desire to stop drinking, they are often unsuccessful in quitting or reducing their consumption (Stacy & Wiers, 2010).

### **Approach-/ avoidance tendencies**

According to the dual-process model, two systems are engaged when processing information such as evaluating a craved object: the impulsive system, which operates rather quickly, automatically, and out of conscious control, and the reflective system, functioning in a deliberate and controlled, yet slower manner (Brockmeyer et al., 2015; Wiers et al., 2013a, Wiers et al., 2013b). In addictive behaviors, there appears to be an imbalance favoring the automatic, impulsive processes over the controlled ones, leading to automated approach tendencies towards craved stimuli (Watson et al., 2012; Wiers et al., 2013b; Fridland & Wiers, 2017). The tendency to approach craved objects over neutral ones, known as the approach bias, has been demonstrated for a variety of addictive behaviors (e.g., alcohol and cigarettes) by the Approach Avoidance Task (AAT; Wiers et al., 2013a; Wiers et al., 2013b). In its original form, the AAT displays pictures of craved (i.e., alcohol) and neutral objects on a computer screen. In the diagnostic version of the AAT, individuals are asked to implicitly react to the format, such as horizontally vs. vertically displayed pictures, or colored frames (e.g., yellow vs. blue) around the neutral and unhealthy/addictive stimuli and to either pull or push the pictures towards or away from oneself with a joystick, depending upon the presented format. The approach bias can be calculated in different ways, such as by comparing the

standardized reaction times for approaching (pulling) unhealthy vs. neutral stimuli or subtracting standardized pull- from push- reaction times. In addictive behaviors, the reaction times are often found to be faster while approaching craved objects (Wiers et al., 2011; Wiers et al., 2010; Kakoschke et al., 2017). In one of the approach bias modification (ApBM) versions of the AAT, the instructions are explicit, in that critical stimuli such as alcoholic beverages are always to be rejected, i.e. pushed away from oneself as quickly as possible with a joystick, while neutral stimuli are to be approached by pulling them towards oneself. ApBM has proven to be effective across multiple addictive behaviors, such as excessive alcohol consumption, smoking or unhealthy eating (Fridland & Wiers, 2017., Wiers et al., 2011, Wiers et al., 2010; Wiers et al., 2015; Kakoschke et al., 2017).

The AAT is partially derived from the concept of embodied cognition, i.e., that movements, cognitions, and emotions are interrelated (Dijkstra & Post, 2015; Fridland & Wiers, 2017) and that the connections between movement and inner states such as thoughts and emotions/ perceptions are automatic (Dijksterhuis & Bargh, 2001) (e.g., adopting a hunched posture when feeling sad, or pushing something away from oneself/ subconsciously distancing oneself in dislike). By pushing or pulling a joystick, the extension or retraction of the arm demonstrates this concept in either rejecting or approaching something physically. Sensorimotor characteristics, such as body postures, play a decisive role in mental processes (Pecher & Zwaan, 2005; Barsalou, 2010; Vallet et al., 2016) and it has been demonstrated that when seeing a stimulus or envisioning it, the same sensorimotor units are involved (Vallet et al., 2016; Vermeulen et al., 2008; Riou et al., 2011; Rey et al., 2014).

### **Imaginal retraining**

The AAT has proven to be effective in retraining approach biases (Fridland & Wiers, 2017; Wiers et al., 2011; Wiers et al., 2010; Wiers et al., 2015), but there are barriers to its implementation: a computer is required for training and the stimuli are almost always



confined to those presented on the screen. Although customizing the computerized AAT to include individualized stimuli is possible in theory and could constitute a promising treatment addition for substance use disorders, it would be a rather time-consuming approach and difficult to implement in diagnostic trials. Additionally, due to its repetitive nature, the computer AAT is a rather tedious task. Imaginal retraining (IR; Moritz et al., 2019a; Moritz et al., 2020; Moritz et al., 2019b) is a type of ApBM which is performed using one's own imagination. This non-computerized technique asks participants to imagine their individually craved stimuli (e.g., favorite alcoholic beverage) at their common place of consumption (e.g., bar, sofa, table), followed by a negative emotion induction (in the form of a crooked posture, frowning in disgust, contemplating negative thoughts such as being sick following alcohol consumption) introducing the concept of embodiment to a much larger extent than the computer AAT. Subsequently, instead of pushing away a joystick like in the computerized AAT, an actual 'pushing movement' is carried out, while simultaneously imagining throwing the craved object away before the 'inner eye'. Conversely, non-addictive healthy drinks (e.g., water, tea) should be imagined while standing upright, slightly looking up, then lifting the object up, while pulling the arm towards the mouth. The non-addictive healthy drinks should be consumed in imagination in an exaggerated way, coupled with positive feelings (e.g., activated by thinking about snuggling a pet), to increase the feeling of appreciation. While the original computer AAT also features embodiment (i.e., pushing and pulling a joystick), imaginal retraining goes further in amplifying the effects of embodied cognition by adding facial expressions such as frowning and smiling, specific motions or postures (i.e., crooked posture vs. standing up right), and thoughts or imagination, intended to evoke negative vs. positive emotions. This hypothetically fortifies the re-scripting and encoding of associations (i.e., alcohol-negative-avoidance vs. healthy beverages-positive-approach) and hence stimulus-response-patterns. Advantages over the original AAT include easier implementation and integration into daily life as no equipment such as computers or joysticks are needed.

Additionally, personalized, preferential stimuli (such as most strongly craved alcoholic beverages) can be imagined, ensuring a customized training directed at reducing individualized approach biases. Previous research examining the effectiveness of imaginal retraining by our research group have shown promising results in samples of problem drinkers (Moritz et al., 2019a), smokers (Moritz et al., 2020), and overweight women (Moritz et al., 2019b).

Up to date, little is known about how imaginal retraining works. Two pilot studies have examined a selection of mechanisms of action (e.g., imaginary arm movement vs. imaginary and actual arm movement combined) in imaginal retraining in a one-time application. One study conducted with a sample of smokers (Moritz et al., 2021) indicated that the execution of the actual arm movement could be dispensable, while another study including a sample of overweight and obese women (Wirtz et al., 2021) suggested that the combination of an imagined and actual arm movement results in a reduction of craving. However, as both studies were restricted in that they only addressed the effects on craving after a one-time application of the imaginal retraining technique, it remains of interest which elements of the original retraining in-sensu are most effective, which may be dispensable, or perhaps even detrimental to its efficacy, in more realistic, long-term training schemes (i.e. randomized-controlled trial; RCT). Furthermore, other components including extensive embodied cognitions besides the arm movement, such as positive and negative emotion indication, alternating body postures and specific thoughts, or psychoeducation as separate interventions (all components of the original technique) were not evaluated in the preceding studies and their importance for effectiveness thus remain unknown.

### **The present study**

As a dismantling study on imaginal retraining has not yet been examined in the scope of an RCT involving repeated training, while segregating different components of the

technique, the present study will focus on this research question. The original imaginal retraining manual as proposed by Moritz and colleagues (Moritz et al. 2019a; Moritz et al., 2020; Moritz et al., 2019b) includes several mechanisms that may alone or in combination contribute to the reduction of craving and drinking amount. These include: 1) [PE] a psycho-educative component, which also contains tips to reduce the amount of drinking and reduce the urge to drink, 2) [AAT0] the original AAT technique itself, as introduced by Wiers and colleagues (Wiers et al., 2013a; Wiers et al., 2013b; Wiers et al., 2011; Wiers et al., 2010), including the movement of pushing alcohol away (vs. pulling non-alcoholic beverages towards oneself), and 3) [AAT+] an emotion induction component, which uses both embodied cognition (e.g., posture of disgust, frowning, exaggerated movements such as throwing the alcoholic drink away aggressively), and emotion induction (e.g., creating preceding negative imaginal scenarios and thoughts) to enhance negative associations in relation to alcohol.

To examine the underlying mechanisms of imaginal retraining, the original manual was dismantled into three different sub-manuals containing the previously mentioned components (for further details, see *Interventions* section). A more detailed overview can be found in Table 1 in the *Supplementary materials*. By the segregation of these three main components, we were hoping to further unravel and differentiate the mechanisms underlying the effectiveness concerning reduced craving for alcohol, as previously reported following an imaginal retraining intervention (Moritz et al., 2019a). We set out to elucidate whether the craving-reducing effects of imaginal retraining are due to the mechanisms of enhanced embodied cognition and emotion induction - or can be achieved by the arm movement alone - or even simple psychoeducation. While other imaginal retraining studies mainly focused on effects on craving and rather vague reports on alcohol consumption were assessed (e.g. “did you consume much more, more, the same, less, or much less alcohol over the course of the last month?”), in this study, in addition to the main outcome variable craving, we took a closer look at imaginal retraining’s effects on alcohol consumption (grams of ethanol, as computed

## DISMANTLING IMAGINAL RETRAINING

from indicated standard drinks consumption) as a secondary outcome variable. In this study we compared three manualized conditions, hypothesizing that both manuals containing imaginal retraining (AAT0, AAT+) would have a stronger effect on alcohol craving- (primary outcome) and consumption reduction (secondary outcome) compared to the manual containing psychoeducation only (PE), and a wait-list control (WLC). Additionally, it was expected that a month-long application of the ‘enriched’ imaginal retraining would be superior to the simplified retraining in reducing craving and alcohol consumption, due to the added element of enhanced embodied cognition/ emotion induction.

## METHODS

### Recruitment, Design, and Study Procedure

The study was approved by a local ethics board prior to study onset (LPEK-0019). Recruitment mainly took place via Google AdWords, flyers and student e-mail distributors between June 2019 and April 2020. The RCT was implemented online using the survey platform Questback/ UniPark® (no IP addresses were stored). To ensure participation was as anonymous as possible, participants were instructed to use e-mail addresses not disclosing their names to conceal their identity (for sending the manuals, reminders for participation in post-test and follow-up, handling study compensation, or questions). Inclusion criteria for the online study entailed self-reported problematic drinking behavior, the wish to reduce the amount of drinking for different reasons (e.g., alcohol consumption having led to problems in social, traffic, or work-related contexts), and age between 18-75 years. No formal diagnosis of alcohol use disorder (AUD) was required for participation; but AUD also did not constitute an exclusion criterion either. Exclusion criteria were acute suicidality, assessed with BDI-II item 9 (rating  $\geq 2$ ), lifetime psychosis (self-reported diagnosis) and previous participation in any of the forerunner trials on the imaginal retraining method. In addition, participants who indicated drinking less than twice a month were filtered out in the survey and those indicating an ethanol intake of 0 (in g) were excluded post-hoc (see *Measures* section for details).

At the beginning of the baseline assessment [T0], informed consent was obtained by participants declaring to have understood and agreed to the data protection regulations and provided study information, being aware that participation was voluntary with the possibility of termination at any given time without need for explanation and agreeing to create a study e-mail address that does not contain personal information. Subsequently, socio-demographics, including age, sex, family status, work status, education level, previous and current therapeutic (incl. medication intake) status, lifetime and current psychiatric diagnoses, experience with psychotherapy and self-help interventions in general were assessed. In addition, alcohol- and

other addiction-related data was acquired, including existing substance use disorders, specification of alcohol/ drug intake, age of onset (regular drinking), and prior treatments. Lastly, a range of psychological and alcohol related questionnaires were administered (see Measures section), and afterwards participants were randomly assigned by the survey platform Questback/ Unipark® to one of four arms (see Interventions section for details).

There were three points of assessment for all groups: baseline [T0], a post-test directly after the training period of four weeks [T1] and at an additional four-week follow-up [T2]. Participants were reminded up to three times via e-mail to take part in post-test and follow-up. After the post-test assessment, participants were informed that they were welcome to keep on practicing their assigned intervention if they liked to. To be able to control for potential differences in training time between the active groups, participants were additionally asked about their average training time at the post- and follow-up assessment. After completion of a minimum of two assessment points, participants received a 20€-Amazon-voucher, sent to the anonymous e-mail-address they created at the beginning of the study. Participants in the WLC group received the AAT+ manual following completion of the study.

As medium to large effects were to be expected (Moritz et al., 2019a), a sample size of 30-40 participants per condition had been determined with the software G\*Power (Erdfelder et al., 1996) ( $\alpha = .05$ ;  $\beta = .80$ ). In total, a sample size of 120-160 participants was needed. Subjects were recruited online until 150 individuals with at least two assessment points (pre-, post- and/or follow up) had participated. Follow-up assessments were completed in August 2020.

### **Interventions**

The study comprised three interventional groups, listed here in ascending order according to the degree of complexity (number of hypothetical components): **1)** psychoeducation + advice only = (PE), **2)** (PE) + a simplified motor in-sensu AAT instruction = (AAT0), and **3)** ‘enriched’ original imaginal retraining technique (including the components

(PE) + a motor (AAT0) + an embodied cognition/ emotion induction instruction) = (AAT+). These were contrasted against a wait-list control [WLC] group. For a comparative overview detailing the different interventions in their modular structure, content, and varying degree of complexity, see Table 1. For AAT0 and AAT+ the instruction was to ideally practice twice a day, for about 10 minutes each time over the course of four weeks (until post-assessment; afterwards ad libitum).

## Measures

### *Alcohol-related problems at baseline*

The Alcohol Use Disorder Identification Test (AUDIT; Babor et al; 2001), issued by the WHO, consists of ten items (e.g., “*How often during the last year have you found that you were not able to stop drinking once you had started?*”) and measures the severity of alcohol-related problems on a 5-point-scale (“*never*”, “*less than monthly*”, “*monthly*”, “*weekly*” and “*daily or almost daily*”) and two items on a 3-point scale (“*no*”, “*yes, but not in the last year*”, “*yes, during the last year*”). Scores are weighed between 0 and 4 points each and summed up to calculate the total score of the individual items. Reliability in the present baseline sample ( $n = 200$ ) was good (Cronbach’s  $\alpha = .85$ ). The possible sum score ranges from 0-40, whereby scores from 0-7 indicates a low-risk, 8-14 hazardous drinking, and  $\geq 16$  likelihood of alcohol-dependence. As the AUDIT focuses on the past year, it was inquired at baseline [T0] only, while at the post-assessments [T1 and T2] participants were asked concerning the outcomes *craving* and alcohol consumption [g of ethanol; see below].

The Alcohol Dependence Scale (ADS; Skinner & Horn, 1984), which was developed to assess both the presence and severity of AUD, comprises 25 items (e.g. “*Following drinking, have you lost consciousness in the past 12 months?*”) to rate the presence of AUD-related problems and symptoms over the past year either dichotomously (i.e., *yes* vs. *no*), or on a 3- or 4-point scale (e.g., quantifying different intensities). The total sum scores can range between 0 and 30. Values  $\geq 20$  indicate a potential AUD. Reliability was good in the present

sample (Cronbach's  $\alpha = .82$ ). Both AUDIT and ADS were assessed, in order to be able to cross-check results concerning baseline severity of alcohol-related problems.

### *Psychopathology assessments*

Additional psychopathology was assessed, as symptoms relating to depression, stress, and anxiety are known to have strong inter-relations with alcohol consumption. The prognosis for patients with AUD and depression is worse compared to the prognosis of patients without depression (McHugh & Weiss, 2019). Thus, checking for baseline differences in the levels of (subclinical) affective symptoms in the present sample across groups seemed highly relevant.

The German version of the Depression Anxiety and Stress scale-21 (DASS-21) is a 21-item measure, whereby each dimension is measured with 7 items on a 4-point Likert scale. The dimensions of depression, anxiety, and stress in the present study were assessed referring to the past 4 weeks (e.g., “*I found it difficult to relax*”). Differential cut-offs are suggested for depression (0-9: *normal*, 10-13: *mild*, 14-20: *moderate*, 21-27: *severe*, 28+: *extremely severe*), anxiety (0-7: *normal*, 8-9: *mild*, 10-14: *moderate*, 15-19: *severe*, 20+: *extremely severe*), and stress (0-14: *normal*, 15-18 *mild*, 19-25: *moderate*, 26-33: *severe*, 34+: *extremely severe*). Cronbach's  $\alpha$  for all scales in the present sample was good (depression:  $\alpha = .83$ , anxiety:  $\alpha = .81$ , stress:  $\alpha = .83$ ).

### *Impulsiveness*

On a meta-analytical level, impulsivity has been suggested to be central to both the emergence and maintenance of addictive behavior (Lee et al., 2019). The Barratt Impulsiveness Scale (BIS-11; Patton et al., 1995) assesses the trait of impulsivity as self-report. Its' 15-item version has shown acceptable psychometric properties (Spinella, 2007; Meule, 2011). The items are assessed on 4-point Likert scales (1 = *rarely/ never*; 2 = *occasionally*, 3 = *often*, 4 = *almost always/ always*), and 5 items each reflect one dimension of impulsiveness: attentional (e.g., item: “*I am restless at lectures or talky*”), motor- (e.g., item: “*I act on*



*impulse*”) and non-planning (e.g., item [reverse coded]: “*I plan tasks carefully*”). For the present study, the total score was used, with the total scale having a good internal consistency (Cronbach’s  $\alpha = .83$ ) in the present sample.

#### *Primary Outcome*

Craving was assessed concerning frequency (“*How often did you crave an alcoholic drink during the last 7 days?*”; 1 = *never* to 11 = *all of the time*), peak intensity and average intensity (“*How intense was the strongest [average] craving for alcohol during the last 7 days?*”; 1 = *non-existent* to 11 = *extremely intense/ strong*) on visual analogue scales. The item score (sum/ 3) was used as the outcome in the present study. Internal consistency was good (Cronbach’s  $\alpha$  at baseline: .89).

#### *Secondary Outcome*

An adapted version of the beverage intake questionnaire (BEVQ; Hedrick et al., 2010) was used, whereby the frequency of intake of seven standard alcoholic drinks (i.e.: beer 330ml, wine 150ml, sparkling wine 150ml, spirits 20ml, liqueur 40ml, long drinks 200ml, and alcohol-containing cocktails 200ml) during the last four weeks was assessed. Ratings were 0 = *never*, 1 = *once per week*, 2 = *2-3 x per week*, 3 = *4-6 times x week*, 4 = *once per day*, 5 = *2-3 x per day*, 6 = *4x or more per day*. Based on the indicated frequencies, first the total number of drinks per category (e.g., total number of beers) within one month, and based on this, the total amount consumed in ml per type of drink, was computed. In a second step, ethanol in grams was computed for the total amount (in ml) consumed, separately for each drink, by multiplying total consumed ml of the respective drink with the %vol (0.05 for beer, 0.12 for wine and sparkling wine, 0.38 for spirits, 0.30 for liqueur, 0.20 for long drinks and cocktails), thus obtaining alcohol consumed in ml. This number was then multiplied by the weight of ethanol, which is 0.8g (per ml of alcohol). In a final step, the ethanol values per drink were summed up to form a total score of pure alcohol consumption in gram ethanol within the last four weeks.

*Amount of training*

Amount of training was assessed to control for ‘participant engagement’ as potential confounder, by means of participants indicating their average training time per day, ranging from *< 5 minutes* to *> 30 minutes* on an 8-step scale, at post-test and follow-up.

**Statistical Analyses***Baseline sample characteristics and group differences at baseline*

In a first step, two separate baseline analyses were performed; once for a Per-Protocol-Analysis sample with complete data on all three assessment points (PPA; Table 2), and once for the entire baseline sample, as originally assigned per group, for the Intention-to-Treat-Analysis (ITTA; Table 3; see *Missing data approach and multiple imputation for the intention to treat analysis* section for details). It was assessed whether the groups differed in any of the baseline characteristics data, i.e., age, sex distribution, school education, AUDIT and ADS (severity of alcohol-related problems), DASS (depression, anxiety, stress), BIS (impulsiveness), the presence of current mental disorder (self-reported in general; AUD according to AUDIT and ADS), and the outcome variables (craving, alcohol consumption in g ethanol). Non-parametric or robust tests were chosen based on checking for normality of the data distribution in the sample with a rule-of-thumb criterion (skew and kurtosis between -1 to 1). Detecting differences at baseline between groups is key to identify potential sources of bias (covariates) that need to be considered for the main analyses, hence no *p*-level adjustments were made (as opposed to the post-hoc tests for the main analyses; see subsection *Post-hoc tests*).

*Per protocol analysis: mixed ANOVA approach to assess intervention effects on craving and alcohol consumption*

All analyses were conducted in SPSS27. In a first step, only fully compliant subjects with complete data were analyzed (Per-Protocol-Analysis; PPA,  $n = 126$ ). If normality (see previous section) was violated, and/or Boxplots exhibited extreme outliers (as defined in SPSS,

which employs the Tukey method: lower bound = 1<sup>st</sup> quartile – 3 \* IQR; upper bound = 3<sup>rd</sup> quartile + 3\*IQR), data transformation approaches were applied first to compensate for these deviations. If data transformation could not solve the problem, a suited non-parametric test (i.e., median-, or rank-based) instead of the intended parametric repeated measures ANOVA (rm ANOVA) was used.

Two separate rm ANOVAs were conducted for the main (craving) and secondary (alcohol consumption in g ethanol) outcome, whereby assumptions were carefully checked. First, within-subject effects were examined to test for global trends over time (time effect). Between-subject effects (i.e., group) were examined to test for global differences between the groups in the (averaged) outcome across assessment points. Of main interest were significant group-dependent, differential changes over time (inner-subject group\*time interaction effects). Effect sizes for main effects (i.e., group, time) and the interaction (group\*time), were interpreted as  $\eta^2_{partial} = 0.01$  *small*;  $\eta^2_{partial} = 0.06$  *medium*;  $\eta^2_{partial} = 0.14$  *large effect*.

#### *Handling of significant baseline differences*

In case of any of the baseline characteristics differing significantly, rm ANCOVA was conducted, with the respective baseline characteristic variable as covariate. In case of significant baseline differences in the outcome variables, multivariate analysis of covariance (MANCOVA) was conducted. The MANCOVA was run including the significantly differing baseline characteristics variable(s) plus the baseline values of the respective outcome as covariates, group as fixed factor, and post-test and follow-up data for craving/ alcohol consumption in g ethanol as outcomes.

#### *Post-hoc-tests*

For the main analyses (rm ANOVAs), within group paired (dependent, across all timepoints) and between-group (independent, within all timepoints) post-hoc t-tests were only run if the respective time and group\*time interaction effect was significant. Hypothetically, between-group post-hoc tests would only need to be run in the case of significant between

subject effects in the previous main analysis step. However, it was reasoned that including between-group contrasts, applying an overall approach to mitigate alpha-error cumulation (see below), would provide a more detailed and conservative picture of findings. Cohen's  $d$  (Cohen; 1988) was estimated, using the rule of thumb:  $\geq 0.2$  small effect,  $\geq 0.5$  medium effect,  $\geq 0.8$  large effect, to facilitate interpretation. Type I error accumulation (or alpha inflation) is a common problem in repeated testing of the same data. Since on the other hand the Bonferroni approach has been heavily criticized for being overly conservative, leading to inflated type II error (Perneger, 1998), a correction for multiple testing was implemented applying the Benjamini and Hochberg (1995) false discovery rate (FDR) method. Thereby, all  $p$ -levels for all post-hoc-tests that were run on the outcomes were entered into a new data file, and an SPSS script to implement the FDR method was run over this data, to distinguish false- from true-positive results (script available at: <https://www.ibm.com/support/pages/does-spss-statistics-offer-multiple-comparisons-using-benjamini-hochberg-method-control-false-discovery-rate>).

*Missing data approach and multiple imputation for the intention to treat analysis (ITTA)*

Missingness (%) was compared across groups. In addition, non-completers and completers were compared at baseline concerning the above-mentioned variables, to be able to detect systematic differences in compliant vs. non-compliant cases, which is relevant for inferring generalizability of results. A missing value analysis was carried out in SPSS27, whereby following a general rule of thumb, data on the dependent variable was not to be imputed if it was missing with  $> 40\%$  (Jakobsen et al., 2009). If this criterion was fulfilled, missing outcome data was imputed for all participants who had been assigned to any the different study arms using multiple imputation (10 imputations, method: automatic selection based on empirical data option). Thereby, all data (i.e., for all baseline characteristics, outcome variables at baseline, intervention assignment) as used in the final analyses, were entered into the imputation model to obtain the estimates for the missing outcome data at post-test and

follow-up. Main analyses were repeated on an aggregated imputed dataset (containing averaged imputed values per participant).

### 3. RESULTS

#### Sample characteristics and group differences at baseline

Out of  $n = 210$  baseline participants, 10 did not fulfil the inclusion criterion (i.e., baseline ethanol  $g = 0$ ) and were excluded from the main analyses. This resulted in a new total baseline sample of  $n = 200$ . For baseline descriptive data, including group comparisons for the PPA and ITTA, on all variables at baseline, see Tables 2 & 3. For dropout within each respective group, and subsequent samples for PPA and ITTA, see the CONSORT flowchart. For inferential statistics concerning dropouts across the groups, see next section.

#### Missing data analysis

We distinguished two ‘types’ of participants in our data ( $N = 200$ ): completers (as also included for PPA, with full data available on all three assessments;  $n = 126$ ), and non-completers ( $n = 74$ ). It was found that in AAT+, descriptively the highest rates of non-compliance occurred ( $n = 28/ 59$  initially enrolled participants; 47.5%), followed by WLC ( $n = 19/ 49$ ; 38.8%), AAT0 ( $n = 16/ 46$ ; 34.8%) and PE ( $n = 11/ 46$ ; 23.9%), yielding  $X^2(3, 200) = 6.31$ ,  $p = .097$ . Please see Table 4 in the supplementary materials, for descriptive data and comparisons between completers vs. non-completers for all relevant baseline variables and outcomes at baseline. The two groups differed significantly concerning age, AUDIT mean sum score,  $AUDIT \geq 16$ , ADS mean, total craving mean, and alcohol consumption in g ethanol, with all scores being higher for non-completers. None of the outcomes at T1 or T2 had more than 40% missingness (T1 missings on craving and ethanol: 29.5%, T2 missings on craving and ethanol: 32.5%), thus multiple imputation, as described above, was performed for ITTA.

#### Main Results

*Per-protocol-analysis 1 repeated measures ANOVA: intervention effects on craving*

There were no (extreme) outliers for craving, skew and kurtosis were all in an acceptable range at all assessment points in the total sample. Since there were no significant baseline differences in the PPA sample (see Table 2), no control for covariates (rm ANCOVA, or MANCOVA) was necessary. The rm AVOVA inner-subject effects analysis revealed a significant time effect [ $F(1.88, 229.59) = 14.00, p < .001, \eta^2_{\text{partial}} = .103$ ], but no significant group effect was found [ $F(3, 122) = 0.45, p = .719, \eta^2_{\text{partial}} = .011$ ]. The omnibus effect for the inner-subject interaction group\*time was significant [ $F(5.65, 229.59) = 2.35, p = .035, \eta^2_{\text{partial}} = .055$ ]. Descriptive outcome data can be found in Table 5. A bar diagram, including denoted post-hoc tests that survived the FDR procedure (for details see section *Post-hoc tests*) can be found in Figure 1.

*Per-protocol-analysis 2 repeated measures ANOVA: intervention effects on alcohol consumption (g ethanol)*

There were several extreme outliers across all three assessment points (T0:  $n = 1$ , T2:  $n = 12$ , T3:  $n = 8$ ). Skew and kurtosis strongly suggested non-normality of the distributions. As some individuals had zero values on g ethanol at post-test or follow-up, log10 transformation was not possible in these cases. Thus, first we performed a linear transformation (+1) on all assessment points. Afterwards, log10 transformation was performed. Albeit the pre-defined skew and kurtosis criterion was still slightly failed (skew<sub>T0</sub> = 0.05, kurtosis<sub>T0</sub> = -0.76; skew<sub>T1</sub> = -1.39, kurtosis<sub>T1</sub> = 1.44; skew<sub>T2</sub> = -1.14, kurtosis<sub>T2</sub> = 0.57), and, due to the participants that changed to zero consumption, 17 extremes at post-test (T1:  $n = 3$  in PE,  $n = 5$  in AAT0,  $n = 6$  in AAT+,  $n = 3$  in WLC) and 18 non-extreme outliers at follow-up (T2; 7 in PE, 5 in AAT0, 2 in AAT+, 4 in WLC), (lower bound = 1<sup>st</sup> quartile - 1.5 \* IQR; upper bound = 3<sup>rd</sup> quartile + 1.5\*IQR) were present, we reasoned that excluding those cases would be inappropriate, as they likely constitute realistic cases. Since there were no significant baseline differences in the PPA sample (see Table 2), no control for covariates (rm ANCOVA, or MANCOVA) was necessary. The rm AVOVA within-subjects effects analysis revealed a significant time effect [ $F(2, 244) =$

28.28,  $p < .001$ ,  $\eta^2_{\text{partial}} = .188$ ]. The between-group effect was non-significant [ $F(3, 122) = 1.16$ ,  $p = .328$ ,  $\eta^2_{\text{partial}} = .028$ ]. The group\*time within-subject effect was significant [ $F(6, 244) = 2.78$ ,  $p = .013$ ,  $\eta^2_{\text{partial}} = .064$ ]. Descriptive data can be found in Table 5. A bar diagram, including denoted post-hoc tests that survived the FDR procedure can be found in Figure 2.

*Intention-to-treat analysis 1 repeated measures ANCOVA: intervention effects on craving*

There were no (extreme) outliers for craving, skew and kurtosis were all in an acceptable range at all assessment points in the total sample. As there was a significant difference in baseline AUDIT scores (see Table 3), this variable was included as covariate in the analysis. The within-subject effects analysis revealed that the time effect was non-significant [ $F(1.79, 349.26) = 1.136$ ,  $p = .318$ ,  $\eta^2_{\text{partial}} = .006$ ], and no significant group effect was found [ $F(3, 195) = 0.59$ ,  $p = .623$ ,  $\eta^2_{\text{partial}} = .009$ ]. The within-subject group\*time interaction effect was significant [ $F(5.37, 349.26) = 4.17$ ,  $p = .001$ ,  $\eta^2_{\text{partial}} = .060$ ]. Descriptive data can be found in Table 6. A bar diagram, including denoted post-hoc tests that survived the FDR procedure can be found in Figure 3.

*Intention-to-treat analysis 2 repeated measures ANCOVA: intervention effects on alcohol consumption (g ethanol)*

Skew and kurtosis were similar as for the PPA, and despite of slight deviations from the proposed rule of thumb and some extreme cases due to reduction to zero alcohol consumption, deemed acceptable. As there was a significant difference in baseline AUDIT scores (see Table 3), this variable was included as covariate in the analysis. The rm ANCOVA within-subject effects analysis revealed a significant time effect [ $F(1.98, 386.85) = 11.20$ ,  $p < .001$ ,  $\eta^2_{\text{partial}} = .054$ ], and, as opposed to the PPA, a significant group effect was found [ $F(3, 196) = 2.85$ ,  $p = .039$ ,  $\eta^2_{\text{partial}} = .042$ ]. The group\*time within-subject effect was significant [ $F(5.95, 386.85) = 5.44$ ,  $p < .001$ ,  $\eta^2_{\text{partial}} = .077$ ]. The ITTA results for drinking amount in g ethanol broadly confirm the PPA results. Descriptive data can be found in Table 6. A bar diagram, including denoted post-hoc tests that survived the FDR procedure can be found in Figure 4.

*Replication of Intention-to-treat analysis 2 using MANCOVA: intervention effects on alcohol consumption (g ethanol)*

MANCOVA (covariates = AUDIT mean sum score, log-transformed alcohol consumption in g ethanol at baseline; fixed factor = group; outcomes = log-transformed post-test and follow-up alcohol consumption in g ethanol) was conducted to check robustness of findings controlling for systematic differences in alcohol consumption at baseline. The according multivariate omnibus test revealed that group significantly explained variance in the outcomes [ $F(6, 386) = 5.75, p < .001, \eta^2_{\text{partial}} = .082$ ], suggesting that the interventions had a significant effect on post-test and/ or follow-up alcohol consumption, after controlling for the identified baseline differences. The overall model explained 16.3% of variance in the outcome data (exact statistic:  $R^2_{\text{adjusted}} = .163$ ).

*FDR-corrected post-hoc results for PPA and ITTA*

In total, as there were four significant interaction effects, 4 (craving: PPA/ ITTA, alcohol consumption: PPA/ ITTA) \* 4 (groups) \* 3 (within-group contrasts between timepoints) = 48 within-group contrasts (paired t-tests) were computed. Concerning between group differences, 4 (craving: PPA/ ITTA, alcohol consumption: PPA/ ITTA) \* 6 (possible comparisons between groups within a given timepoint) \* 3 (timepoints) = 72 between group t-tests were computed. Of the total  $n_{\text{post-hoc}} = 120$  tests, 40 (33%) were significant ( $p < .05$ ) before applying the Benjamini-Hochberg method. After applying the procedure, 15 previously significant tests were discarded as false-positives, resulting in 25 post-hoc-tests identified as true-positives. The findings after FDR-correction, along with effect sizes for the respective contrasts, are indicated in the respective figures for the PPA and ITTA analyses (Figures 1 - 4).

In sum, the evidence was most coherent for craving, with significant decline most clearly arising within the AAT0 group in the PPA analysis from baseline to post-test (large effect), with relative stability towards the follow-up. Furthermore, in the PPA, the AAT+ group only exhibited a moderate, significant baseline-to-post-test effect, which was lost to follow-up (see



Figure 1). In the ITTA, the findings for AAT0 were replicated. However now also the AAT+ group showed significant, baseline-to-post-test and baseline-to-follow-up decline in craving (moderate effect; see Figure 3). For drinking amount in g ethanol, the pattern of findings was not in favor of imaginal retraining, with all groups showing significant decline in drinking amount to some extent, both in the PPA (Figure 2) and ITTA (Figure 4).

*Differences in training duration between the intervention groups*

Training duration was assessed on an ordinal scale level, hence non-parametric K-W tests were computed. Concerning training time reported at post-test, there was no difference between types of interventions (PE, AAT0, AAT+) on training time from pre-to-post ( $H(3, n = 126) = 3.99, p = .136$ ). From post-test to follow-up, training time also did not differ significantly between the active groups ( $H(3, n = 123) = 4.38, p = .112$ ), suggesting that differential engagement of participants may not constitute a source of systematic bias across these groups.

#### 4. DISCUSSION

By segregating the three main components of imaginal retraining, we strived to further elucidate the mechanisms underlying the technique and whether its craving reducing effects are due to the mechanisms of ‘enriched’ embodied cognition or can be achieved by a more simplified version. In addition, to enhance validity, we investigated the effect on alcohol consumption. Overall, the pattern of results suggests that AAT0 most consistently reduces craving in the PPA from baseline to post-test (large effect) with results remaining stable over time, rendering it the most favorable intervention. Moreover, for AAT+, a moderate effect between baseline and post-test, not stable over time (effect lost at follow-up), could be observed in the PPA. The ITTA showed similar results for AAT0, for AAT+ significant reductions in craving could be found from baseline to post-test, as well as post-test to follow-up with moderate effects. However, it needs to be carefully considered that there were baseline differences between AAT+ and PE, with individuals in AAT+ exhibiting significantly higher craving. In addition, on a descriptive level, higher craving was observed in both imaginal retraining groups which needs to be considered for interpretation. Although AAT0 and AAT+ show significant effects with regards to craving reduction, the effects for AAT0 were more robust. For the secondary outcome drinking amount in grams ethanol, the pattern of findings was inconsistent and not in favor of imaginal retraining, with all groups showing a significant decline in drinking amount to some extent, both in the PPA and ITTA. Interestingly, even PE and WLC show reductions in drinking amount. A reason for PE leading to a reduction in alcohol consumption might be that the psychoeducation component is mainly focused on providing advice to reduce one’s drinking amount and habits while the imaginal retraining components from AAT0 and AAT+ set out to actively reduce craving (Moritz et al., 2019a; Moritz et al., 2020; Moritz et al., 2019b; Moritz et al., 2021; Wirtz et al., 2021). Another explanation for this could be that participants in these groups compensated for their lack of intervention by taking other measures to address their drinking issues, as they participated in the study because of their

desire to reduce their alcohol intake. Furthermore, participants of PE and WLC might have been primed to desire and initiate change in drinking amount after participating in a study and being assessed twice beforehand.

Somewhat unexpectedly, the results did not support the hypothesis that the ‘enriched’ version of imaginal retraining, containing enhanced embodied cognitions and emotion induction, is most effective in reducing craving, somewhat contradicting the notion that ‘more is more’. Previous studies on alcohol- and other substance abuse behaviors have shown that substance-seeking can even increase through negative emotion induction, if participants display depressive symptoms and use substances to cope with negative affect (Hogarth et al., 2018; Hogarth et al., 2019), possibly supporting the notion that imaginal retraining can dispense with the component of emotion induction, especially as the results for the simplified retraining manual are more robust than those of the ‘enriched’ version. Additionally, the number of dropouts throughout the assessments was descriptively highest in AAT+ (47.5% ) compared to the other conditions (WLC = 38.8%, PE = 23.9% and AAT0 = 34.8%). Moreover, containing 11 pages and 3,683 words, the AAT+ manual is the longest of the three (621 words more than the AAT0 manual) and the instructions of the ‘enriched’ version are more abstract and complex to perform. More instructions need to be kept in mind, which might distract from focusing on pushing away the imagined beverage while carrying out the movement. As working memory capacity is limited (Proctor & van Zandt, 2008), accompanying imaginal retraining with ‘enriched’ embodied cognition might lead to a heightened cognitive load and take away from the actual exercise. Therefore, shorter instructions involving less factors to integrate might facilitate the exercise and its effectiveness. Thus, the heightened cognitive load of the AAT+ intervention might explain the more consistent findings concerning the simplified imaginal retraining version AAT0, as the focus is on the essence of approach bias modification.

**Limitations and future studies**

Due to the self-help nature of the study, it was not possible to track participants' adherence of following manual instructions and carrying out the training correctly. The results are based on retrospective self-reported information regarding treatment adherence, craving and alcohol consumption. However, prior imaginal retraining studies (Moritz et al., 2019a; Moritz et al., 2020; Moritz et al., 2019b; Moritz et al., 2021; Wirtz et al., 2021) faced the same drawbacks, yet promising results regarding the imaginal retraining technique were observed consistently across all previous trials. However, in all previous studies the original retraining technique (AAT+) has been proven to be effective in reducing craving for AUD (Moritz et al., 2019a), smoking (Moritz et al., 2020) and overeating (Moritz et al., 2019b). A pilot dismantling study examining the immediate effects of a single application of the imaginal retraining technique including an imagined- and an actual arm movement showed reduced craving at statistical trend level for high-calorie food with and without an actual arm movement (Wirtz et al., 2021). A similar study dismantling imaginal retraining in a sample of smokers showed significant reductions in craving only in the condition without an actual arm movement (Moritz et al., 2020). However, both studies examined the effects on craving after a one-time application only, which sets them apart from this RCT. In this trial we also ensured that participants could contact the study coordinators in case any questions and/or uncertainties regarding study participation/ implementation of the intervention remained, by providing an e-mail address and a phone number. As the AAT+ might not be as feasible in its original manual version as previously thought, future studies should evaluate the effect of instructing imaginal retraining in a video format, demonstrating all steps of the technique (learning by model). This way of illustrating the technique might make it easier to follow the instructions and carry out the movements, rather than reading the extensive manual. Increasing comprehensibility might in turn heighten the engagement of the participants.

Moreover, there was a distinct pattern in that non-completers displayed significantly higher levels of alcohol consumption and craving, as well as higher scores on the AUDIT (significantly more non-completers showing a score of  $\geq 16$  indicating a likelihood of alcohol dependence) and the ADS. From this it can be derived that participants who discontinued participation of the study displayed more severe alcohol-related symptoms than completers. This raises the question whether the manual version of imaginal retraining is a suited intervention for individuals with high levels of craving, alcohol consumption and alcohol related issues. Moreover, significantly older individuals dropped out of the trial. This might indicate that younger people are more likely to use the imaginal retraining manual, which poses an additional limitation, as the reach of the intervention may be restricted to an even more distinct demographic. Additionally, this study only included people who actively searched the platform Google for help to restrict alcohol intake, scrolled through Facebook self-help groups, or were actively invited to participate (in case they fulfilled selection criteria) through a student e-mail distributor. This poses the issue of a selection bias, as this form of reactive recruitment of participants may not reach a large representative part of the target population of problem drinkers but is restricted to those actively seeking help. Therefore, the generalizability of the results may be limited to younger individuals actively looking for self-help tools to intervene with their drinking behavior with only mild to moderate craving, alcohol consumption and alcohol related problems.

Furthermore, heterogeneity in drinking amount was very high, causing large variances between the ethanol scores, indicating the need to introduce stricter inclusion criteria for future studies than wishing to reduce alcohol intake. Alcohol consumption should be assessed in a more restricted fashion to attain a more homogeneous sample and to ensure comparability. For example, this could entail only including participants displaying a score of 8 or above on the AUDIT, indicating hazardous drinking behavior, and/or setting a stricter cut-off amount for

grams of ethanol, following established guidelines of problematic or risky consumption. This might enable future studies to specify and generalize results to a certain population.

Lastly, although this study showed the effectiveness of imaginal retraining, it is yet unclear how it compares directly to the original AAT performed on a computer. Future studies should aim to juxtapose the simplified imaginal retraining manual with the computer AAT in a RCT to discern which one is more effective in reducing craving and alcohol consumption.

### **Conclusion**

In sum, this RCT demonstrates that a simplified version of the imaginal retraining manual is more feasible and renders more consistent results in craving reduction compared to the original version, with moderate to large effect sizes. Based on the results of this study it can be concluded that components such as ‘enriched’ embodied cognition (i.e., emotion induction via thoughts and postures) may be dispensable or should perhaps be better instructed (e.g., in the form of a video). Additionally, the complexity of the AAT+ manual may have led to the high drop-out rate in this specific intervention group. Future studies should aim for a more homogenous sample by applying stricter inclusion criteria (clinical cut-offs) with regards to alcohol consumption, to avoid large variance and thus poor interpretability of the results. Based on the findings, it cannot be reasoned that imaginal retraining has a positive effect on reducing alcohol consumption at this point. However, this study provides more evidence that imaginal retraining significantly reduces craving for alcohol.

## References

- Babor, T., Higgins-Biddle, J., Saunders, J., & Monteiro, M. (2001). *AUDIT: The Alcohol Use Disorders Identification Test Guidelines for Use in Primary Care* (2nd ed.). Geneva, Switzerland: World health organization department of mental health and substance dependence.
- Barsalou, L. (2010). Grounded Cognition: Past, Present, and Future. *Topics In Cognitive Science*, 2(4), 716-724. <https://doi.org/10.1111/j.1756-8765.2010.01115.x>
- Benjamini, Y., & Hochberg, Y. (1995). Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal Of The Royal Statistical Society: Series B (Methodological)*, 57(1), 289-300. <https://doi.org/10.1111/j.2517-6161.1995.tb02031.x>
- Brockmeyer, T., Hahn, C., Retz, C., Schmidt, U., & Friederich, H. (2015). Approach Bias Modification in Food Craving-A Proof-of-Concept Study. *European Eating Disorders Review*, 23(5), 352-360. <https://doi.org/10.1002/erv.2382>
- Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences* (2nd ed.). Lawrence Erlbaum Associates.
- Dijksterhuis, A., & Bargh, J. (2001). The perception-behavior expressway: Automatic effects of social perception on social behavior. *Advances In Experimental Social Psychology Volume 33*, 1-40. [https://doi.org/10.1016/s0065-2601\(01\)80003-4](https://doi.org/10.1016/s0065-2601(01)80003-4)
- Dijkstra, K., & Post, L. (2015). Mechanisms of embodiment. *Frontiers In Psychology*, 6, 1-11. <https://doi.org/https://doi.org/10.3389/fpsyg.2015.01525>
- Erdfelder, E., Faul, F., & Buchner, A. (1996). GPOWER: A general power analysis program. *Behavior Research Methods, Instruments, & Computers*, 28(1), 1-11. <https://doi.org/10.3758/bf03203630>

- Fridland, E., & Wiers, C. (2017). Addiction and embodiment. *Phenomenology And The Cognitive Sciences*, 17(1), 15-42. <https://doi.org/10.1007/s11097-017-9508-0>
- Hedrick, V., Comber, D., Estabrooks, P., Savla, J., & Davy, B. (2010). The Beverage Intake Questionnaire: Determining Initial Validity and Reliability. *Journal Of The American Dietetic Association*, 110(8), 1227-1232. <https://doi.org/10.1016/j.jada.2010.05.005>
- Hogarth, L., Hardy, L., Mathew, A., & Hitsman, B. (2018). Negative mood-induced alcohol-seeking is greater in young adults who report depression symptoms, drinking to cope, and subjective reactivity. *Experimental And Clinical Psychopharmacology*, 26(2), 138-146. <https://doi.org/10.1037/pha0000177>
- Hogarth, L., Hardy, L., Bakou, A., Mahlberg, J., Weidemann, G., Cashel, S., & Moustafa, A. (2019). Negative Mood Induction Increases Choice of Heroin Versus Food Pictures in Opiate-Dependent Individuals: Correlation With Self-Medication Coping Motives and Subjective Reactivity. *Frontiers In Psychiatry*, 10, 274. <https://doi.org/10.3389/fpsy.2019.00274>
- Jakobsen, K., Frederiksen, J., Hansen, T., Jansson, L., Parnas, J., & Werge, T. (2009). Reliability of clinical ICD-10 schizophrenia diagnoses. *Nordic Journal Of Psychiatry*, 59(3), 209-212. <https://doi.org/10.1080/08039480510027698>
- Kakoschke, N., Kemps, E., & Tiggemann, M. (2017). Approach bias modification training and consumption: A review of the literature. *Addictive Behaviors*, 64, 21-28. <https://doi.org/10.1016/j.addbeh.2016.08.007>
- Lee, R., Hoppenbrouwers, S., & Franken, I. (2019). A Systematic Meta-Review of Impulsivity and Compulsivity in Addictive Behaviors. *Neuropsychology Review*, 29(1), 14-26. <https://doi.org/10.1007/s11065-019-09402-x>



- McHugh, R., & Weiss, R. (2019). Alcohol Use Disorder and Depressive Disorders. *Alcohol Research: Current Reviews*, 40(1). <https://doi.org/10.35946/arcr.v40.1.01>
- Meule, A., Vögele, C., & Kübler, A. (2011). Psychometrische Evaluation der deutschen Barratt Impulsiveness Scale – Kurzversion (BIS-15). *Diagnostica*, 57(3), 126-133. <https://doi.org/10.1026/0012-1924/a000042>
- Moritz, S., Paulus, A., Hottenrott, B., Weierstall, R., Gallinat, J., & Kühn, S. (2019a). Imaginal retraining reduces alcohol craving in problem drinkers: A randomized controlled trial. *Journal Of Behavior Therapy And Experimental Psychiatry*, 64, 158-166. <https://doi.org/10.1016/j.jbtep.2019.04.001>
- Moritz, S., Göritz, A., Schmotz, S., Weierstall-Pust, R., Gehlenborg, J., Gallinat, J., & Kühn, S. (2019b). Imaginal retraining decreases craving for high-calorie food in overweight and obese women: A randomized controlled trial. *Translational Psychiatry*, 9(1), 319. <https://doi.org/10.1038/s41398-019-0655-7>
- Moritz, S., Göritz, A., Kraj, M., Gehlenborg, J., Hottenrott, B., & Tonn, P., Ascone, L., Pedersen, A. & Kühn, S. (2020). Imaginal Retraining Reduces Cigarette Smoking: A Randomized Controlled Study. *European Addiction Research*, 26(6), 355-364. <https://doi.org/10.1159/000509823>
- Moritz, S., Gehlenborg, J., Wirtz, J., Ascone, L., & Kühn, S. (2021). A dismantling study on imaginal retraining in smokers. *Translational Psychiatry*, 11(1), 92. <https://doi.org/10.1038/s41398-020-01191-9>
- Patton, J. H., Stanford, M. S. & Barratt, E. S. (1995). Factor structure of the barratt impulsiveness scale. *Journal of Clinical Psychology*, 51(6), 768-774. [https://doi.org/10.1002/1097-4679\(199511\)51:6<768::AID-JCLP2270510607>3.0.CO;2-1](https://doi.org/10.1002/1097-4679(199511)51:6<768::AID-JCLP2270510607>3.0.CO;2-1)

- Pecher, D., & Zwaan, R. (2005). *Grounding cognition: The Role of Perception and Action in Memory, Language, and Thinking*. Cambridge, UK: Cambridge University Press.
- Perneger, T. (1998). What's wrong with Bonferroni adjustments. *BMJ*, *316*(7139), 1236-1238.  
<https://doi.org/10.1136/bmj.316.7139.1236>
- Proctor, R., & Van Zandt, T. (2008). *Human factors in simple and complex systems* (3rd ed.). CSR Press.
- Rey, A., Riou, B., & Versace, R. (2014). Demonstration of an Ebbinghaus Illusion at a Memory Level. *Experimental Psychology*, *61*(5), 378-384. <https://doi.org/10.1027/1618-3169/a000258>
- Riou, B., Lesourd, M., Brunel, L., & Versace, R. (2011). Visual memory and visual perception: when memory improves visual search. *Memory & Cognition*, *39*(6), 1094-1102.  
<https://doi.org/10.3758/s13421-011-0075-2>
- Skinner, H. A., & Horn, J. L. (1984). *Alcohol dependence scale (ADS): User's guide*. Toronto, Canada: Addiction Research Foundation.
- Spinella, M. (2007). Normative Data and a short form of the Barratt Impulsiveness Scale. *International Journal Of Neuroscience*, *117*(3), 359-368.  
<https://doi.org/10.1080/00207450600588881>
- Stacy, A., & Wiers, R. (2010). Implicit Cognition and Addiction: A Tool for Explaining Paradoxical Behavior. *Annual Review Of Clinical Psychology*, *6*(1), 551-575.  
<https://doi.org/10.1146/annurev.clinpsy.121208.131444>
- Vallet, G., Brunel, L., Riou, B., & Vermeulen, N. (2016). Editorial: Dynamics of Sensorimotor Interactions in Embodied Cognition. *Frontiers In Psychology*, *6*, 1-3.  
<https://doi.org/10.3389/fpsyg.2015.01929>

- Vermeulen, N., Corneille, O., & Niedenthal, P. (2008). Sensory load incurs conceptual processing costs. *Cognition*, *109*(2), 287-294.  
<https://doi.org/10.1016/j.cognition.2008.09.004>
- Watson, P., de Wit, S., Hommel, B., & Wiers, R. (2012). Motivational Mechanisms and Outcome Expectancies Underlying the Approach Bias toward Addictive Substances. *Frontiers In Psychology*, *3*, 1-12.  
<https://doi.org/10.3389/fpsyg.2012.00440>
- Wiers, R., Rinck, M., Kordts, R., Houben, K., & Strack, F. (2010). Retraining automatic action-tendencies to approach alcohol in hazardous drinkers. *Addiction*, *105*(2), 279-287.  
<https://doi.org/10.1111/j.1360-0443.2009.02775.x>
- Wiers, R., Eberl, C., Rinck, M., Becker, E., & Lindenmeyer, J. (2011). Retraining Automatic Action Tendencies Changes Alcoholic Patients' Approach Bias for Alcohol and Improves Treatment Outcome. *Psychological Science*, *22*(4), 490-497.  
<https://doi.org/10.1177/0956797611400615>
- Wiers, R., Gladwin, T., Hofmann, W., Salemink, E., & Ridderinkhof, K. (2013a). Cognitive Bias Modification and Cognitive Control Training in Addiction and Related Psychopathology. *Clinical Psychological Science*, *1*(2), 192-212.  
<https://doi.org/10.1177/2167702612466547>
- Wiers, C., Kühn, S., Javadi, A., Korucuoglu, O., Wiers, R., Walter, H., Gallinat, J. & Bermpohl, F. (2013b). Automatic approach bias towards smoking cues is present in smokers but not in ex-smokers. *Psychopharmacology*, *229*(1), 187-197.  
<https://doi.org/10.1007/s00213-013-3098-5>
- Wiers, C., Stelzel, C., Gladwin, T., Park, S., Pawelczack, S., & Gawron, C. (2015). Effects of Cognitive Bias Modification Training on Neural Alcohol Cue Reactivity in Alcohol

DISMANTLING IMAGINAL RETRAINING

Dependence. *American Journal Of Psychiatry*, 172(4), 335-343.

<https://doi.org/10.1176/appi.ajp.2014.13111495>

Wirtz, J., Ascone, L., Gehlenborg, J., Moritz, S., & Kühn, S. (2021). A dismantling study on imaginal retraining in overweight or obese women. *Translational Psychiatry*, 11(1).

<https://doi.org/10.1038/s41398-021-01595-1>

World Health Organization. (2018). *Global status report on alcohol and health 2018*. Switzerland.