UNIVERSITÄTSKLINIKUM HAMBURG-EPPENDORF

Klinik und Poliklinik für Psychiatrie und Psychotherapie

Prof. Dr. Jürgen Gallinat

An Investigation of Neuronal Changes and the Role of Response Inhibition in the Approach and Avoidance Task and a Contribution to the Discussion Surrounding the Veracity of the Ego Depletion Effect

Dissertation

zur Erlangung des Doktorgrades Dr. rer. biol. hum.
an der Medizinischen Fakultät der Universität Hamburg.

vorgelegt von:

Dimitrij Tycho Leon Kugler
aus Bielefeld

Hamburg 2024

Angenommen von der
Medizinischen Fakultät der Universität Hamburg am:

22.11.24

Veröffentlicht mit Genehmigung der
Medizinischen Fakultät der Universität Hamburg.

Prüfungsausschuss, die Vorsitzende:

Prof. Dr. Simone Kühn

Prüfungsausschuss, zweiter Gutachter:

PD Dr. Bastian Cheng

Table of contents

1. Introduction

1.1 Ego Depletion and the Rise of Dual-Process Models

Probably each of us knows the feeling of finding oneself in a situation, in which one behaved a certain way despite having had resolutions not to do so. Examples might pertain to prolonged media consumption, food consumption, infidelity, impulse buying, stopping a diet, alcohol-, and drug intake to name just a few. In fact, the phenomenon is as old as mankind and can be traced back to the fall of man in Christian theology, where Adam and Eve ate the fruit from the tree of knowledge of good and evil. All these situations have in common that an immediate gratification needs to be postponed to obtain more distal and abstract long-term goals, such as living a healthy life, saving money, or continuing living in Garden Eden to keep up with the biblical example. Psychologists have used different terms to refer to the ability to delay immediate gratification. Whereas the term inhibition is more frequently used in the experimental and neuroscientific literature, willpower, self-control, and self-regulation are more common terms in the clinical- and social psychological literature.

Almost 25 years ago, research by Baumeister and colleagues (1998) challenged the way we think of the ability to break the flow from intention to action. They showed that utterly different tasks, relying on the ability to exert self-control, led to worse performance on a second task that drew off the same capacity. Such states of so-called ego depletion were inconsistent with the predominant view of self-control being a stable personality characteristic (e.g., Mischel, Shoda, & Rodriguez, 1989) and provided first evidence for the resource model of self-control (Baumeister, Heatherton, & Tice, 1994). According to this model, self-control is a central and limited resource that can be temporarily exhausted if demanded too much, just like a muscle gets fatigued from prolonged activity (Baumeister, Vohs, & Tice, 2007). Importantly, the central attribute indicates that the same finite resource underlies all behaviors relying on self-control exertion. For instance, resisting temptations to eat desired foods, making an

important personal choice, and suppressing emotions were all shown to decrease the persistency on subsequent tasks tapping into self-control resources (Baumeister et al., 1998). Especially this vast applicability has intrigued researchers from various domains alike, leading to a surge of studies that showed depleted resources to affect behaviors as diverse as drug use, physical exercise, aggression, cheating, empathy, racism, test performance, and eating to name just a few (Baumeister & Tierney, 2011; Inzlicht & Friese, 2019). Another important aspect in understanding why ego depletion captivated social psychology is the possibility to make predictions about when and under which circumstances people would be more prone to act less controlled and throw overboard their intention. According to the resource model, this will be the case whenever control resources were engaged beforehand. This idea was highly influential and soon found its way into other (social) psychological theories, and therefore profoundly impacted psychological thinking over the last 25 years.

Dual-process models are one example of a theory that integrated the ideas of ego depletion into their own propositions. In these models, human behavior is conceptualized to result from the reciprocal engagement of reflective and impulsive processes. Generally, reflective processes are thought to be rule-based, accessible to conscious awareness, evolutionary new, and critically depend on the availability of cognitive resources. Impulsive processes, on the other hand, are evolutionary old, fast, associative, and do not depend on conscious awareness (Strack & Deutsch, 2004; Smith & DeCoster, 2000). Hence, self-control represents a reflective process, and its prior exertion will decrease the chances of more thoughtful, reflective acts soon thereafter therewith increasing the likelihood of impulsively driven behavior. Whereas earlier models claimed that both processes unfold sequentially, so that one cannot operate while the other is active (Smith & DeCoster, 2000), today's models take a more liberal stance, with both processes mostly working antagonistically but also synergistically (e.g., Hofmann, Friese, & Wiers, 2008; Strack & Deutsch, 2004).

1.2 Dual-Process Models and Addiction

Especially psychological disciplines dealing with phenomena in which behavior is disturbed and no longer under voluntary control like addiction, soon embraced ideas from dual-process models. By stressing the importance of unconscious processes in the genesis and maintenance of addictive behavior such models accounted for the long-observed paradox of sustained drug use despite having the wish to stop. More precisely, health-related dual-process models propose that an overly strong impulsive system determines addictive behavior with the influence of reflective control processes continuously weakening the more addiction progresses (e.g., Hofmann et al., 2008; Stacy & Wiers, 2006). Friese and colleagues (2011) used a horse rider metaphor to illustrate the interplay between both systems. While the rider represents the rule-based reflective system, the fast and associative impulsive system is represented by a horse. Even though in principle both systems can operate jointly to achieve the same goal, oftentimes the horse and the rider differ in their goals and steer in opposite directions. Thus, in people with addiction the horse has become increasingly powerful, peculiar, and takes in a more prominent role than the rider. On a functional level and less metaphorically, drug abuse is thought to sensitize reward- and compromise inhibitory control processes related to drug use, which establishes an imbalance between both systems (Wiers et al., 2007). Of course, such theoretical ideas had to be underpinned by empirical evidence and hence needed to be measure- and testable.

1.3 The Role of Implicit Measures in Dual-Process Models

In fact, the success story of dual-process models is closely linked to the rise of implicit measures of cognition which were often presented with the allure of being a window into the unconscious (e.g., Banaji & Greenwald, 2013). While earlier accounts focusing on unconscious processes in human behavior such as Freudian ideas were largely based on rather anecdotal evidence from clinical case studies, having at hand an arsenal of implicit measures, that ostensibly made

unconscious processes measurable and hypotheses empirically testable, is what really set dual-process models apart from earlier accounts. Interestingly, dual-process models parallel Freud's ideas in several ways; even though Freud proposed a triad of different processes in his psychic apparatus, he highlighted the importance of unconscious processes and described how different systems may interact to shape behavior (Freud, 1922). While the super-ego was theorized to contain the person's norms and values, the id was assumed to maintain basic drives and the ego to reconcile the forces between the two of them. Of course, the comparison is not exact, but the super-ego shows resemblance with the reflective- and the id with the impulsive system, which haggle about behavior without engaging the ego as a middleman.

In contrast to their explicit counterparts, implicit tests measure people's thoughts and feelings without having to ask them directly (Brownstein, Madva, & Gawronski, 2019), this way precluding socially desirable answers. According to De Houwer and Moors' (2007) more elaborate definition a test can be labelled implicit if it fulfills at least one automaticity feature. The specified features are processes that are uncontrolled, unintentional, goal-independent, purely stimulus-driven, autonomous, unconscious, efficient, and fast. Importantly a test does not have to meet all these criteria and still would be implicit. The last decades showed a great rush on implicit measures due to their special allure of unravelling determinants of behavior that were immeasurable beforehand. Among these, the implicit association test (IAT) is probably the most famous and has been used in countless studies ever since its introduction (Greenwald, Mc Ghee, Schwartz, 1998). The aim of this implicit test is to disclose associations between different constructs by means of a computerized categorization task. If one, for example, wants to investigate whether people with alcohol addiction hold positive, appetitive associations for alcohol, one would compare performances between blocks requiring the same response for alcohol- and positive words and blocks in which alcohol- and negative words require the same response. The procedure is based on the assumption that responses will be

facilitated if the two constructs sharing a response key are associated in the test taker's semantic network.

1.4 The Approach and Avoidance Task

From a clinical stance, the Approach and Avoidance Task (AAT) has been probably the most influential implicit measure, since it was not only used to measure biased cognition, but also demonstrated favorable effects as a treatment add-on. The assessment version of the task requires participants to execute approach and avoidance gestures conditional on a stimulus-irrelevant feature, such as the frame color or the format of the picture. For instance, participants could be instructed to approach all pictures in portrait- and to avoid all pictures in landscape format. This way the content of the pictures becomes task-irrelevant and therefore should not affect reaction times. If reaction times are influenced by picture content nonetheless, the terms approach and avoidance bias have been coined to refer to comparatively faster approach and avoidance behavior, respectively. First AAT research showed spider fearful individuals to respond more quickly to spider pictures by pushing than by pulling, both when compared to control pictures and reactions from non-spider fearful individuals (Rinck & Becker, 2007). An important characteristic of the task is its zooming feature, which gradually increases or decreases picture size upon pulling and pushing movements, respectively (Rinck & Becker, 2007). Not only does this amplify feelings of approach and avoidance, but also precludes cognitive reinterpretation of the movements. Over the course of the years, biased action tendencies have been shown in various domains such (alcohol) addiction (e.g., C.E. Wiers et al., 2014), heavy cannabis use (Cousijn, Goudriaan, & Wiers, 2011), obesity (Mehl, Mueller-Wieland, Mathar, & Horstmann, 2018) social anxiety disorder (Heuer, Rinck, & Becker, 2007), energy drink consumption (Kemps, Tiggemann, Cibich, & Cabala, 2019), and even social media usage (Juergensen & Leckfor, 2019).

While the potential to measure implicit approach and avoidance preferences attracted plenty attention, the possibility of using AAT-training to alter behavior and enhance treatment outcome is what really made the task a hot topic (Wiers, Eberl, Rinck, Becker, & Lindenmeyer, 2011). From a dual-process perspective this was sensational news. Not only could we measure components of the impulsive system but letting people execute approach and avoidance gestures that are incompatible with their impulsive urges alleviated psychopathological symptoms. The essential difference between AATs aiming to measure biased action tendencies and AAT-trainings lies in the contingency between the stimulus-irrelevant feature (e.g., picture format) and stimulus type (e.g., alcoholic- and non-alcoholic drinks). Typically, AATs that measure action tendencies use a contingency of 50/50. This means that pictures of both target categories are presented equally often in each format, so that they must be pushed/avoided and pulled/approached equally often. During AAT-training, in contrast, this contingency is shifted so that the vast majority of disease-specific stimuli become paired with a desired response. For instance, in the domain of alcohol addiction 90% of the alcohol stimuli would be presented in the format requiring a push response, so that patients with alcohol addiction learn to avoid such stimuli. Importantly, AATs in which responses depend on a stimulus-irrelevant feature (e.g., the format of a picture or its frame color) are called implicit, whereas AATs in which test takers need to respond to the stimulus categories themselves are referred to as explicit. Thus, in such an explicit AAT-training patients with alcohol addiction would be instructed to push away all alcohol- and to pull towards themselves all non-alcoholic pictures, so that the stimulus-irrelevant feature becomes obsolete by establishing a perfect contingency of 100%.

Perusal of the literature disclosed most beneficial treatment effects of AAT-training in the realm of addiction with inconclusive evidence for why that is (e.g., Cristea, Kok, & Cuijpers, 2016; Rinck, Wiers, Becker, & Lindenmeyer, 2018). Probably the most popular training study to this day showed that a very brief training period of four sessions decreased the

relapse probability among people with alcohol addiction by 13% (Wiers et al., 2011). Other early training studies demonstrated similar improvements in treatment outcome with concurrent declines in approach biases that mediated the effects (Eberl et al., 2013; Sharbanee et al., 2014), which established the catchy narrative that bias change is driving treatment effects. This idea is also apparent in the term cognitive bias modification (CBM), which has been coined as an umbrella term to refer to the AAT and related (implicit) tasks aiming to change biased cognition. However, the test of time provided rather little evidence for this narrative, since replication of the aforementioned mediational pattern mostly failed. For instance, the most elaborate clinical trial disclosed improvements in treatment outcome among people with alcohol addiction, but failed to demonstrate simultaneous change of action tendencies (Rinck et al., 2018). Likewise, a meta-analysis in the realm of addiction indexed positive treatment effects, but yielded no support for a mediational effect of bias change (Cristea et al., 2016). Regardless of this contradictory evidence, the storyline of the mechanistic role of action tendencies persisted (e.g., Baird et al., 2017; Eiler et al., 2020; Rinck et al., 2018), thence nurturing the (mis)belief of knowing the mechanism underlying AAT-training. Also terminologies such as CBM or approach bias retraining might have contributed to this misbelief, since both imply that altering biases is key in bringing about treatment effects. Thus, the neglect of contradictory evidence in conjunction with somewhat misleading terminology might be the reason why astonishingly little research addressed mechanisms of AAT-training other than changed action tendencies.

1.4.1 AAT and Neuroimaging Research

Neuroimaging research and its potential to disclose brain areas associated with AAT-training provides a fruitful option to generate new theoretical ideas about the cognitive processes implicated in the task. To this day, such imaging research is very scarce and still in its infancy

with only five studies published on the topic. The very first neuroimaging study used functional near-infrared spectroscopy (fNIRS) in people with alcohol use disorder and showed orbitofrontal cortex activity when approaching and dorsolateral prefrontal cortex (dlPFC) activity when avoiding alcohol (Ernst et al., 2014). Later studies using functional magnetic resonance imaging (fMRI) confirmed prefrontal correlates of alcohol approach behavior and in addition showed the involvement of subcortical structures such as the nucleus accumbens (NAc) (C.E. Wiers et al., 2014; C.E. Wiers et al., 2015a). Research focusing on changes accompanying AAT-training indicated that approach-related activity in the medial prefrontal cortex (mPFC) decreased in patients with alcohol addiction (C.E. Wiers et al., 2015a). Moreover, cue-induced activity in the amygdala declined after AAT-training (C.E. Wiers et al., 2015b). Besides the aforementioned alcohol-related studies, there is only one neuroimaging study targeting a sample of healthy individuals. This study demonstrated activity in utterly different brain regions with lingual, parahippocampal, and fusiform gyri being activated while approaching positive social scenes (Ascheid, Wessa, & Linke, 2019). Thus, rather than relying on a largely identical network of brain areas, the brain areas involved seem to differ across different populations and areas of application (i.e., presented content), with the fysiform gyrus known as the face area of the brain being more important for the appraisal of social scenes and motivational and reward-related structures such as the mPFC, amygdala, and NAc being more specific for drug-related approach behavior.

1.5 Dual-Process Models and Inhibition Training

Relatively independent of AAT-research, but considerably influenced by dual-process theory, trainings that aim to enhance executive control functions have developed over the last decade. Whereas AAT-training is supposed to tap more into the impulsive determinants of addiction, other trainings focused on reflective processes to regain behavioral control. Keeping up with the horse rider metaphor, such trainings set out to strengthen the rider rather than taming the

horse (Friese et al, 2011). Again, most studies focused on alcohol and required participants to withhold prepotent responses to alcohol stimuli in a Go No-Go paradigm. Results showed improved self-control capacity with concomitant decreases in alcohol consumption (Houben, Havermans, Nederkoorn, & Jansen, 2012; Houben, Havermans, Wiers & Jansen, 2011; Jones & Field, 2013). Also, an app-based inhibition training in which participants needed to swipe all items, except those specified, from a buffet to their plate could be shown to enhance response inhibition (Kühn et al., 2017). On closer inspection the app-training shows some similarities with AAT-training. More precisely, collecting items via swiping gestures resembles approach trials of the AAT. While inhibiting responses on the remaining trials is different from avoidance trials, both still share a crucial commonality from a dual-process stance. In this view, drug-related stimuli are thought to elicit relatively automatic approach tendencies and both reactions (i.e., avoidance and inhibition) have in common that behaviour does *not* follow pre-established action tendencies. Because both trainings have previously shown effective at decreasing maladaptive drug consumption, the intriguing possibility that part of AAT-training effects could be due to inhibiting prepotent responses arises. Thus, rather than carrying out movements that oppose one's automatic action tendencies, the inhibition of such implicit approach urges could be (partly) responsible for the beneficial effects of AAT-training. This exciting possibility was tested meticulously over the course of the present dissertation.

1.6 Dual-Process Models and Aggression

Also in the realm of aggression research dual-process models served as useful sources of information, since their theoretical propositions allowed to specify and predict the circumstances under which individuals would behave aggressively. Anderson and Bushman (2002) have defined human aggression as any intentional behaviour that is directed towards other individuals who want to avoid this harm. In addition, their General Aggression Model (GAM) proposes that personality- and situational factors in conjunction with the present

internal state will determine whether a person will engage in thoughtful or more impulsive, aggressive acts. Importantly, the present internal state encompasses the current affect, arousal, and cognition with depleted self-control resources captured by the latter. Thus, ideas from ego depletion as well as dual-process models appear in the GAM and lower amounts of cognitive resources are theorized to increase the chances of acting via the impulsive rather than thoughtful route. Even though these impulsive and thoughtful routes are not identical to the two systems proposed in dual-process models, the resemblance is obvious. In fact, research has indicated theory conform increases in aggressive behaviour after the depletion of self-control resources (e.g., Barlett, Oliphant, Gregory, & Jones, 2016; Osgood & Muraven, 2016; Stucke & Baumeister, 2006).

1.7 Contradictory Evidence

After the initial honeymoon, all the aforementioned research fields faced findings that contradicted basic theoretical ideas. Most notably, the honeymoon of ego depletion research was followed by a replication crisis. Meta-analyses pointed out that the field was substantially distorted by publication bias and that the effect was much smaller (Carter & McCullough, 2014) or even indistinguishable from zero (Carter, Kofler, Forster, & McCullough, 2015) if analyses used techniques to correct for its influence. Hence, despite more than 600 studies of support, a debate whether ego depletion is real or not overtook the field (Inzlicht & Friese, 2019). While publication bias can be tackled relatively easy by preregistering all future studies, the so-called conceptual crisis of ego depletion research is more intricate to handle (Lurquin & Miyake, 2017). Most seriously, self-control tasks used in ego depletion research were not independently validated. Rather, the observation of ego depletion effects was used to deem both tasks used in the sequential-task paradigm as suitable and validated measures of self-control. Therefore, self-control tasks need careful attention and their capability to measure self-control needs to be re-evaluated through systematic research.

Obviously, if the existence of ego depletion is doubted likewise its role in dual-process models must be called into question. Another issue that parallels problems in ego depletion research is the need to carefully re-evaluate the measurement instruments used throughout the dual-process literature. According to dual-process models impulsive processes are thought to be qualitatively different from reflective processes and implicit measures were proposed as suitable candidates to capture such processes (Hofmann et al., 2008). In contrast to reflective processes, which can be investigated by introspective, explicit measures, impulses are more difficult to capture and study and mostly rely on measuring automatic responses. Obviously, basing conclusions on automatic responses holds greater potential for misinterpretations than directly asking people in standardized explicit measures. Moreover, implicit tests and implicit cognition were often conflated so that implicit tests were inferred to measure qualitatively different aspects than explicit measures just because of being implicit. Yet, whether implicit and explicit tests in fact measure qualitatively different aspects needs to be verified in the process of construct validation. Here, implicit- and explicit measures of the same construct should show convergent- and discriminant validity, respectively. Unfortunately, construct validation is lacking for a great deal of implicit measures and whether these tests in fact measure implicit cognition is still uncertain and remains to be demonstrated.

1.8 Aims and Relevance

The main aim of this dissertation was to obtain a more comprehensive understanding of the cognitive and neuronal changes which accompany AAT-training and explain its effectiveness in the best case. Here, special emphasis laid on understanding the role of response inhibition. Earlier studies that successfully decreased drug consumption through response inhibition training suggest that the inhibition of relatively automatic responses to drug stimuli might play a role in addictive behaviour (e.g., Houben et al., 2011; Houben et al., 2011; Jones & Field, 2013). From a dual-process stance, AAT- and response inhibition training have in common that

people must refrain from acting upon automatic, drug-elicited approach urges. Thus, the success of both trainings might be (at least partially) rooted in the same process of response inhibition. To test this, we developed a new AAT-training that was compared to regular- and sham AAT-training. In this new training reactions to control stimuli were identical to regular AAT-training (i.e., avoidance gesture), however for drug stimuli participants were required to withhold responses rather than carrying out avoidance gestures. Importantly, the AAT-typical zooming feature shrank picture sizes upon avoidance gestures as well as withheld responses (Rinck & Becker, 2007), so that active avoidance gestures constituted the essential difference between both trainings. The comparison of both trainings in two separate projects served as the main source of information concerning the role of response inhibition in AAT-training.

In the first project inhibition- and regular AAT-training were compared in people with alcohol addiction. For this purpose, we randomly assigned inpatients with alcohol addiction to one of three conditions (regular AAT-, inhibition AAT-, or sham AAT-training). Besides seeking to replicate basic AAT-training effects of improved treatment outcome, we put special emphasis on studying both the differences and communalities between inhibition- and regular AAT-training. We hypothesized that after the training participants in both training groups would consume less alcohol than people in the sham control group. Yet, somewhat more pronounced effects were expected for patients undergoing regular AAT-training, since this group must not only inhibit prepotent approach responses but acted in opposition to such approach urges by actively avoiding alcohol images. In sum, sham-training was expected to be least effective while regular AAT- was thought to be slightly more effective than the inhibition AAT-training.

In the second project the same training was used to support people in their attempt to quit smoking. Experimental conditions were identical to the previous project and the

effectiveness of each training in smoking cessation was investigated. However, this time special emphasis was devoted to monitoring potential changes of inhibitory control. To this end, a cognitive psychological response inhibition task was employed to track changes of inhibition over the course of the project. Again, differences between all groups, and the inhibition- and regular AAT group in particular, were tested to become a more complete picture than the one obtained from examining changes in treatment outcome in isolation. Two other objectives were to investigate the psychometric properties of the smoking-AAT and to compare a mobile and PC-based version of the task. Perusal of the literature indicates that studies investigating the psychometric properties of the AAT are a scarce good. Moreover, the AAT possesses different levels of reliability for different constructs (Loijen, Vrijsen, Egger, Becker, & Rinck, 2020). Therefore, reliability cannot be generalized from one construct to the other and needs to be determined independently for each area of application instead. To the best of our knowledge, we were the first to assess the task's reliability in the realm of nicotine dependence. Self-evidently, knowing about such parameters is crucial for the appropriate interpretation of research findings, thence helping to put future findings into the right perspective. Likewise, knowing whether the precision of mobile and PC-based (smoking) AATs is about equal is important for future research, since their equivalence would open a door for various new applications and more sophisticate research designs by employing mobile versions. Therefore, both versions were directly compared over the course of the present dissertation.

Contrasting with most AAT-research that either attempted to measure biased action tendencies accompanying (mental) health conditions or sought to alleviate such conditions through AAT-training, a third project took a different stance and tested whether the AAT can also be used to establish more healthy preferences among healthy participants. If so, this would suggest that the AAT can be used as a preventive tool to help at risk individuals, and more generally to support individuals in living up to their resolutions. More precisely, it was

investigated whether AAT-training can induce a preference for healthy beverages. To this end, participants were randomly assigned to an AAT- or sham-training. In the sham-training control group healthy and unhealthy drinks were presented equally often in both formats, so that pictures of both categories had to be approached and avoided equally often. In the AAT-training group, this contingency was shifted so that 90 % of the healthy stimuli required a pull response and 90 % of the unhealthy stimuli must be avoided. Functional brain scans as well as beverage intake were assessed before and after the training. On the behavioural level, we investigated whether the experimental group made comparatively healthier beverage choices after the training than participants in the control group. On the neural level, we were primarily interested in brain changes that might be associated with such behavioural change. Generally, we expected addiction-specific brain areas linked to the AAT by previous research to play a small role if any and assumed a more prominent role of motor and subcortical areas among healthy participants. Insights from healthy participants will help to enhance our understanding of general rather than disease-specific neuronal changes associated with the training.

Another aim was to contribute to the ongoing debate whether ego depletion is a real phenomenon or not. This discussion is vital for AAT research as well, because ego depletion had a decisive impact on dual-process theory and the latter serves as the most common theoretical framework within which the AAT is situated. Among more than 600 studies, demonstrating ego depletion effects, a small proportion showed increases in aggressive acts after the exhaustion of self-control resources (e.g., Barlett et al., 2016; Osgood & Muraven, 2016; Stucke & Baumeister, 2006). Of course, if the effect is questioned altogether, the small fraction of evidence showing a link between depleted self-control resources and aggression needs a critical revision, too. In a first step and separate project, a German version of the word completion task (WCT) for aggression was developed and tested for its suitability as an aggression measure. The elegance of this task lies in its extremely low level of face validity and

its successful construct validation would enable future research to measure the socially sensitive topic of aggression without any priming or self-presentation effects. This unawareness of what is being measured is also highly advantageous when studying the relationship between depleted self-control resources and aggression. Therefore, the task was used in a subsequent ego depletion project to determine whether previous self-control exertion adversely affected aggression indices (WCT and IAT) or general cognitive performance in a response inhibition task. Evidence from this pre-registered project will contribute to the discussion whether ego depletion is real or whether its existence was mostly an artefact of publication bias.

2. Measures

This chapter gives an overview over the most important measurement instruments that were used to answer the aforementioned research questions over the course of the present dissertation. Most roughly, the instruments can be categorised into measures aiming to track changes in consumption, severity symptoms of addiction, neuronal activity, and response inhibition, and automatic action tendencies.

Reports of consumption

All AAT-research was similar in the sense that the influence of AAT-training on later consumption was tested. Thus, even though slightly different methods were used, the objective of obtaining an as accurate account of the participants' consumption was the same across the different projects and served as main outcome measure.

Timeline Follow-Back:

The Timeline Follow-Back interview was used to assess the alcohol consumption of people with alcohol addiction (Sobell & Sobell, 1992). In this interview, participants were asked to recall their daily consumption within specified periods. If consumption for some days could not be recalled right away, several procedures were used to aid in recall. Among others, a visual calendar and key dates sought to reconstruct memories for such days.

Number of Cigarettes Smoked:

The number of cigarettes smoked within different periods was determined by means of a comparable procedure. First, participants needed to indicate days of abstinence within specified periods. Next, they were asked to recall the days on which they smoked and to indicate the number of cigarettes smoked per day. Finally, the total number of cigarettes smoked was computed by multiplying the days on which participants smoked with the number of cigarettes smoked per day.

Beverage Consumption:

Two different measures assessed the extent of unhealthy and healthy beverage consumption. While the diary report resembled the previous attempts to recall consumption within a certain period, the second measure assessed actual consumption in a bogus taste test.

Diary Report:

In the diary report of beverage consumption participants were presented with an extensive list of beverages from which they had to choose the ones they had consumed within the last two weeks. If a beverage was consumed, in-depth questions assessed the frequency and average amount of consumption. Thus, the total amount of consumption could be obtained by multiplying both indices. Lastly, the so obtained numbers were multiplied with the beverage's sugar content and added up across all beverages consumed to yield an index of total sugar consumption within the two weeks. To enhance the accuracy of this measure participants were instructed to keep track of their beverage consumption after the training.

Bogus Taste Test:

Rather than assessing consumption via retrospective self-reports, the bogus taste test measured actual consumption. During an ostensible soft drink tasting participants were asked to try each of four beverages in order to give an estimate of its sugar content. In fact, we measured the amount of consumption for each drink. This amount was calculated by subtracting the cups weight after trying from its initial weight. To minimize the influence of thirst on consumption, participants were constantly provided with the opportunity to drink water during the testing session.

Severity indices

Fagerstrom Test for Nicotine Dependence (FTND):

The FTND (Heatherton, Kozlowski, Frecker, & Fagerstrom, 1991) entails six questions that provide insight into the severity of nicotine dependence.

Cigarette Dependence Scale (CDS):

The CDS (Etter, Houezec, & Perneger, 2003) is a slightly more extensive measure of nicotine dependence comprising a total of 15 questions.

FMRI as a neuroimaging method

Over the course of the present dissertation fMRI was used to track changes in brain activity unfolding after AAT-training. Measuring and understanding brain activity lies at the core of cognitive neuroscience and different brain-imaging methods each coming with its own advantages and disadvantages evolved over the last decades. Since electricity is the language of the brain, we differentiate between measures that assess electrical impulses directly and those indirectly inferring brain activity from other parameters. Neural communication takes place via electrical impulses (action potentials) that release chemicals (neurotransmitters) at synaptic junctions to communicate with and shape the activity of the next cell (e.g., Hyman, 2005). FMRI as a non-invasive brain-imaging method is based on the observation that cerebral blood flow and neuronal activity are systematically coupled (Ogawa et al., 1990). Thus, fMRI constitutes an indirect measure of neural activity, in which brain activity is inferred from homeostatic adjustments of the blood flow, so-called haemodynamic responses. Such responses guarantee that stressed tissue is replenished with essential nutrients such as oxygen and glucose so that continuous and smooth functioning of these brain areas is ensured after neural activity occurred. Because oxygenated and de-oxygenated blood differ in their magnetic susceptibility, we can infer changes in the ratio of oxygenated and de-oxygenated blood by assessing blood-

oxygen-level dependent (BOLD) signals in fMRI. And since brain activity and blood flow of oxygenated blood are systematically coupled, we can use BOLD signals to infer where and somewhat less accurately when neuronal activity occurred.

Response Inhibition

Go/No-Go task:

A classical response inhibition task was used to track changes of inhibitory control. In this task participants were required to react with an immediate button press to all the numbers from one to nine, except the number three. Thus, trials presenting the number three constituted No-Go trials, while all remaining trials were Go trials and thence required a response. The short interstimulus interval (1150ms) and the high proportion of Go- to total trials (218 of 236) established response automatisms making it hard to abort already initiated responses during the No-Go trials of the task. The number of wrong button-presses on No-Go trials, so-called commission errors, served as index of response inhibition with fewer mistakes expressing better inhibitory capacity.

Automatic action tendencies

Assessment AAT:

The repeated measurement of automatic action tendencies over time was achieved by an implicit assessment AAT. Thus, a stimulus-irrelevant feature such as the format of a picture (portrait/landscape) or its frame color (blue/yellow) determined whether participants needed to respond with approach or avoidance gestures. Importantly, every picture was presented once in each format or with each frame color so that an approach as well as an avoidance gesture had to be made for each picture. Subtracting avoidance from approach responses for the different stimulus categories (e.g., drug-related and control) built the basis for computing approach bias scores.

3. Results

Kugler, D., Scholz-Hehn, D., Schäfer, I., Strate, P., Gallinat, J., & Kühn, S. (2022, February 15). The Role of Response Inhibition in the Approach and Avoidance Task: No Evidence for Approach and Avoidance Task-training effects in Inpatients with Alcohol-Use Disorder. https://doi.org/10.31234/osf.io/8mj34

3.1 The Role of Response Inhibition in the AAT among People with Alcohol Addiction

Addiction is characterized by repeated drug intake despite the knowledge of its harmful consequences (Stacy & Wiers, 2010). Paradoxically this is sometimes accompanied by an urge to stop. The rise of dual-process models in recent years can in part be attributed to the fact that these models present an explanation for this paradox. Despite slightly different characteristics, all these models juxtapose two systems of explicit and implicit cognitions in order to describe, explain and predict human behaviour. As we have seen, the reflective system is the rule-based, deliberate and slow system, whose activities are accessible to conscious awareness and require cognitive resources, whereas the impulsive system is the associative, unconscious and fast system that operates independent of the amount of cognitive resources available (e.g., Hofmann et al., 2008; Strack & Deutsch, 2004). In view of these models, addiction is characterized by an imbalance between sensitized and increasingly strong associative activities of the impulsive system and impaired executive processes of the reflective system (Wiers et al., 2007). More precisely, repeated drug intake changes reward-related learning processes (Hyman, Malenka, & Nestler, 2006; Wrase et al., 2002), so that cues associated with a drug increase in incentive salience (Robinson & Berridge, 1993), drug use motivational processes progressively start to operate outside of conscious awareness (i.e. they become implicit) (De Houwer, 2006), and drug cues may then automatically trigger approach behaviour (e.g. Curtin, McCarthy, Piper, & Baker, 2006; Robinson & Berridge, 2003; Tiffany, 1990). In support of this, people with alcohol addiction have been shown to exhibit automatic approach tendencies (approach bias) for alcoholic stimuli (e.g., Ernst et al., 2014; C.E. Wiers et al., 2014; Wiers et al., 2011).

Considerably influenced by the ideas of dual-process models, two different trainings aimed at decreasing problematic drug consumption evolved over the last decades. AAT-training attempts to change automatic, cognitive-behavioural pattern characteristic of (alcohol) addiction by letting patients avoid stimuli related to their addiction and approach non-alcoholic alternatives. Other trainings have focussed on learning to inhibit prepotent responses to alcoholic stimuli in Go / No-Go paradigms and indicated that fostering self-control can decrease alcohol intake among heavy drinkers (Houben et al., 2012; Houben et al., 2011; Jones & Field, 2013). Thus, from a dual-process perspective, AAT-training focusses on changing automatic processes of the impulsive system, whereas inhibition training seeks to regain behavioural control by strengthening the reflective system. An interesting possibility, however, is that AAT-training engages response inhibition, also. This supposition is based on a study by Kühn and colleagues (2017), which successfully used computer game elements to train response inhibition. In this game participants were asked to swipe all (food) items, except those specified as inhibition items, from a buffet to their plate. Because far more non-inhibition items were presented with the intervals between the trials becoming successively shorter, the chances were high that participants had to disengage from, and thus inhibit, already planned motor responses during the remaining inhibition trials. While approach and Go trials share the resemblance of fetching items via hand movements, the similarity between avoidance and No-Go (inhibition) trials is less obvious. Yet, from a dual-process perspective both the inhibition- as well as avoidance response have in common that actions do *not* follow pre-established action tendencies. Moreover, stimuli are perceived as moving away in both tasks with the disappearance depending on the inhibition of (prepotent) responses and active avoidance during gameplay and AAT-training, respectively. Thus, from this stance active avoidance gestures during AAT-training constitute the essential difference between the two of them. In light of the similarities between the trainings and their shared potential to reduce problematic drug

consume, the question whether both engage the same mechanism arises. More specifically, (part of) the positive training effects of the AAT might be due to successfully inhibiting drug-induced approach urges rather than actually avoiding the drug. If this was true, the effectiveness of AAT-training would more heavily rely on reflective processes than previously thought.

Interestingly, some research already lent support to the involvement of reflective and conscious aspects of cognition in the task. For example, more pronounced bias scores were shown when stimuli were task-relevant (Lender et al., 2018; Meule et al., 2019). Similarly, more pronounced training effects were demonstrated when participants became aware of the stimulus-action contingency within the AAT-training (Van Dessel, De Houwer, & Gast, 2016). Furthermore, Koch and colleagues (2008) showed that avoidance movements can foster reflective aspects of cognitive control. Most interestingly for the current project, Van Dessel and colleagues (2015) found instruction-based approach and avoidance effects, i.e., stimulus evaluations changed through the mere instruction of approaching or avoiding stimuli without the actual need to execute these gestures at any point of the experiment. Self-evidently, the fact that the execution of an actual avoidance gesture was not required for changing cognition is further water on the mills of the assumption that active avoidance might be less important than widely assumed.

Hence, the main question addressed in the present project is whether AAT-training effects manifest themselves by carrying out movements that conflict with an individual's automatic action tendencies or whether the mere inhibition to act upon these drug-elicited approach urges might suffice to improve treatment outcome. Or in other words, an attempt is made to disentangle the role of active avoidance and response inhibition in AAT-training. Another focus lies on obtaining a better understanding of the longevity of training-induced bias change. To this end, 76 inpatients with alcohol-use disorder were randomly assigned to one of

three AAT-training groups (avoidance-, inhibition-, and sham AAT-training) and the effectiveness of each training was investigated over the course of a three month period entailing three appointments (pre/post/follow-up assessment). Participants receiving regular AAT-training were asked to avoid alcoholic drinks, while they had to approach pictures of non-alcoholic beverages. Likewise, participants in the inhibition AAT group were instructed to approach non-alcoholic beverages, however for alcoholic drinks they were requested to give *no* responses, thus to inhibit their responses. The active control group had to swipe alcoholic pictures to the left and non-alcoholic pictures to the right or vice versa (depending on counterbalancing). All three groups received a three-week training in addition to standard cognitive-behavioural therapy (CBT).

It was hypothesized that participants receiving either regular AAT- or inhibition AAT-training would show a better treatment outcome than those in the active control group with treatment outcome being defined as the total amount of alcohol consumption over the course of the project. Yet, due to a potential additive effect of performing active avoidance gestures, participants receiving regular AAT-training were hypothesized to show slightly better treatment outcomes than participants in the response inhibition group. Differences between both groups will be of particular interest since these could shed light on shared and distinct mechanisms of both trainings. The same general pattern of results was expected for changes in alcohol approach biases. Thus, participants receiving either regular or inhibition AAT-training were expected to show more pronounced reductions in alcohol approach bias than participants in the active control group, and reductions were supposed to be slightly higher in the group receiving regular AAT-training.

Information on Procedure and Project

Participants

The project was funded by the European Research Council (ERC-2016-StG-Self-Control-677804) and approved by the Ärztekammer Hamburg (PV5189). It was pre-registered at Clinical Trials.gov (NCT04054336) and we were aiming for a sample size of 90 participants. Recruitment of participants started in January 2017 and lasted until February 2020. A final sample of 76 inpatients with alcohol-use disorder, who completed all parts of the project, could be recruited from two study sites (University Clinic Hamburg-Eppendorf and the Asklepios Clinic Nord, Hamburg). Time constraints with another project, targeting the same sample, starting in 2020 and a higher dropout rate than expected (36 percent; see Fig. 1 for a detailed description), let us fall short of our designated sample size.

A confirmed diagnosis of alcohol-dependence and the absence of any current withdrawal symptoms were the main inclusion criteria for taking part in the project. Suffering from another substance-use disorder than alcohol served as an exclusion criterion. Other criteria precluding participation were a current major depressive episode, the presence of psychotic disorders, alcohol consumption in between withdrawal and the first appointment, the presence of pronounced withdrawal symptoms, a somatic or neurological disease that precludes the proper execution of the program, long-term intake of neuroleptics and/or hypnotics, presence of axis-II disorders, and a planned rehabilitation therapy after the three-week detoxification program.

Diagnoses of alcohol-use disorder were confirmed using the Structured Clinical Interview for DSM-IV Disorders (SCID) section E (First & Gibbon, 2004) and the absence of pronounced withdrawal symptoms was assessed by means of the Clinical Institute Withdrawal Assessment for Alcohol (CIWA-A) (Sullivan, Sykora, Schneiderman, Naranjo, & Sellers,

1989). In addition, the Mini Neuropsychiatric Interview (MINI) (Sheehan et al., 1998) was used to assess eligibility.

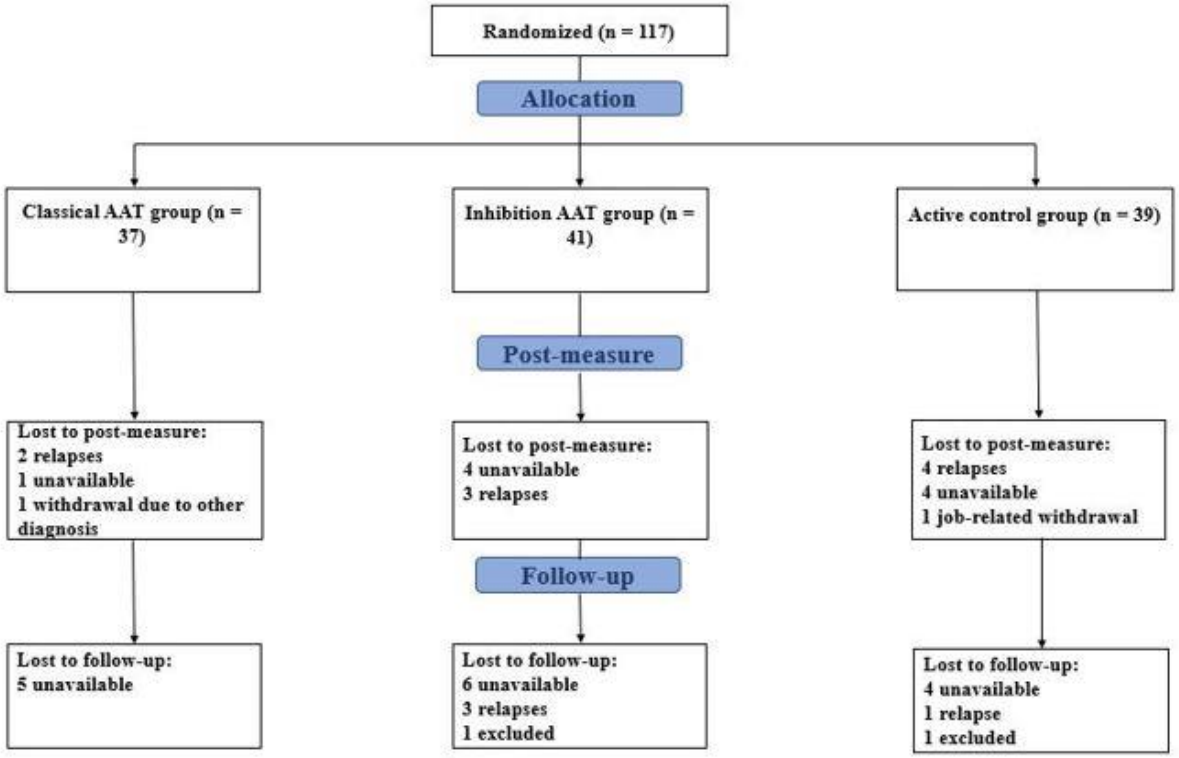

Figure 1: Consort flow diagram. NB: Patients were not excluded because they relapsed, but felt incapable to complete the study due to their relapse.

Procedure and research design

The project employed a longitudinal randomized-controlled trial design with three arms and three points of measurement. The amount of alcohol consumption, the severity of alcohol-addiction symptoms, and automatic action tendencies for alcohol were assessed during each of the three cognitive-behavioral testing sessions (see below for in-depth information). The first testing session occurred during hospitalization (after physical withdrawal and before discharge) at the end of which participants were provided with information about the upcoming training. More precisely, participants were instructed to spend 15 minutes daily on training until the post-

measure three weeks later. Noteworthy, the training was administered on a tablet device (IPad Air) with an AAT-app installed for training purposes. A final appointment (follow-up) aimed at assessing potential long-term changes marked the end of the project and took place nine weeks and three months after the post measure and initial appointment, respectively. At the end of this last appointment patients were debriefed and received 100 € monetary compensation for their participation.

Stimuli

The stimulus set consisted out of 144 (72 alcoholic and 72 non-alcoholic) pictures. All stimuli were presented in front of a white background with a subset of 72 stimuli being used for the AAT-training.

AAT-training

Participants were asked to train 15 minutes daily with the app (3 rounds, 5 minutes each) and instructed to react differently to pictures with and without alcoholic content, thence using explicit task instructions. Participants in the regular AAT group were requested to push away alcohol-related pictures (avoid) and to pull towards themselves pictures with non-alcoholic content (approach) (see Fig. 2). Participants in the inhibition AAT group had to withhold responses when alcohol-related images were presented, while non-alcoholic stimuli had to be approached just like in regular AAT-training. Importantly, a zooming feature gradually shrank picture size upon pushing and withheld responses, so that both groups perceived alcohol stimuli as moving away. Inversely, picture size was enlarged when participants swiped the stimuli towards themselves (Rinck & Becker, 2007). The active control group was asked to swipe alcoholic pictures to the right and non-alcoholic pictures to the left or vice versa (depending on counterbalancing). All trials terminated with given responses or timed out after 2,06 seconds otherwise.

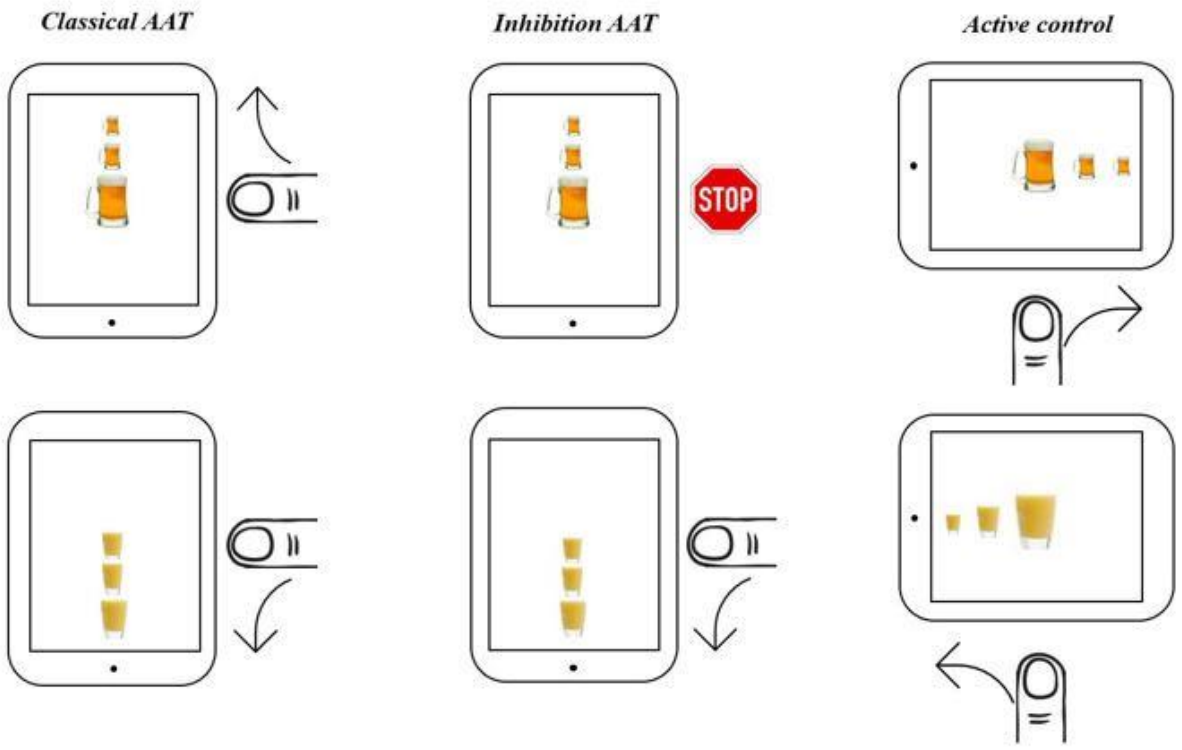

Figure 2: Graphical representation of the task instructions during the AAT-training.

AAT for approach bias assessment

The strength of automatic approach tendencies for alcohol and soft drinks was measured by an implicit AAT on a tablet device with the actions depending on the format of the pictures. The task was programmed in unity (https://unity.com/) and administered to the participants on an IPad Air. Pictures of 72 alcoholic and 72 non-alcoholic drinks were presented twice, once in portrait- and once in landscape format. Picture format to response assignment was counterbalanced, so that half of the participants had to swipe away pictures in portrait format and swipe towards themselves pictures in landscape format (or vice versa). Hence, irrespective of counterbalancing every participant had to execute a push and a pull reaction for each of the 144 pictures, resulting in a total amount of 288 responses. Task completion took approximately 5 minutes. The aforementioned zooming feature was incorporated as a means to enhance perceptions of approach and avoidance (Rinck & Becker, 2007).

Alcohol Timeline Follow-Back

The Timeline Follow-Back (TLFB) interview was used to assess our dependent variable alcohol consumption. The TLFB is the most established method for the retrospective assessment of drinking behavior (Sobell & Sobell, 1992) and relies on several procedures to aid recall, such as using a visual calendar and key dates, in order to obtain precise drinking amount estimates for a given period. Participants' TLFB data was gathered on each of the three testing sessions in order to obtain all-encompassing drinking amount estimates from study enrollment till study completion as well as the 30-day period before hospitalization, which served as a baseline estimate of drinking behavior.

Secondary outcome measures and potential mediators

Several questionnaires were assessed at each of the three testing sessions in order to determine training-induced changes. Measures included the Fagerström Test for Nicotine Dependence (Heatherton et al., 1991), Beck's Depression Inventory (Beck, Steer, & Brown, 1996), State Anxiety Inventory (Spielberger, 1983), Barratt Impulsiveness Scale (Patton, Stanford, & Barratt, 1995), Brief Self-Control Scale (Tangney, Baumeister, & Boone, 2004), Behavioral Inhibition/Activation Scale (Carver & White, 1994), Positive and Negative Affect Schedule (Watson & Clark, 1999), Alcohol Dependence Scale (Skinner & Allen, 1982), Obsessive Compulsive Drinking Scale (Anton, Moak, & Latham, 1996), as well as the Alcohol Urge Questionnaire (Bohn, Krahn, & Staehler, 1995).

Data analysis

TLFB:

Daily drinking reports were converted into standard drinks and subsequently aggregated for each measure, yielding a baseline, a post-intervention, and a follow-up index of alcohol consumption. Outliers were not removed from the data, as long as reports lay within

biologically plausible ranges. Subsequently, a repeated measures analysis of variance (RM-ANOVA) with timepoint (post-measure/follow-up) as within-subject factor and group membership as between-subject factor was fitted. Baseline alcohol consumption and age were entered as covariates into the model.

Assessment AAT:

Raw data were cleaned from outliers by removing all RTs, which deviated more than 3 standard deviations from an individual's mean RT (irrespective of whether these were approach or avoidance trials). Subsequently, the median RT of avoidance trials was subtracted from the median RT of approach trials for both alcohol- and soft drink stimuli. Lastly, approach bias scores were obtained by subtracting the soft drinks from the alcohol score so that larger scores express a more pronounced approach bias for alcohol ((alcohol push – pull) – (soft drink push – pull)). Data inspection showed non-normally distributed data, as indicated by kurtosis and skewness and a significant Kolmogorov-Smirnov test with: D (75) = .201, $p < .0001$, D (75) = .174, $p < .0001$, D (75) = .145, $p < .0001$ on the first, second and third point of measurement, respectively. Log-transformation has been applied to render the data more normally distributed. Subsequently, a 2x3 RM-ANOVA with point of measurement as within-subject factor and group membership as between-subject factor has been executed. Again, age served as covariate in the model.

Results

Participants did not differ between groups in gender F = (2, 73) = 1.31, p = .277, and education F = (2, 72) = 1.705, p = .21. However, there was a significant between group difference in age with participants in the avoidance group being younger than participants in the other two groups F = (2, 73) = 6.442, p = .003 (AAT group: M = 44.89, SD = 7.54; inhibition group: M = 51.71, SD = 8.58; control group: M = 53.13, SD = 10.58).

The RM-ANOVA, investigating whether participants in the experimental groups consumed less alcohol over the course of the study, yielded no significant effects. Neither the main effect of time (F (1, 65) = 1.57, p = .22), nor the hypothesized interaction effect time x group was significant (F (2, 65) = 2.36, p = .10), indicating that alcohol consumption did not change significantly from post- to follow-up measure and that none of the groups drank significantly less than any other group over the course of the study (see Table 1 for descriptive statistics).

Table 1

Descriptive statistics of alcohol consumption and alcohol approach bias

	Alcohol consumption			Alcohol approach bias		
time point	AAT	Inhibition	Control	AAT	Inhibition	Control
baseline	378,66	310,48	302,92	-,12	-,13	-,12
post-test	6,87	3,96	1,15	-,08	-,09	-,12
follow-up	58,49	89,44	27,98	-,08	-,09	-,09

Note. Alcohol consumption in standard drinks and approach bias scores before log-transformation

RM-ANOVA showed no significant changes of alcohol approach bias over time F (1, 72) = 3.90, p = .052. More importantly, the hypothesized time x group interaction was non-significant F (2, 72) = .61, p = .55, indicating that none of the groups profited more from AAT-training than any other group when it comes to changes of alcohol approach bias.

Furthermore, none of the secondary outcome measures showed significant time x group interaction effects, indicating that AAT-training did not affect any of the assessed constructs over the course of the project.

Implications

The AAT has become an accepted treatment add-on in the realm of addiction (Ernst et al., 2014; C.E. Wiers et al., 2014; Wiers et al., 2011), but its working mechanisms remain hitherto poorly understood. The present project sought to improve this understanding by investigating whether response inhibition, as a predominantly reflective process, might play a more pronounced role than previously thought.

However, results lent no support to this notion, since neither conventional AAT-training nor inhibition AAT-training led to better treatment outcomes than sham training. Thus, no significant differences in alcohol consumption were observed between the groups, although the overall alcohol consumption decreased decisively over the course of the project. More precisely, a sharp reduction was observed from baseline to post-training. Since this decrease applied to all groups, it seems most likely that detoxification and abstinence-oriented CBT during hospitalization caused this reduction. After the pronounced decline from baseline to post-measure, alcohol consumption increased again across all three groups. Interestingly and against our expectations, patients receiving sham-training had consumed the least amount of alcohol at follow-up. Although this difference was not statistically significant, visual inspection of the data revealed that patients in this group were also the ones who had consumed the least amount before hospitalization, which serves as one explanation for the observed pattern.

Self-evidently, the search for potential mediators became unnecessary given that AAT-training did not improve treatment outcome. Nonetheless, analyses testing for between-group differences in alcohol approach bias over time were executed in order to evaluate whether automatic action tendencies were differently affected by a certain training. Yet, just like with drinking consumption, none of the trainings changed action tendencies more than any other.

Both, the longitudinal design, allowing to assess alcohol approach biases and other potential mediators over a longer period of time, as well as, having an extensive training of 63 sessions were strengths of the present project. The field of AAT research has come to agree that rather few testing sessions suffice to induce AAT-training effects. Besides mere pragmatic and economic reasons, this conception is mostly based on a study by Eberl and colleagues (2014), which found six to be the optimum number of training sessions. This study used changes in alcohol approach biases to derive the optimum number of sessions. However, since recent research has questioned the role of approach bias change in AAT-trainings (e.g., Cristea et al., 2016; Rinck et al., 2018), basing the number of training sessions on this criterion seems questionable, as well. Even though the present project used a more elaborate training procedure, the question whether participants profited from additional training cannot be answered, as basic AAT-training effects could not be replicated. On the contrary, the theoretical possibility that the failure to show any AAT-training effects was due to having more training sessions must be acknowledged. Thus, future research should compare differences in treatment outcome after experimentally manipulating the number of training sessions for more clear-cut conclusions.

Another potential reason for absent training effects is the rather small sample size of the current project. Sample size calculations were based on an inhibition study by Kühn and colleagues (2017), which successfully showed transfer effects from a response-inhibition training to a stop-signal task with an even smaller sample of elderly people and AAT studies with comparable sample sizes. For instance, approach biases and drinking consumption have been shown to decrease as a consequence of AAT-training in a sample of 42 hazardous drinkers (Wiers, Rinck, Kordts, Houben, & Strack, 2010) and 21 alcohol dependent patients were identified to possess larger alcohol-related approach biases than matched healthy controls (Ernst et al., 2014). In a nutshell, power problems are one possible explanation for the failure

to show any findings, yet other studies with comparable sample sizes have effectively demonstrated inhibition- and AAT-training effects.

The choice of our experimental design might have negatively affected dropout rates and by extension statistical power. As already mentioned above, we chose a longitudinal design, assessing outcome parameters on several occasions and in person rather than via telephone or postal follow-up. While such a design allows to track changes in potential mediators over time, a higher dropout rate can be expected. Future research should be aware of the implications and carefully weigh the pros and cons of each option, when selecting experimental designs.

Lastly, the absence of any observable training effects could be explained by the usage of a mobile version of the AAT, as there might be differences between mobile and conventional joystick versions of the AAT. Replications of basic approach and avoidance preferences for happy and angry faces, respectively, by means of a mobile AAT contradict this notion, however (Zech, Rotteveel, van Dijk, & van Dillen, 2020). Also, mobile AAT-trainings have previously proven effective in reducing alcohol consumption in problem drinkers (Laurens et al., 2020) and unhealthy food choices and approach biases in people with obesity (Kakoschke, Hawker, Catine, & Courten, 2018). Thus, empirical evidence suggests that mobile versions of the AAT can improve treatment outcome just like their joystick analog. Nonetheless, future research should address this issue by comparing both versions directly.

To conclude, no evidence for the involvement of response inhibition in AAT-training effects has been found. Also, basic AAT-training effects, such as decreased alcohol consumption and approach biases, were not replicated in a group receiving classical AAT-training. Reasons for this failure were discussed, the most notable being the rather small sample size. However, issues relating to sample size should not draw off the attention from the possibility that AAT-training in this project might have quite simply been ineffective at

decreasing alcohol consumption and approach biases. Critical voices concerning the AAT have accumulated in the last couple of years. For instance, a meta-analysis casted some doubts on AAT-training in the realm of addiction (Cristea et al., 2016). However, these doubts applied to approach biases only, and a small yet positive effect has been reported for treatment outcome parameters. This is in line with recent findings from the most extensive clinical trial, which found reductions in relapse rates, but no changes in approach biases (Rinck et al, 2018). Thus, according to the present state of the art, AAT-training can most likely improve treatment outcome, although this effect does not seem to be mediated by reductions of approach biases. This makes the search for potential working mechanisms even more important and future research should address this issue by looking for other mediators than changes of approach biases.

3.2 The AAT in Smoking Cessation: The Role of Response Inhibition and Psychometric Properties

Cigarette dependence is among the most frequent substance use disorders worldwide and the leading source of preventable death in Western societies (Begg et al., 2007). Most notably, smoking increases the chances of developing cardiovascular diseases, cancer, impotence/infertility, and respiratory diseases and thus places a huge burden on health care systems worldwide (US Department of Health and Human Services, 2006). Consequently, there is a high need for proper treatment. Historically, treatment approaches have tried to induce behavioral change via one or the other sort of rational argument, while more recent approaches also focused on changing automatic aspects of behavior by implementing training regimes.

Among such treatments AAT-training has shown the most promising results and hence received the greatest attention (e.g., Eberl et al., 2013; Sharbanee et al., 2014; Wiers et al., 2011). In the computerized training people learn to avoid objects related to their addiction by making pushing movements on a joystick device. Thus, training focusses on establishing motoric patterns of avoidance for smoking-related cues. To heighten perceptions of approach and avoidance the test uses a zooming feature, which downsizes or enlarges pictures upon pushing and pulling movements, respectively (Rinck & Becker, 2007). Findings in the realm of addiction research indicated that a relatively brief training of four sessions substantially reduced the relapse probability among patients with alcohol addiction who received AAT-training in addition to treatment as usual (Wiers et al., 2011). Subsequent research showed that AAT-training effectively assists in smoking cessation, as well (e.g., Machulska, Zlomuzica, Rinck, Assion, & Margraf, 2016).

Another effective form of computer-based training has focused on improving response inhibition by withholding prepotent responses to drug-related cues. In such Go/No-Go task trainings people are requested to respond to all except drug-related stimuli by instant button

presses. Since Go trials constitute a high percentage of trials (usually around 80%), participants will likely be in a sort of response automatism when drug-related stimuli appear on the remaining No-Go trials. Hence, response inhibition needs to be exerted to abort already initiated motor responses. Just like AAT-training, Go/No-Go task training was shown effective in decreasing problematic drug consume (Houben et al., 2011; Houben et al., 2012; Jones & Field, 2013).

An intriguing possibility is that both tasks, although different at the surface, might share a common process of response inhibition. Throughout the AAT literature, we often come across the notion that people with addiction build automatic approach preferences for stimuli related to their addiction and that AAT-training counteracts such tendencies by learning to avoid these stimuli. In Go/No-Go task trainings, in contrast, people learn to inhibit their urge to approach the drug by withholding responses. Hence, both have in common that people refrain from acting upon their urge to approach the drug. Interestingly, the same process of inhibition might play a role in both trainings, and AAT-training effects might depend on active avoidance to a lesser extent than we take for granted. In sum, (part of) the AAT-training effect might be grounded in successfully inhibiting the urge to respond to the drug rather than its active avoidance. To test this, we longitudinally compared a newly developed inhibition AAT-, a conventional AAT-, and a sham-training in their effectivity to reduce nicotine dependence.

The possibility of response inhibition involvement is particularly interesting given that the mechanisms underlying the effectiveness of AAT-training are still poorly understood. Most mechanistic research has concentrated on the possibility that approach and avoidance tendencies change concurrent with training. In support of this, initial addiction research showed that patients exhibit an approach bias for drug-related stimuli (e.g., Ernst et al., 2014; C.E. Wiers et al., 2014), and that increased levels of abstinence resulting from AAT-training were mediated

by declines in approach bias (Eberl et al., 2013; Sharbanee et al., 2014). Yet, later unsuccessful attempts to replicate the aforementioned mediational pattern between declines in approach bias and improved treatment outcome (e.g., Rinck et al., 2018) indicate that the story is more complex. Because implicit tests in general have been indicated to possess comparatively poor reliability (e.g., Enkavi et al., 2019), some researchers argued that low reliability might explain replication failures (Gawronski, Deutsch, & Banse, 2011). Yet, surprisingly little research has addressed the task's psychometric properties with a recent review identifying only four studies to do so (Zech, Gable, Van Dijk, & Van Dillen, 2022).

To bridge this gap, split-half and test-retest reliability data were obtained over the course of the present project with the former assessing the task's temporal stability within one measure and the latter across different measures. Knowing how well the task measures a construct is crucial for the interpretation of its findings. For instance, the fact that a recent study showed 58% of alcoholic inpatients to exhibit an approach bias for alcohol (Piercy, Manning, & Staiger, 2021) could indicate that individual differences in approach and avoidance behavior are a relatively poor indicator of alcohol addiction, but likewise a stronger relationship between both could have gone unnoticed because of low test-retest reliability. Hence, knowledge about the task's psychometric properties is indispensable for the correct interpretation of its findings. In addition, research has shown that the precision of AAT measurements differs across populations with reliability being highest in people with phobias (Loijen et al., 2020). To the best of our knowledge, the current project was the first to investigate the task's psychometric properties in the realm of nicotine dependence. Another important question concerns the comparability of PC- and mobile versions of the AAT, as the latter hold the potential of making measures accessible in the field and future research more flexible in terms of research designs. Therefore, both versions were directly compared.

To test our assumptions, the effectiveness of different AAT versions in smoking cessation was compared in a longitudinal project. 75 participants were randomly assigned to one of three groups, with one group receiving regular AAT-, one group inhibition AAT- and the other sham-training. Just like the previous project, the inhibition AAT-training omitted active avoidance gestures and asked participants to inhibit responses when drug stimuli were presented instead (for a detailed description see the methods section). Thus, the main aim of the project was to compare both trainings in terms of their effectiveness to decrease the severity of nicotine dependence. We hypothesized that compared to sham-training both avoidance- and inhibition AAT-training would be more effective. Moreover, differences between the avoidance and the inhibition AAT group were of particular interest as these could hint at shared and dissociable mechanisms. Because we do not rule out that actively avoiding drug-related stimuli is an important aspect of AAT-training, we hypothesized that participants receiving regular AAT-training would show slightly better treatment outcomes than the inhibition AAT group. Lastly, we hypothesized that results from a PC- and tablet-based AAT would yield comparable results when it comes to the sensitivity of approach bias measures and their psychometric properties.

Information on Procedure and Project

Participants

75 healthy participants between 18 and 55 years of age were recruited via flyers, newspaper-, and online advertisements. In the project ad we searched for people who want to stop smoking and people that have just stopped (maximally two weeks ago). Participants' mental health was inquired using a Web Screening Questionnaire for Mental Disorders (Meuldijk et al., 2017). Moreover, participants were excluded if they met any fMRI contraindications, were left-handed or were diagnosed with a neurologic disease.

Procedure and Research Design

All participants signed informed consent at the start of the project. We employed a longitudinal experimental design with random assignment to one of three experimental conditions (regular AAT-, inhibition AAT-, or sham-training). Importantly, the existence of other conditions was concealed, and our instructions guaranteed that all participants received the same information about the aim of the project. In the cover story they were told that the intervention seeks to change unconscious processes of nicotine dependence by responding differently to pictures with and without smoking content. The project entailed three points of measurement each separated by a four-week interval. Thus, the project started with a baseline testing, at the end of which participants were provided with a tablet device. After engaging in the training for four weeks, the post-measure took place and aimed to capture training-induced changes. Another four weeks later a follow-up measure sought to determine the longevity of potential effects and marked the end of the project. At the end of this appointment participants were debriefed and monetarily compensated for their participation with 100€. The project was funded by the European Research Council (ERC-2016-StG-Self-Control-677804) and approved by the Ärztekammer Hamburg.

Stimuli

The stimulus set consisted of 144 images (72 smoking-related and 72 control pictures). The smoking-related images showed people smoking and the control images were chosen to match these pictures as closely as possible by showing people with a pen in their mouth or brushing their teeth. All pictures were matched in size and displayed in front of a white background.

AAT-training

We used an AAT-training with explicit task instructions on a Samsung tablet device. All participants received a tablet device after the baseline testing and were instructed to train three

rounds per day (5 minutes each) during the four weeks till the next appointment. Participants in the regular AAT-training group were instructed to push away (avoid) all smoking related images and to pull towards themselves (approach) images unrelated to smoking. Likewise, people in the inhibition group were required to "approach" non-smoking control pictures. However, for smoking-related images they were requested to give no response and thus to inhibit their response. A zooming feature enlarged image sizes if a pull reaction was given and decreased image size if participants made a pushing or no response in order to enhance perceptions of approach and avoidance (Rinck & Becker, 2007). The control group was instructed to swipe smoking pictures to the left and control pictures to the right or vice versa (depending on counterbalancing). A subset of 60 images (30 smoke and 30 control stimuli) were used for the training.

Outcome measures and experimental tasks

Number of cigarettes smoked:

At each of the three appointments participants gave an account of the number of cigarettes they smoked within the last four weeks. This way we were able to compare the number of cigarettes smoked at baseline (four week period before the start of the project) with the number of cigarettes smoked after the training and at follow-up.

Nicotine dependence indexes:

Likewise, the severity of nicotine dependence was monitored using the FTND (Heatherton et al., 1991) and the CDS (Etter et al., 2003). Again, data was obtained from all participants at each of the three appointments.

Assessment AAT:

Automatic approach tendencies for smoking were measured using an implicit AAT with responses depending on the frame color of the presented pictures. Since picture color to response assignment was counterbalanced, half of the participants had to push away pictures with blue frames and pull towards themselves pictures with yellow frames, while the other half had to execute the exact opposite reactions for the same frame colors. Each of the 144 stimuli was presented twice (once with a blue and once with a yellow frame), so that a push and a pull response had to be executed for each picture. Task completion took approximately five minutes for each of the PC and the tablet-based version. Again, the order of the tasks was counterbalanced to control for unwanted and systematic order effects. Hence, half of the participants did the PC-based version first and completed the tablet-based version afterwards, while the other half performed the tasks in the reversed order. Once again, the essential zooming feature was incorporated into all tasks to prevent cognitive reinterpretation of approach and avoidance movements and to strengthen perceptions of approach and avoidance (Rinck & Becker, 2007). The PC version was programmed in Inquisit4 and the tablet version in c# with unity (https://unity.com/).

Go/No-Go task:

A Go No-Go task was used to track inhibitory control over the course of the project. Participants were instructed to react to all the numbers from 1 to 9 by pressing the space bar, except for the number 3, for which responses should be withheld. A short interstimulus interval (1150ms) and the fact that only 26 out of 234 trials were inhibition trials increased the likelihood that participants were in a sort of response automatism on the No-Go trials of the task. The number of erroneous button presses on such No-Go trials served as the outcome measure of inhibitory control, with more errors representing lower inhibitory control.

Data analysis

Number of cigarettes smoked and nicotine dependence indexes:

Number of smoked cigarettes, as well as FTND and CDS scores were entered into a 3x3 RM-ANOVAs with the within-subject factor time (pre/post/follow-up) and between-subject factor group (avoidance/inhibition/sham).

Assessment AAT: In a first step the data were cleaned from all reaction times (RTs) that deviated more than 3 standard deviations from an individual's mean RT. Subsequently, the median RT of avoidance trials was subtracted from the median RT of approach trials for both tobacco and control stimuli. The so-obtained values were subtracted once more to calculate approach bias scores for smoking ((tobacco push – pull) – (control push – pull)). Thus, larger values represent a larger smoking-related approach bias. Subsequently, 2x3 RM-ANOVAs with timepoint of measurement (pre/post) as within-subject factor and group membership (avoidance/inhibition/sham) as between-subject factor were executed on both the mobile- and the PC AAT data.

Go/No-Go task:

The number of commission errors was normally distributed and entered into a 3x3 RM-ANOVA to investigate whether participants in the experimental conditions made fewer mistakes than participants in the control group after the training. Once again, time point of measurement (pre/post/follow-up) served as within-subject- and group membership (avoidance/inhibition/sham) as between-subject factor.

Split-half reliability:

Spearman-Brown coefficients were calculated by comparing trials with odd and even trial numbers.

Test-retest reliability:

Correlations between tests and retests were computed for each version (tablet and PC) separately. Moreover, F-tests on intraclass correlational coefficients were calculated based on the assumption of consistency.

Results

Neither age F (2, 72) = .03, p = .968, gender F (2, 72) = .50, p = .606, nor education F (2, 72) = 1.22, p = .302 significantly differed between the three groups. The RM-ANOVAs comparing the trainings' effectiveness to decrease nicotine dependence generated a consistent picture. We found significant main effects of time for the number of cigarettes smoked, F (2, 144) = 32.25, p =.0001, FTND, F (2, 142) = 34.78, p = .0001, and CDS scores, F (2, 142) = 49.36, p = .0001, while all the hypothesized time x group interaction effects were non-significant with F (4, 144) = .60, p = .66, F = (4, 142) = .29, p = .882, F = (4, 142) = .83, p = .510 for number of cigarettes, FTND, and CDS scores, respectively. Hence, post-training severity of nicotine dependence lessened in all groups, but decreases were about equal for all three groups.

Although the absence of the hypothesized interaction effects made the search for potential mediators superfluous, we investigated whether smoking approach bias or response inhibition capacity (as measured by the number of commission errors in the Go/No-Go task) changed more in any of the groups. Again, RM-ANOVAs indicated that neither smoking approach bias nor the number of commission errors underwent any group specific changes with F (2, 70) = 1.52, p = .227 (tablet), F (2, 71) = .64, p = .533 (PC) for the former and F (4, 142) = .74, p = .568 for the latter. Also, the main effects for all three measures were nonsignificant (tablet AAT: F (1, 70) = .017, p = .987; PC-AAT: F (1, 71) = .95, p = .334; commission errors: F (2, 142) = .55, p = .58).

Overall analyses revealed low psychometric properties of the assessment AAT. Split-half reliability estimates were low with Spearman-Brown coefficients ranging from .19 to .45 for the tablet- and from -.39 to .26 for the PC version. Likewise, test-retest reliability in the control group was low with values ranging from -.265 to -.072 for the tablet- and -.089 to .166 for the PC version. Lastly, slightly more than half of the participants exhibited an approach bias for smoking at baseline (52.6% for the PC and 61% for the tablet version).

Implications

Because post-training nicotine dependence parameters did not differ between the groups it is hard to tell whether response inhibition plays a role in AAT-training. Although cigarette consumption and severity of dependence decreased over all groups, declines were about equal across the three groups (avoidance/inhibition/sham). Hence, it seems as if the intention to quit itself rather than any specific form of treatment reduced nicotine dependence. Because of the absence of training-specific effects the search for mechanisms underlying treatment effects became superfluous. Nonetheless, we investigated whether variables that we expected to mediate treatment effects changed more in any of the groups. However, neither response inhibition capacity nor smoking approach bias underwent more pronounced changes in any of the groups. In sum, our results could provide no evidence for response inhibition involvement in AAT-training, yet we still encourage future research to re-address this possibility rather than abandoning it prematurely. Obviously, observing typical AAT-training effects such as improvements in treatment outcome would be a prerequisite for meaningful group comparisons.

The poor psychometric properties of the smoking AAT cast serious doubts on the task's usefulness as a measure of individual differences. Both the observed split-half and test-retest reliability were low, indicating absent temporal stability within one measure and across different measures. Thus, rather than being a stable construct smoking approach bias seems to

be subject to pronounced fluctuations. A recent study indicated that accumulating multiple instances of bias measures can increase reliability by reducing the impact of measurement error (Zech et al., 2022). In this study healthy participants' avoidance bias for sad and disgusting faces was measured monthly over a period of seven month. It should be clear that such designs are not easily implemented in addiction research, however. Oftentimes someone's intention to quit and the concomitant willingness to receive treatment in order to break with maladaptive habits have a very limited time window. Hence, researchers in addiction research have much shorter intervals at their disposal to obtain multiple approach bias measures. In sum, repeating measures over the course of months is rather unrealistic in addiction research but acquiring multiple instances over a shorter period (say within several hours or days) seems a challenging, yet fruitful path for future research.

Good news is that our results indicated that mobile and PC-based AATs were about equally effective at identifying smokers, which gives rise to the possibility of acquiring multiple measures in the field without the need for having appointments in the lab. Interestingly, the percentages of smokers who exhibited an approach bias (61% for the mobile AAT and 52,6% for the PC-AAT) resembled recent numbers from alcohol addiction, where 58% showed an alcohol approach bias (Piercy et al., 2021). Thus, the precision with which bias scores can identify people with addiction seems to lie somewhere in between 52,6 and 61%. Self-evidently, such numbers do not qualify to base therapeutic decisions upon, and it should be clear that the term assessment AAT refers to the task's intent to measure differences in action tendencies rather than it being an established diagnostic tool.

A more general concern that arises is the need to refine the concept of approach and avoidance bias and what it reflects. As we have seen, smoking approach bias fluctuates considerably over time. Understanding what kind of ephemeral state-like aspects contribute to

bias scores is an interesting route for future research. For instance, it makes intuitive sense that urges to approach tobacco might be weaker if a cigarette has been smoked just before taking the test. Investigating whether this and other comparable hypotheses indeed hold true will help to sharpen our understanding of approach bias scores and what these encompass. In addition, it will help to tell apart situational influences and measurement error. For now, we should acknowledge that smoking approach bias fluctuates considerably for reasons we do not know. Corollary, we should refrain from interpreting individual differences in smoking approach bias if these factors remain unknown. Between-group comparisons are still informative though since time-varying influences will likely even out with greater numbers of participants.

When looking critically at the present project three main issues can be identified. First, the power with 75 participants is certainly improvable. While larger sample sizes are generally important and desirable, it should be noted that the data acquisition period of the current project lasted more than five years, indicating how tedious it can be to recruit participants with nicotine dependence who are just about to quit smoking. Thus, research striving for higher power critically relies on ample access to people in (sub) clinical populations who are willing to quit.

Secondly, the ratio between smoking-related and unrelated images (50/50) in the inhibition AAT can be considered a double-edged sword. Research design-wise this ratio was elegant because inhibition training differed in only *one* respect from conventional AAT-training (i.e., inhibiting responses to smoking-related stimuli rather than avoiding them), but at the same time such a ratio differs from conventional inhibition training, in which No-Go trials constitute a much smaller proportion (usually around 20%). However, since our goal was to determine whether response inhibition might be implicated in AAT-training, opting for an equal ratio was preferred to increase the similarity and therewith preclude alternative explanations.

Lastly, it must be acknowledged that our data might not have been ideal for test-retest purposes. Even though the control group received no real training, we cannot rule out that other factors such as the intent to quit might have affected bias scores. Therefore, future longitudinal research in the realm of addiction should acquire consecutive bias measures without any form of intervention and/or intent to quit. Such studies could tell whether our results painted a too dark picture of the smoking AAT's test-retest reliability.

To conclude, we could find no evidence that AAT or inhibition AAT-training was superior to sham-training when it comes to decreasing the severity of nicotine dependence. In addition, neither approach bias nor response inhibition changed more due to a specific training and split-half and test-retest analyses casted doubts on the psychometric properties of the smoking AAT. Encouragingly, comparable amounts of precision found between mobile and PC-based AATs argue for the equivalence of both versions and hence provide new flexibility for future research designs.

Kugler, D., Bremer, T., & Kühn, S. (2022, February 15). Decreased Rostral Cingulate Zone Cue-reactivity to Unhealthy Stimuli after Avoidance Training in Healthy Participants. https://doi.org/10.31234/osf.io/an8kq

3.3 AAT-Training Induced Neuronal Changes among Healthy Participants

Ever since its emergence the AAT has received considerable attention mainly for two reasons. First, the possibility of overcoming some of the intrinsic problems associated with the direct assessment of attitudes, such as social desirability, by indirectly assessing action tendencies sparked great interest (Rinck & Becker, 2007). Consequently, the task has been used in various fields of application and biased action tendencies were demonstrated in alcohol addiction (Eberl et., 2013; Wiers et al., 2011), cannabis abuse (Cousijn et al., 2011), anorexia nervosa (Veenstra & De Jong, 2011), nicotine addiction (Machulska, Zlomozica, Adolph, Rinck, & Margraf, 2015; C.E. Wiers et al., 2013), social anxiety (Heuer et al., 2007), depression (Radke, Güths, André, Müller, & de Bruijn, 2014), and phobias (Klein, Becker, & Rinck, 2011; Rinck & Becker, 2007). Secondly, and probably even more importantly, the AAT has proven effective as a treatment add-on with positive effects in alcohol addiction (Eberl et al., 2013; Manning et al., 2021; Rinck et al., 2018; Wiers et al., 2011) and heavy cannabis use (Jacobus et al., 2018).

Because AAT-research failed to identify stable cognitive mediators of AAT-training effects to this day, the investigation of neural substrates seems to be a promising approach, as it might generate hypotheses about cognitive processes involved in and altered by AAT-training. First research dedicated to the neural correlates of the AAT in people with alcohol use disorder showed the engagement of the orbitofrontal cortex while approaching and the dlPFC while avoiding alcohol stimuli during AAT performance (Ernst et al., 2014). However, due to employing fNIRS as a neuroimaging method, the search was restricted to higher cortical areas. In line with findings by Ernst and colleagues (2014), first fMRI research showed more pronounced approach-related mPFC activity in people with alcohol use disorder (C. E. Wiers et al, 2014; C. E. Wiers et al, 2015a). Moreover, subcortical structures such as the NAc were

shown to be implicated in approaching alcohol (C. E. Wiers et al, 2014). Lastly, AAT-training was shown to reduce approach-related activity in the mPFC (C. E. Wiers et al, 2015a) and alcohol cue-induced activity in the amygdala (C.E. Wiers et al, 2015b). Taken together, areas involved in reward and motivational processing such as the NAc and the (m)PFC seem to play a key role in approaching alcohol and activity in the latter can be reduced by means of alcohol avoidance training.

What almost all AAT-research has in common, though, is a strong focus on clinical populations with two main fields of application. Either the task is used to distinguish between clinical and healthy populations or AAT-training is used to achieve favorable health effects in clinical populations, e.g., people with addiction. This focus is also reflected in the four neuroimaging studies mentioned above, all focusing on patients suffering from alcohol addiction. In two of them results were based on differences in AAT-related brain activity between individuals with and without alcohol addiction, while the other two compared changes in brain activity in a group of individuals with alcohol addiction, receiving either AAT- or sham training. Thus, until now only two neuroimaging studies investigated AAT-training induced changes in brain function and those may have been specific to alcohol use disorder. Although research on this issue is still lacking, it seems plausible that reward-related structures, activated in individuals with alcohol use disorder, might play a minor role if any in healthy individuals. It, therefore, seems worthwhile to investigate AAT-training induced neural changes in a nonclinical sample in order to identify general rather than disease-specific mechanisms.

We suppose that a more basal brain network than in alcohol addiction is involved in behavior change among healthy participants. In a seminal study, Solarz (1960) showed the general human preference to approach positive and avoid negative stimuli. Over the course of the present project, we seek to induce such preferences by approaching and avoiding healthy

and unhealthy drinks, respectively. We assume that emotional value is inferred from approach and avoidance movements through a brain network consisting of motor areas and the amygdala. In this network actions are thought to be realized by motor areas while emotional value is thought to be assigned to the stimuli based on inferences about these actions (approach vs. avoidance gesture). Due to the fact that the amygdala has been shown to be involved in the AAT as well as in emotional learning (Everitt, Cardinal, Parkinson, & Robbins, 2003; Maren, 1999; Morris, Öhman, & Dolan, 1998), we consider amygdala activity to be a suitable candidate when it comes to assigning emotional value to stimuli during AAT-training. Even though empirical evidence failed to deliver support for the apparent link between motor areas and the AAT until now, we assume that motor-related brain areas will play a more pronounced role when establishing action preferences among healthy participants. On the other hand, brain areas such as the NAc and the mPFC have been theorized to possess key roles in addiction (Quintero, 2013), and are therefore expected to play a minor role if any in a healthy sample.

The decision to study AAT-training induced changes in beverage consumption is based on previous research, indicating the applicability of basic AAT principles in the domain of food. For instance, approach biases were found in people with obesity (Kemps & Tiggemann, 2015), a tendency for bulimic eating (Brockmeyer, Hahn, Reetz, Schmidt, & Friederich, 2015), high energy drink intake (Kemps et al., 2019), and high chocolate consumption (Meule et al., 2020). Likewise, AAT-training has been shown to be effective in reducing food intake, approach biases for palatable foods (Schumacher, Kemps, & Tiggemann, 2016), bulimic symptoms (Brockmeyer et al., 2015), as well as energy drinks (Kemps et al., 2019). To the best of our knowledge, the current project is the first to investigate AAT-training induced neuronal changes among healthy participants.

To this end, 62 healthy participants were randomly allocated to an AAT- or a sham training group. Beverage consumption in a bogus taste test and regular consumption during the two weeks after the intervention served as behavioral outcome measures. It was hypothesized that participants in the AAT-training group would consume more healthy- and less unhealthy drinks than participants in the control group. Furthermore, it was hypothesized that participants in the experimental group would develop approach preferences for healthy and avoidance preferences for unhealthy beverages after the training, while no such change was expected for the control group.

In order to assess AAT-training induced changes in brain function, cue-reactivity and Go/No-Go fMRI paradigms with photos containing healthy and unhealthy beverages were acquired both before and after the AAT-training. It was hypothesized that compared to the control group, amygdala activity in the experimental group would decrease for unhealthy and increase for healthy beverages. Similarly, cue-induced motor activity in the experimental group was thought to decrease for unhealthy and to increase for healthy beverages, with no changes expected for the control group. The same pattern of results was expected for the Go/No-Go task.

Information on Procedure and Project

Participants

62 healthy participants between 18 and 45 years of age were recruited via online advertisements. The Web Screening Questionnaire for Mental Disorders was used to assess eligibility (Donker, van Straten, Marks, & Cuijpers, 2009). Moreover, exclusion criteria were any fMRI contraindications and fructosemia, since some drinks used in the bogus taste test contained fructose.

Procedure and Research Design

Participants were randomly assigned to the experimental- or sham-training group at project start. They were led to believe that the aim of the project was to investigate whether computer game elements can be used to predict diet and brain structure. The project started with a cognitive-behavioral testing session, comprising questionnaires assessing dietary patterns, a diagnostic AAT, and a bogus taste test. Subsequently, a baseline magnetic resonance (MR) measure was acquired, prior to the 30-minute AAT-training (3 rounds 10 minutes each). Another MR measure followed the training in order to assess post-training changes in brain function. Potential post-training cognitive-behavioral changes (bogus taste test and diagnostic AAT) were acquired thereafter. Noteworthy, participants were offered water throughout the whole appointment to minimize the potential influence of thirst on the results. The re-assessment of beverage consumption two weeks after the appointment marked the end of the project (see Fig. 1 for a graphical representation of the project). Participants were debriefed and monetarily compensated with 45 € for their participation.

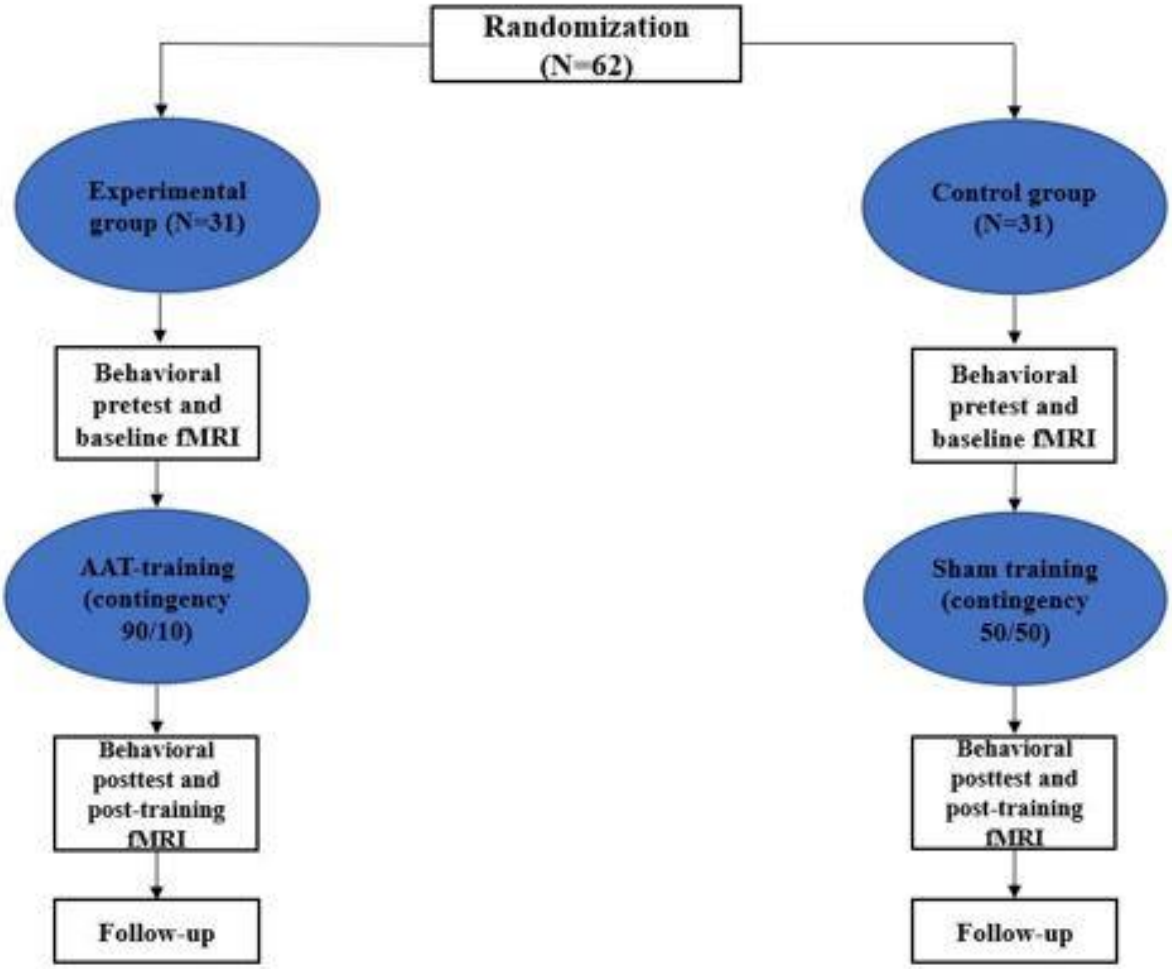

Figure 1: Graphical representation of study design.

Stimuli

The stimulus set consisted of 152 photos (76 healthy and 76 unhealthy beverages). The category of healthy beverages consisted of mineral water exclusively, whereas diverse soft drinks were chosen as unhealthy stimuli. All pictures were matched in size and displayed in front of a white background.

AAT-training

We used an AAT-training with implicit task instructions. The contingency between picture format and picture content was 90/10 in the experimental group, so that 90% of the healthy

beverages were presented in the format requiring a pull response and only 10% a push response. The reversed pattern was applied for unhealthy beverages (90% push and 10% pull). The control group received a sham-training with a contingency of 50/50, so that both healthy and unhealthy beverages had to be approached and avoided equally often. A subset of 60 images (30 healthy and 30 unhealthy stimuli) were used for the training.

Experimental tasks

Diagnostic AAT:

Automatic approach and avoidance tendencies for healthy and unhealthy beverages were measured by an implicit AAT (contingency 50/50) on a tablet device (IPad Air). 106 stimuli were presented once in portrait- and once in landscape format, thereby ensuring that both an approach and an avoidance gesture had to be performed for each stimulus. Again, the impressions of approach and avoidance were further amplified by increasing or decreasing picture size in response to approach and avoidance movements, respectively (Rinck & Becker, 2007). The task was programmed in unity (https://unity.com/) and completion of all 212 trials took approximately 4 minutes.

Bogus taste test:

The bogus taste test was employed as a dependent variable (Robinson et al., 2017). Three different beverages (Afri Kola, lemonade, and water) were provided to the participants, and they were asked to taste all of them and give an estimate of how much sugar each drink contained. The actual variable of interest, however, was the amount of consumption. This was measured by subtracting each drink's weight after tasting from its initial weight.

Diary report of beverage consumption:

Beverage consumption in the two weeks before and after training was assessed through a memory recall procedure resembling the widely used Food Frequency Questionnaires (Willett et al., 1985). If participants indicated that they had consumed a certain drink within the last two weeks, in-depth questions assessed the frequency and amount of consumption of it. The dependent variable was computed by multiplying the total amount of consumption for each drink (in ml) with its respective sugar content (per ml) and subsequently aggregating all indices to create an overall score of sugar consumption.

FMRI task description

Cue reactivity task:

106 (53 healthy and 53 unhealthy beverages) were presented to the participants during fMRI data acquisition. A subset of 60 stimuli was identical to those presented during the AAT-training, while the remaining 46 were not included in the training. Each picture was presented for 5 seconds and intertrial intervals were jittered between 3 and 5 seconds. Five trials, in which participants had to respond to a red fixation cross at the centre of the screen, were used to guarantee focussed attention throughout the task.

Go/No-Go task:

Participants were asked to respond to healthy beverages with a button press, while no reaction was required for unhealthy drinks. The proportion of healthy and unhealthy images presented was 80% and 20%, respectively. The stimulus set, duration of stimulus presentation, and the intertrial interval were identical to those in the cue reactivity task.

Data analysis

Assessment AAT: All reaction times (RT's) that deviated more than 3 standard deviations from an individual's mean RT were removed from further analyses. Approach bias scores were built by subtracting the median reaction times of pull from push trials for healthy and unhealthy beverages, respectively, so that larger values indicate approach biases, while negative values express avoidance preferences. Subsequently, 2x2 RM-ANOVAs with time point of measurement (pre/post) as within-subject factor and group membership (experimental/control) as between-subject factor were executed for both healthy and unhealthy beverages.

Bogus taste test and diary report of beverage consumption:

Kolmogorov-Smirnov tests showed non-normally distributed data for both outcome measures. Log-transformation was applied to render the data more normally distributed. Subsequently, 2x2 RM-ANOVAs with time point of measurement as within-subject factor and group as between-subject factor were fitted with the total amount of sugar consumption as a dependent variable for the retrospective report of beverage consumption and the consumption amount of healthy and unhealthy beverages as dependent variable for the bogus taste test. Body mass Index (BMI) was entered as a covariate in all models in order to control for its potential influence.

Scanning Procedure

Brain images were collected on a 3 Tesla Siemens Magnetom Skyra Magnetic Resonance Imaging (MRI) scanner system (Siemens Medical Systems, Erlangen, Germany) using a 20-channel radiofrequency head coil. The structural images were collected using a three-dimensional T1-weighted magnetization prepared gradient-echo sequence (MPRAGE) (repetition time (TR) = 2500 ms; echo time (TE) = 2.12 ms; TI = 1100 ms, acquisition matrix = 256 × 256 × 192, flip angle = 9°; field of view (FOV) = 240 mm, voxel size = 0.8 mm× 0.8

mm× 0.9 mm). Functional images were obtained using a T2*-weighted echo planar imaging (EPI) sequence sensitive to BOLD contrast (TR = 2000 ms, TE = 30 ms, image matrix = 64 × 64, FOV = 216 mm, flip angle = 80°, voxel size = 3 mm × 3 mm × 3.0 mm, 36 axial slices).

FMRI data analysis

Functional imaging data was analyzed using Statistical Parametric Mapping software package (SPM12). EPIs were corrected for slice timing and head motion and transformed into the stereotactic normalized standard space of the Montreal Neuroimaging Institute (MNI) using the unified segmentation algorithm. Finally, EPIs were spatially smoothed with a 3D Gaussian kernel of 8 mm full width at half maximum.

On the single subject level, the following event-related separate regressors were included for the cue reactivity task: healthy trained, unhealthy trained, healthy untrained, unhealthy untrained, response. Finally, the six rigid body movement parameters were also included in the single subject General Linear Model (GLM). We ran whole brain analyses testing for group by time interactions with changes assumed for the experimental- and stability for the control group (ordinal interaction). The results were corrected for multiple comparisons using a family-wise error threshold of $p < 0.05$.

Moreover, we ran region of interest (ROI) group by time by condition (2x2x2) analyses in brain regions for which we had a priori hypotheses, namely the NAc, mPFC, dlPFC, amygdala and brain regions related to response preparation at pretest namely left motor cortex (lMC) and rostral cingulate zone (RCZ). The lMC and RCZ were determined using pretest data across both groups contrasting the response in the cue-reactivity paradigm against the implicit baseline; after thresholding the whole brain contrast using a family-wise error threshold of $p < 0.05$, the two resulting clusters were chosen as ROIs. All ROIs were overlaid onto a normalized T1-weighted MNI template (colin27) and BOLD signals were extracted using marsbar (Brett

et al., 2002). We tested ordinal group x time interactions for all ROIs, with changes expected in the experimental group and stability in the control group.

Bonferroni correction was applied to account for this multiple testing problem and the associated alpha inflation. Yet, as Bonferroni correction has been criticised as overly conservative and since the two interactions of interest were not independent of one another we additionally conducted a false discovery rate correction (Benjamini & Hochberg, 1995).

Results

Behavioural outcome measures

None of the hypothesized behavioural effects was significant. Thus, there were no significant post-training group differences in beverage consumption in the bogus taste test (neither for healthy (F (1, 59) = 2.39, p = .127), nor for unhealthy beverages (F (1, 59) = 3.13, p = .082) or in regular consumption in the two-weeks after AAT-training with (F (1, 59) = .791, p = .37) for healthy and (F (1, 59) = .17, p = .68) for unhealthy beverages, respectively. Likewise, no significant post-training differences in action tendencies were found between the groups with (F (2, 60) = .99, p = .323) for healthy and (F (2, 60) = .001, p = .971) unhealthy beverages, respectively.

Lastly, RM-ANOVAs performed on the Go/No-Go task data that was acquired during the fMRI scanning procedure yielded no significant time x group interaction effects (neither when looking at reaction times of the Go condition (F (1, 60) = 0.388, p = .536), nor when testing for differences in the amount of correct responses in No-Go trials (F (1, 60) = 1.87, p = .179).

FMRI

Whole brain group by time interaction analyses did not disclose any significant clusters showing an ordinal interaction, neither when focussing on the contrast of healthy and unhealthy

stimuli that were presented during training, nor those which we intended to test potential transfer effects i.e., untrained stimuli in the cue reactivity task. Similarly, no effects were found for the Go vs. No-Go contrast of the second task.

Interaction analyses in each of the ROIs for the cue reactivity task yielded no significant effects in the amygdala (F (1,60) = 2.73, p = 0.10), the mPFC (F (1,60) = .21, p = 0.65), the dlPFC (F (1,60) = 0.03, p = 0.87), the NAc (F (1,60) = 0.2, p = 0.66), nor in the lMC (F (1, 60) = 0.021, p = 0.884). However, in the RCZ we observed a significant time x group x condition interaction (F(1,60)=4.09, p=0.048) with RCZ motor preparatory activity being reduced for unhealthy, not healthy drinks from pre- to post-test in the experimental group (t(30) = 2.70, p=0.011), but not in the control group (t(30) = 0.93, p = 0.358) and post-test group differences in the RCZ signal in response to unhealthy drinks (t(30) = -3.23, p = 0.002) (see Fig. 2 for graphical representations).

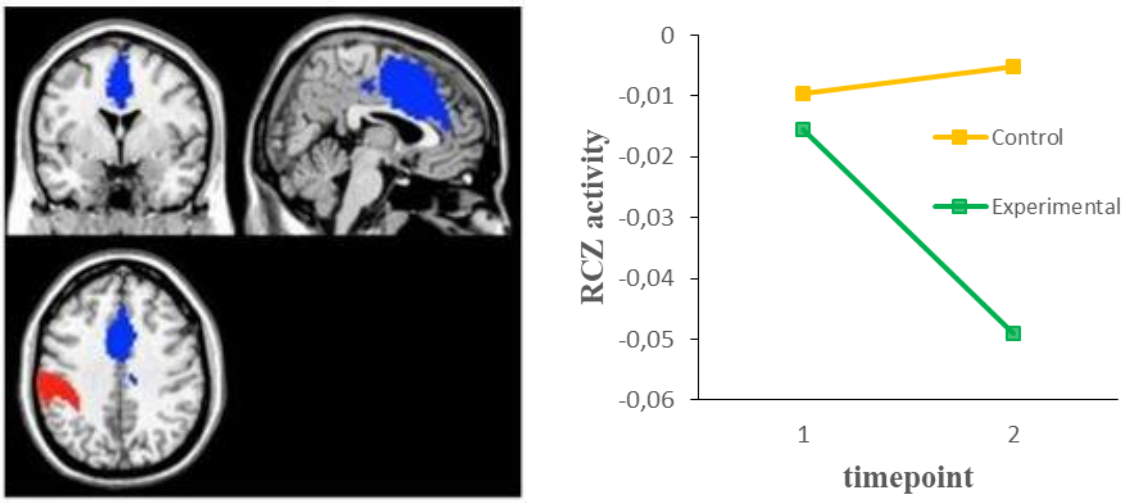

Figure 2: Left: Graphical representation of RCZ (blue) and IMC (red) ROIs.
Right: Change in RCZ bold signal as a function of time.

Implications

To the best of our knowledge the present project investigated AAT-training induced neuronal changes in a sample of healthy participants for the first time. Participants in the experimental group showed less cue-induced activity in response to unhealthy beverages in the RCZ after the AAT-training, while no changes were observed for healthy drinks. The RCZ is a fronto-median structure that is located anterior to the pre-supplementary motor area with pronounced projections to the motor cortex (Ullsperger & von Cramon, 2004). Activity in the RCZ has been associated with internally selected action (Mueller, Brass, Waszak, & Prinz, 2007) and the what-stream of intentional action, which is thought to play a key role in selecting appropriate actions (Brass & Haggard, 2008). Decreased RCZ activity among participants in the experimental group suggests that actions for soft drinks became less internally selected after avoidance training. Thus, the findings could be interpreted to indicate that consistently pairing unhealthy drinks with avoidance gestures led to relatively automatic action selection (i.e., avoidance) for unhealthy drinks. In line with this interpretation, Jocham and colleagues (2009) proposed the RCZ to be implicated in the establishment of action-outcome associations and coding the subjective value of actions. Clearly, the AAT requires monitoring action-outcome relationships across multiple trials and changes in RCZ reactivity suggest that this area was involved in observing such patterns. While Solarz (1960) showed the general human tendency to either approach or avoid stimuli based on affective evaluations, more recent research has shown that the relationship also works the other way around, i.e., stimulus evaluations were shown to change as a consequence of approach and avoidance movements (e.g., Van Dessel, Eder, & Hughes, 2018; Woud, Becker, & Rinck, 2008). This leads to another, yet more speculative, interpretation of decreased RCZ reactivity in response to unhealthy stimuli. Repeatedly pairing unhealthy drinks with avoidance gestures might have stimulated a devaluation process that assigns less subjective value to such stimuli. If this interpretation was

valid, future AAT fMRI-research should provide evidence for a relationship between altered RCZ activity and changes in implicit evaluations.

Integrating the present findings into the existing neuroscientific AAT literature is not trivial since all existing AAT-neuroimaging evidence stems from research on people with alcohol use disorder, while we focused on healthy participants. Even though no behavioral effects were found, fMRI findings for unhealthy stimuli suggest that AAT-training facilitated the inhibition and avoidance of motor preparation-related activity in response to unhealthy drinks rather than approach behaviour towards the more healthy option. This is in line with previous fMRI findings, showing decreased mPFC (C.E. Wiers et al., 2015a) as well as alcohol cue-induced activity in the amygdala after AAT-training (C.E. Wiers et al., 2015b). Both studies showed decreased activity in reward-related structures and in conjunction with our data seem to support the notion that AAT-training effects primarily manifest themselves by learning to avoid detrimental appetitive action alternatives. This interpretation would also explain why we failed to show changes in cue-induced activity for healthy drinks.

We could demonstrate no post-training changes in any of the four brain areas linked to the AAT by previous research (NAc, mPFC, dlPFC, and/or the amygdala) (Ernst et al., 2014; C. E. Wiers et al, 2014; C. E. Wiers et al, 2015a; C. E. Wiers et al, 2015b). Because mPFC and NAc are thought to play a key role in addiction (Quintero, 2013), we anticipated no decisive changes in those areas. However, based on findings showing less post-training alcohol cue-induced activity in the amygdala (C.E. Wiers, 2014), we would have expected decreased cue-induced activity especially for unhealthy stimuli. Even though such changes were not found, the possibility of amygdala involvement among healthy participants should not be repealed prematurely. We had no clear-cut assumptions concerning dlPFC involvement. While the majority of studies found no evidence for dlPFC involvement, first neuroimaging research

showed avoidance-related activity in both alcohol-dependent patients and healthy controls (Ernst et al., 2014), suggesting that activity was not only disease-specific. Yet, the data obtained from our healthy participants once more opposed the general involvement of the dlPFC in AAT-training.

The potential reasons why AAT-training may have failed to induce behavioral changes are complex. One reason may relate to the relatively dense testing procedure. It can be speculated that the fMRI post-measure, acquired after the training and before the assessment of the behavioral outcome measures, interfered with AAT learning effects or the translation of neural into behavioral changes. Another reason pertains to characteristics of our sample. The participants' average BMI of 23,35 was lower than in the general population and the concomitant healthy dietary pattern might have limited the variance in our outcome measures.

Several limitations of the present study must be acknowledged. As already mentioned above, it must be admitted that especially the dependent variable of regular beverage consumption during the two weeks after the training might have been affected by floor effects, resulting from healthy dietary patterns in our sample. Future research should come up with alternative cover stories, which omit stating that the project investigates dietary pattern, in order to not only attract people with a disposition for a healthy lifestyle.

To conclude, partial support for the hypothesized brain network, consisting of the amygdala and motor areas has been found. While no changes in amygdala activity were observed, we found decreased post-training activity in response to unhealthy stimuli in the RCZ. Future research should clarify whether changes in RCZ are only observed among healthy participants or whether these changes unfold in different psychopathologies also.

Kugler D, Ascone L, Pohlmann K, Weichenberger M, Kuehn S. Construct Validation of a German Version of the Word Completion Task for Aggression. PsyArXiv; 2022. DOI: 10.31234/osf.io/3dh5r.

3.4 Construct Validation of the Word Completion Task for Aggression

Human aggression has been defined as any purposeful and harmful behavior that is directed towards another individual that is motivated to avoid that harm (Anderson & Bushman, 2002). Almost all aggression theories distinguish between stable personality characteristics and situational components that may result in aggressive acts. According to the GAM these personality and situational factors in conjunction with the present internal state (i.e., current affect, cognition, and arousal level) determine whether a person will behave aggressively or not.

Historically, aggression has been measured predominantly by means of explicit measures such as self-report questionnaires. The drawback of these tests is, that participants can easily guess what they intend to measure, giving rise to the possibility of strategically faking answers or generating more socially desirable responses. Lately, implicit measures of aggression have become more and more popular (e.g., Greenwald et al., 1998). This shift in focus is closely linked to the ascent of dual-process models, which describe, explain, and predict behavior by juxtaposing a reflective and an impulsive system (e.g., Strack & Deutsch, 2004). While the reflective system is slow, rule-based, deliberate, accessible to conscious awareness and critically depends on cognitive resources; the impulsive system is fast, associative, unconscious and operates independent of the amount of cognitive resources available (e.g., Hofmann et al., 2008; Strack & Deutsch, 2004). In general, explicit measures display reflective aspects, whereas implicit measures are meant to capture more impulsive aspects of behaviors such as aggression.

Opinions on what exactly a test needs to qualify as implicit differ. Some have argued that tests can be labeled implicit if participants do not know what is being measured (Brunel, Tietje, & Greenwald, 2004), while others have ascribed the attribute to tests that do not rely on verbal reports (Fazio & Olson, 2003). De Houwer and Moors (2007) put forth a more elaborate definition. They equate implicit with automatic and define a list of automaticity features. These features include processes that are uncontrolled, unintentional, goal-independent, purely stimulus-driven, autonomous, unconscious, efficient, and fast. Importantly, a test does not have to fulfill all of these criteria, but rather can possess one or the other and still would be termed implicit. The famous Stroop effect, demonstrating that people have difficulties naming the color of presented letters of a color word when the presented color and the word-meaning do not match, serves as a good example (Golden, 1978). Word processing clearly occurs unintentionally, but still is non-automatic in the sense that attention needs to be allocated to the presented letters (i.e., the word). Thus, De Houwer and Moors' definition seems to enable a more nuanced and gradual assessment of the implicitness of measures.

A couple of implicit aggression measures have evolved over the past decades. Among these, the IAT is probably the most well-known (Greenwald et al., 1998). In this computerized test participants are asked to categorize presented words referring to the concept of self or others and aggressive or peaceful words. Responses from block(s) that require the same response for aggressive and self-related words and block(s) calling for the same response to aggressive and other-related words build the basis for the computation of IAT scores. The assumption is that responses, that require the same response key, will be facilitated to the degree to which the two categories are associated within the semantic network of the participant. Thus, comparatively faster responses in the block in which aggressive and self-related words require the same response, are taken to be indicative of an implicit tendency for aggression.

Other implicit measures of aggression are the conditional reasoning measure for aggression (James et al., 2005) and the lexical decision task (LDT) (Meyer & Schvaneveldt, 1971). The former presents conditional reasoning problems to the participants and purports to measure inductive reasoning skills on the surface. In fact, some of the answers to each item are based on implicit assumptions (e.g., a hostile attribution bias) that aggressive individuals tend to possess, and the number of such responses is conceptualized to reflect levels of implicit aggression. In the LDT, participants are asked to indicate whether a presented series of strings builds a word or not. However, a proportion of the words have an aggressive meaning and the reactions times to such words are used as a reflection of implicit aggression with the assumption being that more aggressive individuals will react faster to aggressive words.

The word completion task (WCT) has been used as another implicit measure of aggression (Anderson et al., 2003). Here participants are asked to build words by filling in the blank spaces of each word stem. For example, the word stem KI_ _ could be completed to build the aggressive word KILL or the non-aggressive word KIND. The proportion of total aggressive solutions given by the participants are used as a measure of the accessibility of aggressive cognitions.

Up until now, the WCT for aggression lacks a rigorous validation. As correctly pointed out by Benjamin and McCarthy (2019), the WCT was validated through a procedure that resembles a Penrose staircase, with two studies each referencing the other as evidence for its validation (Anderson, Carnagey, & Eubanks, 2003; Anderson et al., 2004). In a subsequent third study the WCT was then presented as a valid measure of aggressive cognitions (Carnagey & Anderson, 2005). Despite this shortcoming, all studies successfully demonstrated that participants build more aggressive solutions on the WCT if they were exposed to violent media beforehand and several studies have generated similar findings ever since (e.g., Barlett &

Rodeheffer, 2009). The logic inherent in the designs of all these studies is that the WCT seems to reflect state aspects of aggression or what would be subsumed under the term situational component in the GAM. Even though differences in aggressive word completions between groups being exposed to violent and nonviolent media suggest that the WCT measures aggressive states, this does not rule out the possibility that scores also express trait aspects of aggression. The inclusion of trait hostility as a covariate into their analyses shows that Anderson and colleagues (2003) had the assumption that WCT scores are affected by trait aggression and was confirmed by disclosing a relationship between both in one of two studies. In fact, it seems very obvious that if the WCT actually measures the accessibility of aggressive cognitions, then it should also discriminate between highly aggressive and nonaggressive individuals, as aggressive individuals in the GAM are theorized to possess comparatively more aggressive cognitions.

Within the scope of the current project, we present a German version of the WCT and use both state and trait measures of aggression to test its construct validity. To this end, 88 participants completed the WCT as well as several well-known state and trait measures of aggression. We assume that aspects of trait as well as state aggression are captured in the WCT.

Moreover, we present a new, more sophisticate scoring method that goes beyond the conventionally used proportion of aggressive solutions. Rather, the new method considers the number of alternative possible solutions and the word frequency of each solution. For instance, the word stem EXPLO_E has only two solutions to it (i.e., it can be completed to build the nonaggressive word EXPLORE or the aggressive solution EXPLODE). We assume that an aggressive solution to this item will be less indicative of aggression than an aggressive answer to an item that has only one aggressive solution among say 20 non-aggressive alternative solutions. Thus, the smaller the proportion of aggressive to non-aggressive solutions for an

item, the more weight is given to an aggressive answer based on the proposed scoring methodology. Moreover, accounting for how often solution words occur in the test language could further increase the reliability of the WCT. Aggressive solution words that are rarely used within the test language will be treated as being more indicative of aggression than other aggressive solutions that are very frequently used in the language.

Information on Procedure and Project

Participants

The data of the present project were baseline data from a large-scale study, investigating the effects of long-term exposure to violent videogames (Kühn et al., 2019). In this context, 88 healthy participants were recruited via flyers and internet advertisements. Participants had an average age of 28.34 (*SD*: ± 7.69, min = 18, max = 51) with 46 and 42 female and male participants, respectively. In the recruitment ad of the project, it was relatively vaguely searched for volunteers in a longitudinal study on video gaming, this way guaranteeing that participants were unaware of the fact that we looked for effects related to aggression. The fact that participants were told only after the baseline testing which game they had to play, as well as the abundance of different tests acquired in between the aggression measures increased the probability that participants were in fact unaware of the hypothesis being tested. The project was approved by the local ethics committee of the Charité University Clinic, Germany and all participants signed informed consent. At the end of the large-scale intervention study participants were debriefed and monetarily compensated for their participation.

An additional sample of 40 healthy participants was recruited via online advertisements in order to obtain aggression indices for all the solution words of the WCT. Participants in this separate online survey were 28.72 years of age on average (*min* = 18, *max* = 62, *SD* = 9.38) with 58.1 and 41.9 percent being female and male, respectively.

Procedure

All participants completed the Buss-Perry Aggression Questionnaire (BPAQ) (Buss & Perry, 1992), the State Hostility Scale (SHS) (Anderson, Deuser, & DeNeve, 1995), the Revised Conflict Response Questionnaire (RCRQ) (Richardson & Green, 2003), the Rosenzweig Picture Frustration Test (RPFT) (Rosenzweig, 1945), the German version of the WCT, and the LDT in the order just mentioned. All tasks were separated by other questionnaires and experimental tasks included in the larger project.

Measures

Buss-Perry Aggression Questionnaire:

The BPAQ (Buss & Perry, 1992) is a 27-item scale encompassing measure of trait aggression, which captures four different aspects of aggression, i.e., physical aggression (BP_PA), verbal aggression (BP_VA), anger (BP_A), and hostility (BP_H) that can be subsumed to build a total score.

State Hostility Scale:

The SHS (Anderson et al., 1995) assesses state aspects of aggression. More precisely, participants are presented with 35 emotions and they are asked to indicate the extent to which they feel the respective emotion in the very moment on a 5-point Likert scale. The answers are used to build four subscales (sociable, lack of positive feelings, aggravation, and mean).

The Revised Conflict Response Scale:

In the RCRQ (Richardson & Green, 2003) participants are asked to recall all the instances during the last month in which other people said or did something that made them feel angry. Subsequently, questions assess how participants reacted and whether they were more inclined to respond with direct or indirect aggression.

The Rosenzweig Picture Frustration Test:

The RPFT (Rosenzweig, 1945) presents pictures of 24 social encounters to the participants. In each scene participants are required to respond to what another person said by filling in blank speech bubbles. Usually something unfortunate or bad is happening to the participants and the test is meant to test their level of frustration tolerance. Responses were rated along three different types (i.e., ego-defense, obstacle-dominance, and need-persistency) and three directions of aggression (i.e., extra punitive, intra punitive, and impunitive). Aggregate scores were computed for all six aggression indices.

Lexical Decision Task:

The LTD was used as an implicit measure of aggression (Meyer & Schvaneveldt, 1971). In this test participants are required to decide whether presented strings of letters constitute existing words or not. Among the existing words some were aggressive, and reaction time differences between these aggressive and non-aggressive words are conceptualized to reflect levels of implicit aggression. Hence, the outcome measure was computed by subtracting the mean RT for aggressive words from the average RT for non-aggressive words, so that larger values indicate a higher amount of implicit aggression.

Word Completion Task:

A German version of the WCT was developed for the current project. In total the test includes 100 word stems out of which 75 have (partially) aggressive solutions to them. Rather than defining what counts as an aggressive solution, all solution words generated over the course of the project were rated on their aggression level in a separate online survey by 40 raters. Based on these ratings, all words were classified as either being aggressive, partially aggressive (in the case of ambiguity), or non-aggressive.

Data Analysis

Aggregate scores were built for the BPAQ, SHS, and RCRS and all the subscales belonging to the respective scales. The RPFT was scored by three different raters to increase validity through interrater reliability. In case of disagreement answers were discussed and a majoritarian decision was made.

Conventional WCT score: The scores were computed by summing all aggressive solutions and dividing the count by the total number of valid solutions given, i.e., how many real words were built during task performance.

The computation of the advanced WCT scores relies on two calculations at the item-level. First, the number of aggressive and non-aggressive solutions to each item was assessed based on all solutions that were acquired over the course of the entire large-scale project. Subsequently, we computed the reciprocal of the proportion of aggressive to total solutions. For illustrative purposes let us assume that two items have three aggressive solutions each, while one has eight and the other only two non-aggressive solutions. Thus, the proportion of aggressive to total solutions would be three to 11, while this ratio would be three to five for item two. The obtained reciprocals (rec) are:

$rec(\frac{3}{11}) = \frac{11}{3} = 3.66$ and $rec(\frac{3}{5}) = \frac{5}{3} = 1.66$

Thus, more weight is given to aggressive solutions when there are more alternative non-aggressive solutions and/or fewer aggressive solutions since aggressive solutions are less likely qua statistics in both cases. After the computation of reciprocals for every item, the obtained inverse values were rendered more normally distributed through log-transformation and subsequently transformed into Z-scores.

Secondly, word frequencies for each possible solution were determined using the online search engine Google Ngram, which uses a large corpus of German books to chart word frequencies. Frequencies of all solution words were added up at the item-level to build an item total. The frequency proportion of the respective aggressive solution to the item total was the basis for the computation of reciprocals. Again, taking the inverse values ensured that aggressive answers received more weight if they were relatively less frequent in language use than the alternative solutions. The reciprocals of all aggressive solutions were normalized through log-transformation and then Z-transformed.

Advanced WCT score: This alternative index was computed by multiplying both the number of solutions and frequency scores with 1 in case of an aggressive solution and with 0.5 if the answer was partially aggressive. The final score was obtained by summing the scores from all items and dividing this sum by the number of valid solutions.

Pearson correlations were computed for both WCT scores and all the established aggression measures mentioned above.

Results

The conventional- and the advanced WCT score were positively correlated ($r(86) = .705$, $p > .01$). Moreover, we found significant inverse correlations between WCT scores and the obstacle dominance and the intra punitive subtype of the RPFT. Correlational coefficients for the obstacle dominance subtype were ($r(69) = -.262$, $p = .027$) for the conventional and ($r(69) = -.294$, $p = .013$) for the advanced WCT score, respectively. Likewise, the intrapunitive scale was inversely related to conventional and advanced WCT scores with the coefficients being ($r(69) = -.291$, $p = .014$) and ($r(71) = -.288$, $p = .015$), respectively.

Interestingly, only the advanced WCT score was significantly related to aspects of trait aggression. More precisely, we found a positive relationship with the hostility aspect of the

BPAQ, ($r(85) = .226$, $p = .035$), suggesting that the advanced score captures aspects of aggression that the conventional score does not.

WCT scores were neither correlated with state-, indirect-, direct-, or implicit aggression, nor with participants' tendency to respond to conflicts (for a detailed description of all the correlational coefficients, see heatmap in Fig. 1).

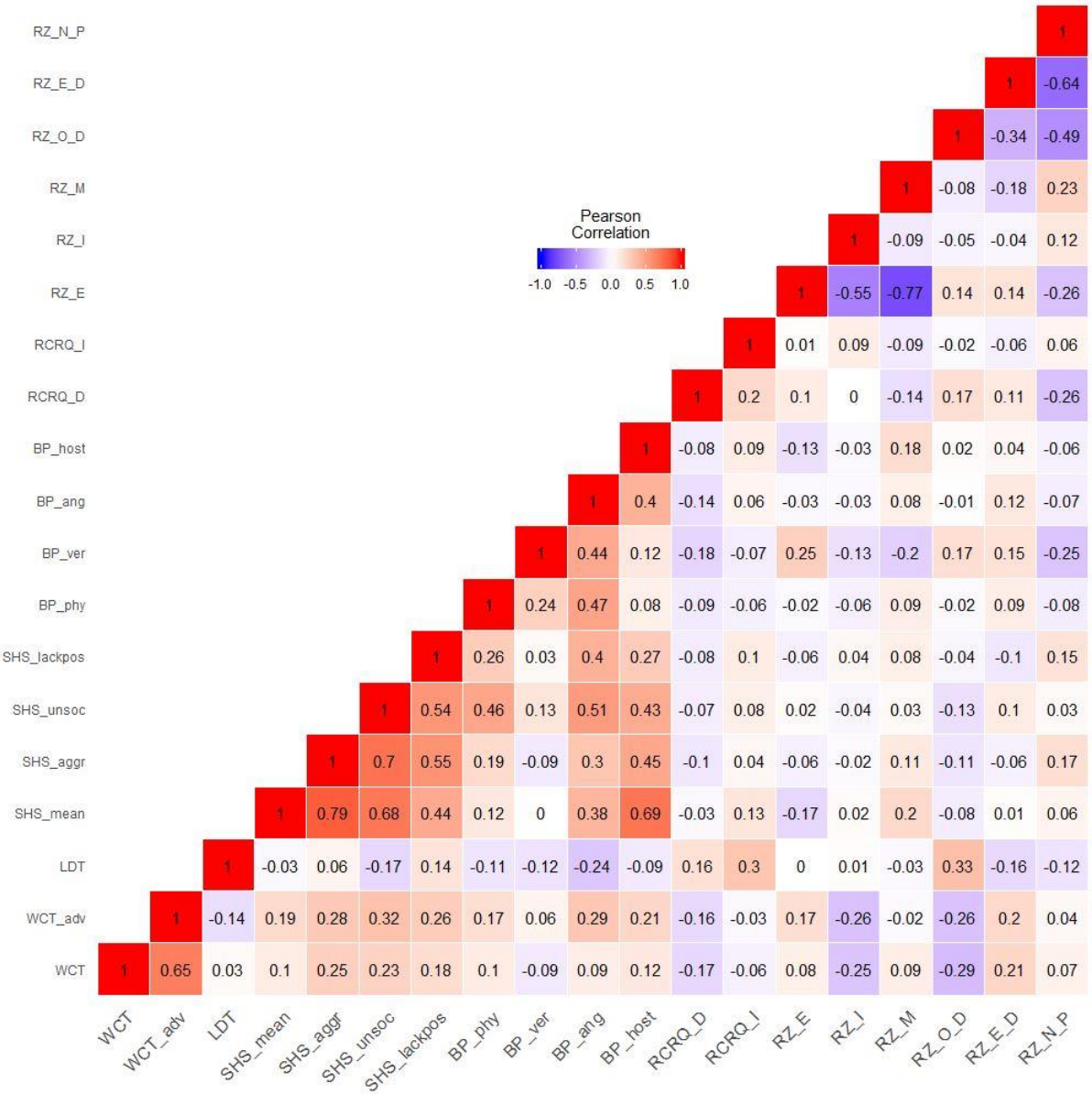

Figure 1: The Pearson correlational matrix of all aggression indices. Higher negative correlations in blue und higher positive correlations in red.

Implications

Even though the WCT for aggression is a commonly used tool in aggression research, it has been lacking a thorough validation. Moreover, its usage was restricted to English speaking countries thus far. Therefore, we developed a German version and set out to formally test its construct validity as an aggression measure. Among the measures used for validation purposes were well-known questionnaires of trait and state aggression as well as a measure assessing direct and indirect forms of aggression from recalled real-life experiences. In addition, verbal reactions to slightly provocative and frustration-inducing social encounters were assessed as a measure that relies on introspection to a lesser extent. Lastly, we included an implicit aggression measure in an attempt to capture automatic and unconscious aspects of aggression, as well. Besides validating the WCT, we also tested whether additional aspects of aggression are captured by a new more in-depth scoring methodology, which weighs answers based on the number of alternative solutions and the frequencies of the solution words.

In accordance with our assumptions, we were able to demonstrate that WCT scores are related to trait aggression. Interestingly, the observed relationship with the hostility aspect of the BPAQ pertained to the advanced WCT score only. This lends support to the assumption that the new score can yield more precise aggression estimates than the classic WCT score. While the classic score is built by summing up all aggressive answers and dividing them by the number of items that participants responded to, the new score goes beyond this by weighing the answers based on the word frequencies and the number of alternative solutions. Even though evidence from the current project suggests that the new score might be superior in capturing aspects of trait aggression, more research is certainly needed for final conclusions in this regard. Nonetheless, our first results are promising and suggest that the implementation of the advanced scoring methodology might not only be beneficial for the German- but also for the English version.

Furthermore, we demonstrated relationships between aggressive word completions and verbal response tendencies in provocative social encounters, i.e., both WCT scores were inversely related to the intrapunitive aspect of aggression of the RPFT. This dimension measures aggression that is directed inwards, against oneself, and overt. Hence, the more aggressive participants answered on the WCT, the lesser the amount of internally directed aggression they showed and vice versa. Likewise, an inverse relationship with the obstacle-dominance type of aggression was found, which has been described as the tendency to direct attention to the frustrating situation of the presented social encounter. Higher scores on the WCT went hand in hand with a decreased tendency to engage in obstacle-dominance and vice versa. While the tendency of directing less aggression against oneself is theoretically plausible, the inverse relationship with obstacle-dominance is puzzling, since one would expect aggressive individuals to have a greater rather than smaller focus on frustrating situations.

Lastly, no significant relationships between aggressive word completions and the other aggression measures could be shown. Neither implicit, indirect nor direct forms of aggression were significantly related to aggressive cognition. Most notably, state aggression was not associated with the WCT, either. However, since the WCT has been predominantly used as a measure sensitive to changes in state aggression, we would have expected a significant relation between the SHS and the WCT. Taking a closer look at the SHS state aggression data revealed that most participants indicated to experience no aggressive emotions. This floor effect and the concomitant low variance in the data most likely explain the absence of a significant relationship. Potentially, the standardized and calm testing situation in the lab prevented participants from reporting strong emotions. We would, therefore, not generally rule out the possibility that the German version is sensitive to changes in state aggression just like the English version, but rather encourage to address this question in future experimental research. Studies could either prime participants with aggressive content to compare WCT scores

between participants that were and that were not primed, or more elegantly could program an aggressive game, in which the number of opponents that one can be aggressive against can be controlled. In such studies, one would expect aggressive word completions to ascend with increasing numbers of aggressive acts during game play.

Evaluating whether or not the results apply to the English version, too, is complex and speculative. It is obvious that both versions can only rely on the same basic idea but are necessarily distinct due to differences in the respective language. We, therefore, strongly recommend a proper validation of the English version to see whether both versions tap into the same aspects or not.

In summary, we have presented a German version of the WCT and a new more in-depth scoring methodology. Moreover, we demonstrated first evidence that using the new scoring formula might possess additional value in assessing trait aggression. Furthermore, the amount of aggressive word completions was related to verbal response tendencies in provocative social encounters. Hence, some first important steps in validating the WCT as an aggression measure were made, but in light of the low overall associations, critical questions arose as well. For instance, we could not show the WCT to be related to state aggression. Since the English version has been predominantly used to assess changes in state aggression, we encourage the use of research designs that allow for more direct comparisons. Likewise, we strongly recommend to formally validate the English version and to implement a scoring methodology analogue to ours.

Kugler, D., & Kühn, S. (2022, March 11). No Effect of Ego Depletion on Cognitive Control and Implicit Aggression. https://doi.org/10.31234/osf.io/j3qra

3.5 Stifled Emotion and Ego Depletion

The view that the constant engagement of self-control leads to a diminished ability to exert this very capacity shortly thereafter has been probably the most popular social psychological finding of the last three decades and the first 20 years of its history read like a fairy tale almost too good to be true. Baumeister and colleagues (1994) first introduced the concept and presented initial evidence for its applicability by showing that utterly different tasks relying on the ability to exert self-control, such as forcing oneself to eat radishes instead of tempting cookies, making a meaningful personal choice, or suppressing emotions all compromised the performance on a subsequent task that drew off the same capability (Baumeister et al., 1998). To this day, more than 600 studies showed depleted self-control resources to affect behaviors as diverse as eating, drinking alcohol, physical exercise, aggression, test performance, smoking, drug use, cheating, empathy, racism, and marital infidelity (Baumeister & Tiarney, 2011; Inzlicht & Friese, 2019), leading to the proposition of a strength model of self-control. According to this model, the exertion of self-control depletes limited self-control resources just like a muscle gets tired from exertion (Baumeister, et al., 2007). Importantly, the adverse effects of ego depletion did not pertain to lab settings only, but also were shown in real life contexts in the health, and educational sector (Dai, Milkman, Hofmann, & Staats, 2015; Persson, Barrafrem, Meunier, &Tinghög, 2019; Sievertsen, Gino, & Piovesan, 2016).

However, the past decade predominantly produced findings, challenging the veracity of ego depletion. First doubts arose from studies showing that motivating participants to excel at the second task of the sequential-task paradigm, be it through kind instructions given by the researchers (Muraven, Gagné, & Rosman, 2008) or monetary incentives (Luethi et al., 2016), extinguished ego depletion effects. These findings contradicted propositions from the strength

model, since the prior engagement of self-control should lead to a state of ego depletion in a later task irrespective of any extrinsic or intrinsic motivations (Baumeister et al., 2007). More seriously, the disclosure of publication bias raised concerns about the effect as such (Carter & McCullough, 2014), since the hitherto accepted medium to large effect size (Hagger et al., 2010) shrank to small (Carter & McCullough, 2014) or even indistinguishable from zero (Carter et al., 2015) when correcting for publication bias and applying more narrow inclusion criteria. Yet another meta-analysis, using the p-uniform method to correct for publication bias, still indicated moderate effect sizes (Blazquez, Botella, & Suero, 2017) and critically highlights how the drawn inferences depend on the chosen meta-analytical method. Taken together, publication bias exaggerated our belief in the ego depletion effect and the extent of reduction depends on the method chosen for correction. Clearly, each correction technique relies on its own assumptions and having unbiased data in the first place is always preferable to relying on correction techniques. A straightforward and powerful way to prevent publication bias is the preregistration of planned research, which fortunately became more and more common over the last couple of years.

Unfortunately, the field of ego depletion research faces problems that are more complex to deal with. First, the measures used to induce and measure states of ego depletion were not formally validated on their eligibility to do so. Rather any combination, leading to decrements in performance on the second task, was taken as evidence that both tasks rely on self-control. Secondly, and somewhat intertwined with the last point, the interdependence between both tasks is almost always overlooked in the propositional logic of ego depletion research. More precisely, ego depletion research is based on evidence derived from so-called sequential-task paradigms, in which a first task is used to deplete and another to measure self-control capacity. In such a design, any inferences drawn will always be the result of the interplay between both tasks. The neglect of this interdependence in conjunction with the double-edged sword of high

applicability and usage of non-validated tasks in the sequential-task paradigm lies at the core of the chaos that is currently to be disentangled and often gets in the way of a proper scientific discussion. To resolve this chaos, the field needs to identify tasks that validly induce and measure decrements in self-control. Next, the fact that inferences will always depend on the combination of tasks used in the sequential-task paradigm needs to be internalized and highlighted. While such an implementation seems cumbersome in the first place, it will advance discussions in the long run since pairs of tasks that can and that cannot be reliably used for the assessment of ego depletion will become identifiable.

The disclosure of publication bias in ego depletion research led to a polarized discussion centering around the question whether ego depletion exists or not. Even though no single study alone can provide a final answer to this question, many research projects drew such far-reaching conclusions. For instance, null results obtained from a multilab preregistered replication study were interpreted as evidence favoring the most conservative meta-analytical model (Carter et al., 2015), since both converged most closely in terms of effect sizes, letting the authors conclude that other meta-analytical techniques inflated results in favor of ego depletion (Hagger et al., 2016). Such inferences are improper, however, because apples are compared with oranges. Clearly, a model comprising the whole variety of a) (non-validated) tasks used to induce a state of ego depletion and b) (non-validated) tasks used to measure this state subsequently cannot be directly compared to a single (multilab) study where arbitrary choices have to be made for one or the other primary- and secondary task. What such large-scale studies will confidentially tell us, however, is whether the selected combination of tasks is effective in inducing and measuring states of ego depletion. While the answer should be no in the aforementioned study from Hagger and colleagues (2016), other studies showed small to medium sized effects when using the Stroop task to deplete and the antisaccade task to measure self-control resources (Dang, Liu, Liu, & Mao, 2017; Dang et al., 2021). From today's

perspective we know that Hagger and colleagues (2016) null findings most likely stemmed from using the letter cancellation task, which has been shown incapable of inducing decrements in self-control in the meantime (Wimmer, 2019).

Consequently, we laid special emphasis on selecting appropriate measures for the sequential-task paradigm. The choice for emotion suppression videos as our depletion task was based on a meta-analysis, identifying this task as the most consistent inducer of ego depletion effects (Dang, 2018). Moreover, we opted for several outcome measures rather than just one. We expected impairments in a classic cognitive multi-source interference task (MSIT), as it was shown to be a sensitive outcome measure recently (Dang, 2016). We also looked for spillover effects onto implicit aggression. Several studies have shown dual-process theory conform elevations of aggression levels after the depletion of self-control resources (e.g., Barlett et al., 2016; Osgood & Muraven, 2015; Osgood & Muraven, 2016; Schmidt, Zimmermann, Banse, & Imhoff, 2015; Stucke & Baumeister, 2006), with one study showing that stifling emotions during an emotion video increased subsequent aggression (Vohs, Glass, Maddox, & Markman, 2011). This rather ancillary branch of ego depletion research might be subject to publication as well, and the ostensible relationship between ego depletion and aggression should be put under scrutiny. Hence, we hypothesized that participants having to suppress their emotions while watching a video of surgeries would subsequently exhibit more implicit aggression and perform worse on the MSIT than participants who watched the same video without having to suppress their emotions. Showing effects in two seemingly unrelated tasks would be more convincing evidence in favor of ego depletion than compromised performance on a second task in a sequence of two relatively similar tasks (Dang et al., 2017; Dang et al., 2021). It should be noted that a multilab large-scale study was beyond the scope of our preregistered project. Although the standardized study protocols of multilab large-scale studies are clearly advantageous in terms of comparability, it seems elusive that all future

evidence will be exclusively obtained from such projects. Rather, multi-site large scale projects should be used in conjunction with many smaller, thought-through experiments with a strong focus on selecting appropriate tasks in the sequential-task paradigm. The so-obtained evidence will enable future meta-analyses to resolve the question whether ego depletion is real or not in the long run.

Information on Procedure and Project

Participants

60 healthy volunteers aged in between 19 and 40 years were recruited via online advertisements and received 7,5 € for their online participation. Participants' eligibility was assessed using a web screening questionnaire (Donker et al., 2009). Moreover, people working in medical professions were not allowed to participate since they would most likely experience no strong emotions such as disgust while watching videos of surgeries.

Procedure and Research Design

The project was preregistered at Clinical Trials (NCT04879719) and approved by the local psychological ethic committee of the University Clinic Hamburg Eppendorf (LPEK-0118). Participants were randomly assigned to the ego depletion or the control condition and were led to believe that the influence of emotional videos on cognition and creativity is being tested. Small video clips with experimenter instructions were used to increase adherence during the online procedure. After answering a questionnaire assessing demographic variables, participants watched the surgery video. Next, all subjects completed the MSIT, a WCT for aggression, an IAT for aggression, the modified differential emotion scale (mDES) (Izard, 1977), and the BPAQ (Buss & Perry, 1992). At the end of the project, participants were thanked for their participation and debriefed through another video.

Measures

We used an emotion video, where emotions were to be suppressed, as depleting- and the MSIT as outcome task of the sequential-task paradigm. The emotion video showed sequences of different surgeries and the duration of the clip (9 minutes) was informed by other studies, successfully demonstrating ego depletion effects. While participants in the ego depletion condition received the instruction to watch the video and suppress any emotions, the control group was merely asked to watch the video attentively.

Multi-source interference task:

The MSIT was used to assess potential decrements in self-control capacity (Bush, Shin, Holmes, Rosen, & Vogt 2003). On each trial of the task a combination of three digits (1, 2, 3, or X) was presented at the center of the screen, with one digit differing from the other two identical digits. Participants were asked to identify the unique number by pressing the corresponding button key as fast as possible. One of two sources of interference was the position of the unique letter which was either congruent or incongruent with the number itself. The sequence 122 is an example of a congruent trial, since the unique number 1 is also spatially presented at the first and leftmost position. The sequence 212, on the other hand, is an example of an incongruent trial since the unique number (i.e., 1) and its position do not match (number 1 presented at the second rather than first position). The letter size constituted the second source of interference with the assumption being that participants' attention will be automatically drawn to items with big rather than small font sizes. Thus, RT's are predicted to be slower if the unique target is presented in a smaller font than the other items. The task consisted of 156 trials in total with 96 being incongruent, 42 congruent and 12 control trials (presenting X's as alternatives). Each trial started with the presentation of a fixation cross (500ms) followed by the stimulus. Trials ended with feedback (500ms) once participants gave a response or timed

out after 2000ms otherwise. The outcome measure was computed by subtracting the mean RT on congruent trials from the mean RT on incongruent trials, so that larger values express larger interference effects.

Word Completion Task for Aggression:

The previously validated German version of the WCT was one of two implicit aggression measures used in the current project. In this task participants had six minutes time complete as many word stems as possible. For example, the word stem KI_ _ could be completed to build the aggressive word KILL or the non-aggressive word KIND. In total the test contained 100 word stems that were presented on 25 separate pages, hence four word stems depicted on each page. These pages were presented randomly, so that all word stems were presented equally often throughout the project. The amount of total aggressive solutions divided by the number of total solutions was used as a measure of the accessibility of aggressive cognitions. In addition, the newly introduced more in-depth score, taking into consideration the word frequencies and the number of alternative solutions to each item, was assessed as an alternative indicator of implicit aggression. Due to the fact that ego depletion has been shown to increase aggression levels (Vohs et al., 2011), we used the WCT as a second outcome measure of ego depletion.

Implicit Association Task for Aggression:

The IAT for aggression is a computerized task that requires participants to categorize presented words referring to the concept of self and others and aggressive and peaceful words (Greenwald et al., 1998). More specifically, the task relies on the assumption that responses, which require the same response key, will be facilitated to the degree the two categories are associated within the semantic network of the test takers. Comparatively shorter RTs on blocks requiring the same response for aggressive and self-related words are taken as evidence for higher levels of implicit

aggression. IAT scores were obtained using an improved scoring algorithm (Greenwald et al., 2003). Just like the WCT, the IAT was used as a third outcome measure of ego depletion.

Modified differential emotions scale:

The mDES comprises 20 questions with half of them assessing positive and the other half negative emotions (Izard, 1977). Participants were asked to indicate the extent to which they feel the respective emotion by 5-point Likert scales ranging from 0 (not at all) to 4 (extremely). Average scores were built for all positive and all negative emotions. The mDES data served as a manipulation check. Because the control group was allowed to express natural emotions as an adaptive coping mechanism (Classen, Koopman, Angell, & Spiegel, 1996), participants in this group were expected to be in a more favorable emotional state after watching the video than participants in the emotion suppression group.

Buss-Perry Aggression Questionnaire:

The BPAQ is a widely used measure of trait aggression that compasses 27-items (Buss & Perry, 1992). The questionnaire assesses four different aspects of aggression, i.e., physical aggression (BP_PA), verbal aggression (BP_VA), anger (BP_A), and hostility (BP_H). Scores were used to control for the possibility that aspects of trait aggression explain potential between group differences in our implicit aggression measures.

Data analysis

A manipulation check was conducted using the mDES data. Independent sample t-tests were used to investigate whether participants with and without suppressed emotions differed in their emotional state after watching the video.

Likewise, independent samples t-tests were used to test whether participants in the emotion suppression (i.e., ego depletion) group performed worse on the MSIT and/or more aggressive on the WCT and IAT.

Lastly, Bayesian independent t-tests were used to answer the same question. Since this approach directly compares evidence in favor of the null and the alternative hypothesis more conclusive inferences can be drawn regarding the likelihood of any effect.

Results

Participants in the final sample had an average age of 27,03 years (*SD* = 4,76) with participants in the ego depletion group being slightly younger (*M* = 25,8) than participants in the control group (*M* = 28,3), *t*(58) = .89, *p* = .38. Gender and years of education did not differ significantly between the groups.

Compared to the control group, participants in the emotion suppression group did not express more negative emotions after watching the surgery video, *t*(58) = .89, *p* = .38. However, we observed a trend towards a significant between group difference in the amount of positive emotions experienced after watching the video, *t*(58) =1.99 , *p* = .052, with participants in the emotion suppression group expressing fewer positive emotions (*M* = 4.9, *SD* = 6.13) than participants in the control group (*M* = 7.57 , *SD* = 3.77).

Independent samples t-tests did not indicate any ego depletion effect, neither in the MSIT, *t*(58) = .37 , *p* = .97, the IAT, *t*(58) = .40 , *p* = .69 or the WCT, *t*(58) =1.49 , *p* = .14. Thus, participants in the emotion suppression group did not perform worse on a cognitive interference task or more aggressive on implicit aggression measures. Bayesian analyses, performed to evaluate whether the null- or the alternative hypothesis was more likely given the observed data, supported the null hypothesis in all cases. Bayes factors were 0.255, 0.201, and

0.117 for the MSIT, the WCT, and the IAT, respectively. Hence, the null hypothesis being 3.91, 4.96, and 8.55 times more likely for the observed MSIT, IAT, and WCT data, respectively.

Implications

Taken together, the findings of the current project raise concerns about using emotion suppression videos to induce ego depletion and add to the increasing number of projects failing to show ego depletion effects (e.g., Carter et al., 2015; Hagger et al., 2016). More specifically and contradicting basic ego depletion principles, participants who had to suppress their emotions while watching a surgery video did neither perform worse on a classic cognitive interference task (MSIT) nor more aggressive on implicit aggression measures such as the IAT or the WCT than participants who watched the same video without any instructions. Especially, the Bayesian analyses that were conducted in addition to classical frequentist statistics yielded evidence against ego depletion effects by showing that the null hypothesis was at least 4 times more probable given the data that we observed. This is moderate support for the null hypothesis and lets us argue against the existence of an ego depletion effect in the current project.

A common problem faced in ego depletion research is to discern whether the task used to induce or the task used to measure ego depletion was responsible for the failure to show ego depletion effects. The same is true for the present project. A final judgment on whether the emotion suppression video, our outcome measures, or a mixture of both was responsible for absent effects is intricate. For several reasons we assume that this failure was primarily due to the emotion suppression video being an ineffective inducer of ego depletion, however. First, our outcome measure (MSIT) has been shown to be a valid measure of ego depletion in a large-scale study previously (Dang, 2016). Secondly, the MSIT was not our only outcome measure so that chances to reveal effects in at least one outcome measure would have been high if suppressing emotions while watching the video had in fact depleted self-control resources.

Moreover, the fact that participants in the emotion suppression group felt less positive after watching the video does not only indicate that suppressing adaptive emotions led to unfavorable emotional outcomes, but also shows that participants in the emotion suppression condition adhered to the instructions. Thus, results are taken as evidence that the emotion video worked properly in all respects other than inducing ego depletion.

Showing ego depletion effects through emotion suppression would have been important in light of the ongoing debate about the existence of ego depletion, however. Preregistered large-scale multilab studies aimed at critically re-examining the effect have provided mixed results so far (Dang et al., 2021; Hagger et al., 2016). Yet, the study that successfully showed an ego depletion effect used two highly similar cognitive tasks to induce and measure ego depletion, i.e., the Stroop and the antisaccade task (Dang, 2021). Much of the initial appeal of the ego depletion effect originated from it *not* being domain specific and showing that seemingly unrelated behaviors such as resisting temptations, making important personal choices, and so forth all drew from the same limited self-control capacity. If two relatively similar tasks are used to deplete and measure decrements in self-control subsequently, this almost automatically leads to the question whether we are still talking about carry-over effects. And even worse whether reduced performance in fact represents depleted self-control resources rather than mere cognitive fatigue in a domain other than self-control. Therefore, we opted for an emotion suppression video, since showing ego depletion effects in seemingly unrelated tasks would be much more compelling evidence in favor of a central and limited self-control resource. Our choice for an emotion suppression video was based on a meta-analysis that identified this task as the most consistent inducer of ego depletion among all the different tasks used to deplete self-control resources (Dang, 2018). Even though the current project yielded no support for the notion that stifling responses during emotionally charged videos induces states of ego depletion, we strongly recommend future research to further address this possibility.

Unlike previous research (Barlett et al., 2016; Osgood & Muraven, 2015; Schmidt et al., 2015; Stucke & Baumeister, 2006; Vohs et al., 2011), we could not show a relationship between depleted self-control resources and aggression. While our focus on implicit rather than explicit aggression measures serves as one explanation for this disparity, another pertains to effects from publication bias. The fact that no other study with a failed attempt to demonstrate a relationship between depleted self-control resources and aggression could be found in the literature raises the suspicion that this strain of ego depletion research is subject to publication bias as well. Hence, caution is advised when interpreting findings from this branch of ego depletion research and future studies should clarify on the magnitude of bias and whether the relationship exists at all.

It must be acknowledged that a bigger sample size would have been desirable for more in-depth conclusions regarding the effectiveness of emotion suppression videos in depleting self-control resources. While a multilab large-scale study was beyond the scope of the present project, the effort of testing emotion suppression videos in such a project seems worthwhile in light of the promising meta-analytical data (Dang, 2018). Alternatively, many smaller and sound preregistered studies with comparably smaller sample sizes like ours could inform future meta-analyses and will allow more far-reaching conclusions than a single project alone.

In addition, the implementation in an online setting might have negatively impacted our results. We opted for such a design due to the worldwide COVID-19 pandemic and data quality checks generally contradicted this interpretation. Nevertheless, it must be acknowledged that online set-ups in general limit the amount of control that the experimenters have over factors such as situational influences.

To conclude, the present preregistered project further fueled critical accounts on the ego depletion effect. We found no evidence for ego depletion effects in the trio of a cognitive

interference task, and two implicit aggression measures when an emotion suppression video was used to deplete self-control resources. Moreover, we strongly recommend the usage of Bayesian statistics in future ego depletion research. In doing so, we could demonstrate that the absence of ego depletion effects was at least four times more likely for all our outcome measures. Nevertheless, we encourage more in-depth examinations of emotion suppression videos, either through a multilab large-scale study or many smaller preregistered projects like the present one. This will increase the power of the inferences and allow for a final judgment.

4. Discussion

In the following paragraphs of this chapter the most consistent findings of the current dissertation are summarized and embedded into the bigger picture of the recent scientific discourse. Moreover, resultant avenues for future research and the broader theoretical implications of the findings are outlined.

4.1 Main AAT Findings

Assembled evidence from all AAT projects conveyed a rather gloomy picture of the AAT's potential as a treatment add-on. Even though previous research most consistently demonstrated the beneficial effects of AAT-training in the realm of addiction (Loijen et al, 2020), no training effects were found among people with alcohol addiction or nicotine dependence. In other words, neither participants' alcohol nor tobacco consumption decreased significantly more than after sham-training. Moreover, no positive health effects were found in a more explorative project, in which AAT-training sought to establish healthier beverage preferences among healthy participants. All in all, AAT-training could not be shown to induce positive health effects that went beyond those of sham-training.

Self-evidently, causes of an effect cannot be identified if no effect was observed. Hence, absent training effects also hampered the search for underlying mechanisms. Nevertheless, trends below the threshold of statistical significance could still be informative since the data could be in line with a priori assumptions, but might fail to reach significance due to factors such as low power. For alcohol consumption this was not the case, however, since consumption decreased decisively across all groups with no special benefit of any training being visible. In other words, pronounced fluctuations and decline occurred in the control group, for which stability was expected. Therefore, it must be admitted that most variance likely stemmed from unknown factors, which lets interpreting differences between conventional- and inhibition

AAT-training seem inappropriate. Another valuable source of information concerning the role of inhibition in AAT-training was obtained from tracking changes in response inhibition over the course of smoking cessation. Although no statistically significant changes in Go/No-Go task performance were observed across the different trainings over time, visual inspection was somewhat informative. As one would expect, response inhibition capacity among participants in the inhibition AAT group improved. This is particularly interesting given that this group presented with the highest inhibition capacity at study start, so that room for improvement was smaller than in the other groups. If the assumption that AAT-training works via learning to inhibit prepotent responses to drug-elicited approach urges was true, also participants receiving conventional AAT-training should have improved in response inhibition capacity. However, in this group response inhibition barely changed, thus arguing against the involvement of response inhibition in AAT-training. In sum, non-significant findings in conjunction with the observed trends rather disconfirmed the involvement of response inhibition in AAT-training.

Furthermore, no link between AAT-training and alterations of (biased) action tendencies was found, since neither alcohol- nor tobacco approach bias decreased after training. Likewise, avoiding unhealthy and approaching healthy beverages during training did not alter action tendencies in a sample of healthy participants. Thus, although unchanged action tendencies can be interpreted in different ways, the most straightforward explanation seems to be that automatic action tendencies were *not* altered by AAT-training.

Interestingly, the percentage of people with nicotine dependence that could be identified based on approach bias scores resembled recent numbers found in alcohol research (Piercy et al., 2021). Reporting such numbers should become common practice since it will help in attaining a more realistic appraisal of the task's efficacy. For instance, when considering that only about 53-61 % of people with addiction presented with an approach bias, it becomes

obvious that approach bias can play *no* causal role in addiction as previously claimed by some authors (e.g., Gladwin et al., 2015). Thus, neither should we perceive biased action tendencies to cause addiction nor regard bias assessment as a diagnostic tool. Self-explanatory, sensitivity would need to be much higher to qualify AAT measures in this regard.

Empirical evidence of the current dissertation showed that (smoking) approach bias scores substantially varied both within the same measure and across different measures. This has major implications for the interpretation of approach and avoidance preferences in the realm of nicotine dependence. Most importantly, one should refrain from interpreting individual differences of bias scores since these are not reliable. However, valuable information can still be obtained from between-group differences because large enough sample sizes will enhance the probability that the extent of unknown variance and measurement error are evenly distributed across the groups.

Moreover, data from people with nicotine dependence indicated that mobile AATs are at least equally effective as PC-based measures when it comes to identifying people who smoke. This adds to recent proof-of-concept research in which basic approach and avoidance preferences for positive and negative stimuli were shown using mobile AATs (e.g., Zech et al., 2020). Hence, for now there is ample reason to believe that mobile AATs work properly not only for healthy participants but also in (sub) clinical populations. This is good news since it provides unprecedented flexibility in terms of research designs. For instance, recent evidence showed that reliability of the AAT can be enhanced by acquiring not one, but many consecutive measures of action tendencies (Zech et al., 2022) and it is obvious how their acquisition will be facilitated by mobile AATs. One interesting path for future research is to investigate whether multiple bias measures can increase the task's reliability in the domain of addiction as well. Moreover, building bias scores from many observations lowers the contribution of unknown

variance and thus holds a lot of potential for research seeking to identify factors related to biased action tendencies. Up until now, attempts to disclose consistent relationships mostly failed, probably because fluctuations led to spurious correlations in some cases and concealed actual relationships in others. Encouragingly, acquiring multiple bias measures can help in overcoming these difficulties and facilitate the identification of factors related to automatic action tendencies.

Yet another important avenue for future research is to compare different versions of mobile AATs and to test whether one is superior to the others. A closer look at all the different mobile AATs revealed that no two were alike. While PC-based set-ups are relatively fixed with the joystick positioned in between the PC and the test taker, set-ups with mobile devices almost logically provide more possibilities for heterogeneity. Most roughly mobile AATs can be divided into those that rely on touchscreen movements and those in which the device itself needs to be moved. Knowing whether one is superior to the other and in a next step which exact movements are best suited is not only important for selecting the most appropriate version in the future, but also could provide fruitful hints about the task's underlying cognitive mechanisms.

Lastly, the current dissertation shed some light on the neural substrates of AAT-training. The results can be considered as promising, yet mixed, given that alterations in cue-induced brain activity in a motor area of the brain confirmed a priori assumptions, but occurred without any concomitant behavioral change. While AAT brain-imaging research almost exclusively concentrated on people with alcohol addiction thus far, the approach to explore training-induced changes among healthy participants was a novelty. In contrast to people with addiction for whom reward-related structures such as the NAc and the mPFC (Ernst et al., 2014; C.E. Wiers et al, 2014; C.E. Wiers et al., 2015a) were associated with AAT-training, healthy participants

showed elevated cue-induced activity in the RCZ, a motor area of the brain, after AAT-training. Importantly, participants learned to avoid unhealthy- and to approach healthy drinks during training and changes of cue-induced activity were found for pictures of unhealthy beverages only. This pattern is in line with evidence from alcohol addiction (C.E. Wiers et al., 2015b) and suggests that learning to avoid detrimental alternatives plays a more important role for the effectivity of AAT-training than approaching healthy ones. Thus, overcoming seductive but detrimental approach urges seems to be more important to achieve behavioral change. In sum, showing changes in motor-related brain function that accompanied AAT-training is promising first evidence, but should be replicated in future research. Moreover, it would be important to find behavioral change that co-occurs with the change in brain function.

4.1.1 Broader Theoretical Implications

A pretty straightforward implication of the current findings is that we should critically rethink the usage of terms such as CBM and approach bias retraining. In alignment with a previous meta-analysis (Cristea et al., 2016), the attempt to disclose a mechanistic role of automatic action tendencies failed once more. Therefore, we question the common practice of interchangeably using the terms CBM, approach bias retraining, and AAT-training, because in doing so the misbelief of already knowing the mechanism via which AAT-training changes behavior becomes wrongfully fostered.

Also, views ascribing a causal role in addiction to biased action tendencies seem no longer tenable when considering the relatively low percentages of people with addiction who present with a drug-related approach bias. The relatively low sensitivity found in the present dissertation is in line with recent numbers found in alcohol addiction (Piercy et al., 2021) and argues against both the causal role of biased action tendencies in addiction and their suitability as a diagnostic tool.

An interesting question is whether the aforementioned relationships are in fact absent or only concealed by factors such as situational influences, unwanted variance, and measurement error. Giving a conclusive answer is currently impossible since research on this topic is still too scarce. For instance, very little is known about how situational influences might affect approach and avoidance behavior. However, it seems plausible that someone who drank too much alcohol the other day will show weaker approach biases the day after. Likewise, someone who just smoked a cigarette might feel less inclined to approach cigarettes right after. Another source of unwanted variance might originate from the selection of study stimuli. For example, people with alcohol addiction might have a predisposition for a couple of regularly consumed drinks only and possess aversions for others. Thus, if the favorite drinks are not included in the stimulus set or constitute only one among say 30 alternatives, the findings might be relatively imprecise reflections of alcohol approach biases. An interesting solution to deal with this source of unwanted variance is to tailor the stimuli to individual consumption in future research. In such an AAT, participants could select their individualized stimulus set based on the drinks they regularly consume. Generally, research into how situational factors and research design choices affect task performance seems worthwhile, since such insight could increase the task's reliability by improving research designs and minimizing unwanted sources of measurement error.

Somewhat intertwined, the general question arises what automatic action tendencies reflect? The most widespread view holds that automatic action tendencies reflect affective evaluations. This proposition is based on Solarz' seminal finding (1960) that positively valenced stimuli elicit automatic approach and negatively valenced pictures avoidance behavior. Thus, facilitation is supposed to occur when actions are compatible with preexisting affective evaluations and impeded if both are incompatible. However, it is important to note that what counts as positive and negative will likely vary from person to person and is shaped

by personality traits, personal experience, cultural beliefs, and so forth. Evidence for the involvement of personality traits comes from people high in reactive aggression who were shown to possess approach rather than avoidance preferences for negatively valenced attack-related stimuli (Lobbestael, Cousijn, Brugman, & Wiers, 2016). Thus, the tendency for reactive aggression seems to overwrite affectively driven avoidance tendencies that are observed among most individuals. Also, prosocial tendencies have been shown to reverse compatibility effects of stimulus valence. For instance, when presented with negatively valenced pictures of crying or feared people, prosocial participants show approach biases presumably because they want to help in relieving the adversity of the other (Gračanin, Krahmer, Rinck, & Vingerhoets , 2018). In addition, several mental disorders are characterized by reversed or absent automatic action tendencies. Highly socially anxious individuals were shown to exhibit automatic avoidance rather than approach behavior in response to positively valenced pictures of smiling people (Heuer et al., 2007) while people with depression often lack regular approach and avoidance preferences in response to both positive and negative facial expressions (Radke et al., 2014). Thus, personality characteristics and the accompanying motivations seem to affect automatic action tendencies. Systematic research as to whether reversals of action tendencies in fact go hand in hand with reversals of affective evaluations could provide further support for the crucial role of affective evaluations or demonstrate that factors such as personality traits influence action tendencies independently. Some evidence for such independent influences exists already and showed contextual effects on approach and avoidance tendencies (e.g., Bamford & Ward, 2008; Saraiva, Schüür, & Bestmann, 2013). Thus in a nutshell, affective evaluations seem to contribute decisively to automatic action tendencies, but explaining action tendencies exclusively in terms of such evaluations bears the risk of overlooking other determinants.

Importantly, the answer to this question has implications for construct validation of the task, as well. If automatic action tendencies are perceived to predominantly reflect affective

evaluations, the IAT seems to be most appropriate as another implicit measure of the same construct. If one, however, sticks to interpreting approach and avoidance time differences, the manikin task could be the more adequate choice. Hence, construct validation should include both to see which one is more adequate in terms of convergent validity. It is important to stress how important construct validation of the AAT as an implicit measure is. Such efforts can provide answers to the question whether approach biases indeed meet the criterion of reflecting automatic, unconscious aspects of cognition that differ qualitatively from explicit measures as is proposed in dual-process theory. Therefore, proper construct validation of the task is strongly recommended.

Lastly, careful perusal of the AAT literature disclosed surprisingly little evidence for the mechanistic role of automatic action tendencies (Eberl et al., 2013; Sharbanee et al., 2014). Nonetheless, the storyline that changes of action tendencies are responsible for the beneficial effects of AAT-training remained largely unaffected by contradictory evidence and is still ubiquitous (e.g., Baird et al., 2017; Eiler et al., 2020; Rinck et al., 2018). As we have seen, this (false) conviction is further amplified by using the terms CBM and approach bias retraining interchangeably with AAT-training, since both imply that changing biases lies at the core of the beneficial training effects. Thus, refraining from using such terms will help to raise the awareness that the mechanisms underlying the effectiveness of the task are largely unknown.

4.2 Main Ego Depletion Findings

An important contribution of the present dissertation is the development and construct validation of a German version of the WCT for aggression. The elegance of this implicit aggression measure is rooted in a low level of face validity and not requiring priming any aggressive cognition so that people are commonly completely unaware of what is being measured. This differs from most other (implicit) aggression measures in which the concept of

aggression is primed in one way or the other. In addition to the task itself, a new scoring formula was introduced and some first evidence for its superiority over the regularly used proportion of aggressive solutions was found. This formula takes both the number of alternative solutions and the word frequencies into account to derive implicit aggression indices. Future use of the new scoring formula and the WCT in general is recommended and should further prove their suitability. Moreover, implementing the new scoring formula in the English version to enhance its sensitivity as well seems promising. Yet, beforehand this version should undergo proper construct validation since this is still lacking.

After successful construct validation, the WCT was used in a sequential-task paradigm to investigate whether depleting self-control resources adversely affected subsequent cognitive performance and aggression levels. In line with critical accounts (e.g., Carter & McCullough, 2014; Carter et al., 2015; Hagger et al., 2016) no evidence for the existence of ego depletion was found since neither cognitive performance (MSIT) nor aggression levels (WCT and IAT) worsened after suppressing emotion during a surgery video. Logically, answering the question whether ego depletion is real or not is beyond the scope of a single project and many (large-scale) studies are required to resolve this issue. Nonetheless, the null findings add to a growing body of preregistered research that failed to demonstrate ego depletion effects (e.g., Hagger et al., 2016) and can feed into future meta-analyses on the topic.

The probably greatest intricacy in ego depletion research is that self-control is an elusive concept with no direct means to assess it. Therefore, indirect indices such as manipulation checks and subsequent task performance must be employed to infer whether self-control was engaged. Although the amount of positive emotion experienced after the video served as a manipulation check, it is obvious that this should not be equated with self-control. Strictly speaking, differences in the amount of positive emotion experienced after watching the video

merely express that the participants in the emotion suppression group adhered to the task instructions. Whether this adherence also consumed self-control is unknown. An interesting avenue for future research is to assess pupillometry while watching the video. If pupils of people who suppressed their emotions during the video dilate more, this would hint at mental effort- and self-control involvement.

It must be acknowledged that having had more than two tasks in the sequential-task paradigm added some level of extra complexity. Generally, in a sequence of "only" two tasks, a hypothesized effect in the outcome measure is inferred to reflect that both tasks relied on self-control exertion, while there are two options in case of nonsignificant findings; either both tasks did not tap into self-control or only one of them required its exertion. For instance, the first task might have successfully depleted self-control resources, but depletion went unnoticed due to an insensitive second task. Conversely, the first task might have failed to induce ego depletion although the second task could have assessed levels of depletion theoretically. If more than two tasks build the sequence things become even more intricate, since interjacent tasks might obscure the relationships. For instance, the MSIT (second task) might have been insensitive to assess states of ego depletion and even worse might have provided time to replenish depleted resources. Therefore, a state of ego depletion between the initial depletion task (emotion suppression video) and the later aggression measures might have been undetectable. Unfortunately, very little is known about the longevity of ego depletion effects and whether approximately five minutes spent on another task might suffice to recover self-control resources. Taken together, due to the theoretical possibility of replenishing self-control resources during interjacent task performance there is somewhat more certainty for the absence of ego depletion between the initial emotion suppression and the following MSIT than between emotion suppression and the two implicit aggression measures acquired later in the sequence.

4.2.1 Broader Theoretical Implications

Given that our findings and many recent large-scale studies further fueled the doubts about ego depletion's veracity, the implications for other related theories must be reviewed. For instance, when considering the interweavement of ego depletion and dual-process theory, it is hardly surprising that both theories are facing very similar issues. In an attempt to fight the conceptual crisis of ego depletion, Lurquin and Miyake (2017) called for a concise operational definition of self-control, independent validation of self-control tasks, and models that specify how self-control resources are depleted as well as the extent of exertion that is needed to impair subsequent self-control acts. Using an example, the following section will outline the importance of such theoretical adaptations for both theories and depict how their relevance will be threatened otherwise.

For illustrative purposes, let us consider an example: Imagine me sitting at my desk at the first day of summer, hearing the birds singing, and people enjoying themselves outside. I am not going to lie, the opportunity to join them and relax in the sun is tempting, but in order to obtain my long-term goal of attaining my PhD degree I have to refrain myself from doing so. Later that day, on the way back home I went to the supermarket to do some grocery shopping. Even though I had the resolution to have no dessert after dinner, as if by magic a box of walnut ice cream with maple syrup landed in my cart.

Adopting an ego depletion perspective, the sustained energy exerted to keep writing my dissertation can be conceived of as a first instance relying on self-control, while the situation at the supermarket, in which I acted contrary to my resolutions, can be thought of as another instance tapping into that same resource. Thus, the act of buying ice cream is taken as evidence that both situations drew from a shared self-control resource. This story line exemplifies both the appeal and problems associated with ego depletion theory. Most of the theory's appeal arose

from the assumption of a domain-general self-control resource, which allowed researchers to link many at first glance unrelated behaviors with the main assumption being that self-control will suffer if preceding acts relied on its exertion already (Baumeister et al., 2007).

Probably the most far-reaching problem in ego depletion research is that any inferences will be subject to alternative explanations without independent validation of self-control tasks. So rather than being depleted into temptation, I might have bought the ice cream simply because I was very hungry while being at the supermarket. In fact, it was close to midnight and a long time had elapsed after lunch. Also, research findings from studies with random assignment will be subject to alternative explanations if self-control tasks were not independently validated beforehand. For instance, the comparatively shorter time that people spent on unsolvable puzzles after eating radishes instead of tempting alternatives might again reflect differences in hunger states rather than ego depletion (Baumeister et al., 1998). Before study start all participants were food deprived and eating cookies in an ostensible taste test most certainly alleviated hunger more than eating radishes. Somewhat intertwined, sugar supply derived from eating cookies was higher and constitutes another alternative explanation for the observed difference in persistency. One way to disambiguate such findings is to provide independent validation through careful correlational research such as latent-variable analysis. Based on the domain-generality assumption one would expect to find an underlying commonality between ego depletion tasks, namely self-control resources (Lurquin & Miyake, 2017). Self-evidently, tasks must possess measurable properties to qualify for correlational analyses and unfortunately many self-control tasks do not. Along the lines of the previous examples, the duration that participants spent on solving the puzzle is quantitatively measurable, whereas resisting temptation (be it the urge to relax outside or to eat seductive food) lacks quantifiable properties. For the latter two it is therefore hard to tell how much self-control was exerted if at all.

In a similar vein, models of ego depletion need to be more concise when it comes to specifying how self-control resources are depleted and when resources are low enough to impair subsequent performance. Otherwise, clear and falsifiable predictions about whether or not to expect an ego depletion effect will remain difficult (Lurquin & Miyake, 2017). Ideally, the tasks used to induce a state of ego depletion should possess a parametrically manipulable characteristic that is thought to tap into self-control. The stop-signal task seems particularly promising in this respect, since it comprises a staircase algorithm to adjust stop-signal delays on the fly. Basically, these stop-signal delays set the task's difficulty level with successful abortion of already initiated motor responses becoming more difficult with longer delays. Most importantly, the staircase algorithm adjusts stop signal delays based on the participants' individual performance rather than having to select a somewhat arbitrary, one-size-fits-all level in advance, which increases the chances of it being sufficiently demanding for everyone. A lot of self-control tasks miss this important characteristic and how demanding a certain task actually was is therefore often a matter of debate. For instance, the chances are high that some participants in Baumeister and colleagues (1998) study were not tempted by chocolate cookies, thence fueling concerns about the main assumption of the paradigm. In sum, ideally tasks should possess a parametrically manipulable and measurable characteristic that provides the basis for falsifiable hypotheses. For tasks lacking measurable properties, assessing pupillometry could be fruitful in determining whether mental effort actually was engaged.

Dual-process models need very similar specifications and will remain hard to falsify otherwise. According to dual-process theory, I bought the ice cream because its sight activated positive, hedonic associative clusters that were formed through previous consumption. These prevailed over my symbolically represented restraint standards since situational boundary conditions (ego depletion) and internal states (homeostatic dysregulation) impaired the monitoring of ongoing behavior in accordance with symbolical representations (Hofmann et

al., 2008). Hence, dual-process models consider situational influences (such as hunger) and how these might shift the focus from one towards the other system. While this is distinct from ego depletion theory, both theories share a striking commonality. Although not explicitly referred to as resource models, dual-process models are based on the key assumption that reflective processes are effortful and require cognitive resources, whereas impulsive processes are thought to operate seamlessly without the need for much resources. Thus, whenever cognitive resources are sparse, impulsive processes are theorized to exert a greater influence on overt behavior. Hence, uncontrolled impulsive acts such as buying the ice cream are thought to unfold with greater ease if preceding behavior mainly implicated reflective processes (e.g., resisting temptation).

However, knowing how much load on reflective processes suffices to cause a switch from reflective towards impulsive processes would be extremely important to derive precise and falsifiable hypotheses. Unfortunately, dual-process models are currently too vague when it comes to specifying when behavior will be driven by the impulsive system. According to Strack and Deutsch's (2004) influential reflective-impulsive model behavioral schemata are activated through a final common pathway and schema conform behavior will be executed whenever a schema is activated above a certain threshold. However, where do these thresholds lie and how can we measure processes that feed into the pathway? Just like ego depletion theory, dual-process models need to state more precisely which amount of activation is needed to trigger schema conform behavior. Furthermore, processes feeding into the pathway would need to have measure- and thus testable properties in order to be able to derive clear-cut and falsifiable hypotheses. Otherwise, we might tap into confirmation biases and end up finding what we want to find. For instance, in the ice cream example we might have erroneously inferred that resisting temptation throughout the day heavily relied on reflective processes which in turn increased the chances of impulsively buying the ice cream. Taken together, operationalization that allow to

falsify propositions rather than ascribing reflective and impulsive attributes to behaviors in a post-hoc fashion are indispensable.

Implicit tests were proposed as suitable measures of impulsive processes in dual-process models (e.g., Hofmann et al., 2008). Almost self-evidently, tests in which we do not directly ask people but rather infer aspects of cognition from (automatic) responses are open to interpretation. Oftentimes implicit tests were introduced as a kind of panacea with grandiose prospects such as being a window into the unconscious (Banaji & Greenwald, 2013). Thus, they were supposed to capture qualitatively different aspects than introspectively accessible cognition. However, just because a test is implicit it does not automatically follow that it assesses qualitatively different aspects than self-report measures. Rather, this must be indicated in the process of construct validation, where convergent- and discriminant validity should be demonstrated for other implicit and explicit measures of the same construct, respectively. Unfortunately, in many cases construct validation was omitted and tasks were (probably incorrectly) inferred to assess implicit cognition on the basis of being implicit, or in other words implicit tests and implicit cognition were often conflated.

Admittedly, in light of the replication- and conceptual crisis maintaining trust and an accurate overview is challenging. It is impossible to separate “true findings” from those promoted by chance factors such as an uneven distribution of experimentally-relevant characteristics in studies with random assignment. Importantly, no meta-analytical (imputation) technique will provide a final answer, either, since one simply cannot know whether the findings reflect the effect’s real magnitude. Consequently, new projects that are based on new standards aiming to prevent the fallacies of the past are the only way forward to eventually yield trustworthy and compelling answers. Noteworthy, this does not only pertain to the field of ego depletion, but also encompasses dual-process models and implicit cognition tasks. In principle,

all the efforts that are currently undertaken to regain trust were the ones needed in the first place. Hence, an interesting idea to learn from history and to prevent its repetition is to install a sort of quality assurance through large-scale proof-of-concept studies once an effect presented sufficiently robust initial evidence. This way its veracity could be verified before building ramifications with other theories.

4.3 Final Thoughts

To conclude, where do we stand after the projects of the present dissertation? Well, there is a cynical voice in my head saying more or less where we started from. However, bearing in mind the disastrous consequences of publication bias in the realm of ego depletion research and the efforts that are currently necessary to resolve them, it becomes easier to adopt a more positive stance. Positively framed, there is now one potential AAT-training mechanism less to inquire, since no evidence for the involvement of response inhibition could be demonstrated. Because AAT-training did not systematically change automatic action tendencies either, the doubts about their mechanistic role were sowed once more, making the search for other mechanisms more important than ever.

While the failed attempts to induce behavioral change via AAT-training could be regarded as disproving the effectivity of AAT-training as a treatment add-on, other explanations seem more plausible when taking a wider view. For instance, compared to other AAT-research the sample sizes were relatively low and the comparatively low power might explain absent effects. It is important to note though, that the focus lay on explorative- rather than large-scale proof-of-concept purposes, letting comparisons appear kind of unfair. When considering proof-of-concept evidence, Rinck and colleagues (2018) study still serves as the gold standard. Here, AAT-training in a large sample of 1108 patients with alcohol addiction was shown effective in decreasing the relapse rate one year later by 8,4 %. Thus, the effect of AAT-training seems to

be robust, yet somewhat smaller than initially thought (Wiers et al., 2011), with our understanding of why that is still being poor. Therefore, projects aiming to contribute to a better understanding of the why are more important than ever.

Admittedly, maintaining trust and an accurate overview is challenging in face of the overarching threat of replication crisis and the conceptual problems mentioned above. For instance, in the domain of ego depletion it is impossible to separate "true findings" from those promoted by chance factors such as an uneven distribution of experimentally-relevant characteristics in randomized studies. Importantly, no meta-analytical (imputation) technique will provide a final answer, either, since one simply cannot know whether the findings reflect the effect's real magnitude. The steps necessary to either regain trust in the concept or to unambiguously disprove its existence were discussed in detail and mainly entail independent validation of self-control tasks in order to generate clear and falsifiable predictions about ego depletion states, which should then be tested in large-scale proof-of-concept research.

Rather than conceiving of the discussion surrounding ego depletion as a red rag, we should think of it as being fruitful for deducing precautions to avert similar patterns in the future. Essentially, all the efforts that are currently undertaken to regain trust were the ones needed in the first place and thus can be used to establish new standards to stop history from repeating itself. The most obvious precaution is to preregister planned research to prevent future publication bias. Fortunately, the need for preregistration has been widely accepted by now and thence is on a good way to becoming a mandatory scientific practice. Another interesting idea is to install large-scale proof-of-concept studies as a sort of quality assurance, once an effect demonstrated sufficiently robust initial evidence. This way its veracity could be verified before building ramifications with other theories. Importantly, the focus of such collective attempts is not on weakening well-established effects, but on testing how much trust, if any, is warranted

into a given theory. Self-evidently this will be somewhat disillusioning in some cases, but on a large scale it will help to identify empirically tested and trustworthy theories, upon which fruitful future research can be build. Of course, psychology is not comparable to mathematics, but in a certain sense gatekeeping via large-scale proof-of-concept projects can be viewed as the psychological equivalent of a mathematical proof. Moreover, such research will foster a sincere and open dialogue as well as collegial networks in general since multiple labs need to collaborate and strive towards the same goal(s). Noteworthy, the new standards do not only pertain to the field of ego depletion research, but in principle can help to increase the trust into psychological research as a whole.

5. Summary

Investigating which cognitive and neural mechanisms might be implicated in bringing about the beneficial effects of Approach and Avoidance Task (AAT)-training was the main aim of the present dissertation. A special focus lay on investigating the role of response inhibition in the task. However, the comparison of regular AAT-training, in which stimuli related to addiction must be avoided, and a newly developed inhibition AAT-training, in which responses needed to be inhibited, lent no support to response inhibition involvement. More precisely, neither regular nor inhibition AAT-training decreased consumption among people with alcohol- or nicotine addiction more than sham training. In addition, response inhibition capacity was not differently affected by any of the trainings, which argues against its involvement. The psychometric properties of the task (split-half and test-retest reliability) were low and raise concerns about interpreting individual differences of (smoking) AAT-scores. Encouragingly, PC and mobile versions of the task showed comparable precision when it comes to identifying people who smoke, opening the door to new, more flexible research designs. Among healthy participants brain-imaging research disclosed alterations in motor activity for unhealthy stimuli that were avoided during the training. Future research should replicate the role of motor-related activity and ideally show behavioral change which accompanies the alterations in brain activity.

To contribute to the ongoing debate about ego depletion's veracity it was investigated whether depleting self-control by stifling emotions during a surgery video worsened subsequent cognitive performance and/or aggression levels. However, in line with critical accounts no evidence for depleted self-control resources could be demonstrated. Noteworthy, before being used to measure self-control resources, a German version of the word completion task for aggression was successfully construct validated. Moreover, a newly introduced scoring formula showed initial evidence for its superiority over the conventionally used score.

Die Hauptzielsetzung der vorliegenden Dissertation war zu erforschen welche kognitiven und neuronalen Mechanismen zu den positiven Effekten des Approach und Avoidance Tasks (AAT)-Trainings beitragen. Ein spezieller Fokus lag hierbei auf der Untersuchung der Rolle von Reaktionsinhibierung. Der Vergleich von regulärem AAT-Training, in dem suchtrelatierte Stimuli vermieden wurden, und einem neu entwickelten Inhibitions-AAT-Training, in dem Reaktionen für eben diese inhibiert wurden, sprach jedoch gegen die Beteiligung von Reaktionsinhibierung. Genauer gesagt verringerte weder reguläres noch Inhibitions-AAT-Training den schädlichen Konsum von Menschen mit einer Alkohol- oder Nikotinabhängigkeit mehr als ein Täuschungstraining. Ebenso spricht die Tatsache, dass keins der Trainings die Fähigkeit zur Reaktionsinhibierung in besonderem Maße veränderte gegen deren Beteiligung. Die psychometrischen Eigenschaften des Tasks (splithalf und test-retest Reliabilität) waren niedrig und werfen Zweifel über die Interpretation individueller Unterschiede des (Rauch-) AATs auf. Erfreulicherweise zeigten PC und mobile Versionen des Tasks vergleichbare Präzision bei der Identifizierung von Menschen die rauchen, was Möglichkeiten für flexiblere Forschungsdesigns eröffnet. Unter gesunden Probanden enthüllten bildgebende Verfahren veränderte Motoraktivität gegenüber ungesunden Getränken, die während des Trainings vermieden wurden. Zukünftige Untersuchungen sollten die Rolle von Motoraktivität replizieren und idealerweise Veränderungen im Verhalten aufzeigen, welche damit einhergehen.

Um zu der Debatte über die Existenz von ego depletion beizutragen wurde untersucht, ob Selbstkontrollerschöpfung durch Emotionsunterdrückung kognitive Performanz und/oder Aggressivität verschlechterte. Mit kritischen Standpunkten übereinstimmend konnte jedoch keine Selbstkontrollerschöpfung demonstriert werden. Die in dem Projekt verwendete deutsche Version des Wortkomplettierungstasks für Aggressivität wurde vor der Nutzung als Selbstkontrollmaß separat konstruktvalidiert. Zusätzlich zeigte eine neuentwickelte Auswertungsmethode Überlegenheit gegenüber der konventionell benutzten Methode.

6. Abbreviations

AAT	Approach and Avoidance Task
BMI	Body Mass Index
BOLD	Blood-Oxygen-Level Dependent
BPAQ	Buss-Perry Aggression Questionnaire
BP_A	Buss-Perry Anger
BP_H	Buss-Perry Hostility
BP_PA	Buss-Perry Physical Aggression
BP_VA	Buss-Perry Verbal Aggression
CBM	Cognitive Bias Modification
CBT	Cognitive-Behavioural Therapy
CDS	Cigarette Dependence Scale
CIWA-A	Clinical Institute Withdrawal Assessment for Alcohol
dlPFC	Dorsolateral Prefrontal Cortex
EPI	Echo Planar Imaging
fMRI	Functional Magnetic Resonance Imaging
fNIRS	Functional Near-Infrared Spectroscopy
FOV	Field Of View
FTND	Fagerstrom Test for Nicotine Dependence
GAM	General Aggression Model
GLM	General Linear Model
IAT	Implicit Association Test
lMC	Left Motor Cortex
LTD	Lexical Decision Task
mDES	Modified Differential Emotion Scale
MINI	Mini Neuropsychiatric Interview
MNI	Montreal Neuroimaging Institute
mPFC	Medial Prefrontal Cortex
MPRAGE	Magnetization Prepared Gradient-Echo Sequence
MR	Magnetic Resonance
MRI	Magnetic Resonance Imaging
MSIT	Multi-Source Interference Task

NAc	Nucleus Accumbens
RCRQ	Revised Conflict Response Questionnaire
RCZ	Rostral Cingulate Zone
RM-ANOVA	Repeated Measures Analysis Of Variance
ROI	Region Of Interest
RPFT	Rosenzweig Picture Frustration Test
RT	Reaction Time
SCID	Structured Clinical Interview for DSM Disorders
SHS	State Hostility Scale
SPM	Statistical Parametric Mapping
TE	Echo Time
TLFB	Timeline Follow-Back
TR	Repetition Time
WCT	Word Completion Task

7. References

Anderson, C. A., & Bushman, B. J. (2002). Human aggression. *Annual review of psychology*, *53*(1), 27-51.

Anderson, C.A., Carnagey, N.L., & Eubanks, J. (2003). Exposure to violent media: The effects of songs with violent lyrics on aggressive thoughts and feelings. *Journal of Personality and Social Psychology*, *84*, 960-971.

Anderson, C. A., Carnagey, N. L., Flanagan, M., Benjamin, A. J., Eubanks, J., & Valentine, J.C. (2004). Violent video games: Specific effects of violent content on aggressive thoughts and behavior. *Advances in Experimental Social Psychology*, *36*, 199-249.

Anderson, C. A., Deuser, W. E., & DeNeve, K. M. (1995). Hot temperatures, hostile affect, hostile cognition, and arousal: Tests of a general model of affective aggression. *Personality and social psychology bulletin*, *21*(5), 434-448.

Anton, R. F., Moak, D. H., & Latham, P. K. (1996). The obsessive compulsive drinking scale: A new method of assessing outcome in alcoholism treatment studies. *Archives of general psychiatry*, *53*(3), 225-231.

Ascheid, S., Wessa, M., & Linke, J. O. (2019). Effects of valence and arousal on implicit approach/avoidance tendencies: A fMRI study. *Neuropsychologia*, *131*, 333-341.

Ascone, L., Wirtz, J., Mellentin, A. I., Kugler, D., Bremer, T., Schadow, F., ... & Kühn, S. (2023). Transferring the approach avoidance task into virtual reality: a study in patients with alcohol use disorder versus healthy controls. *Virtual Reality*, 1-12.

Baird, S. O., Rinck, M., Rosenfield, D., Davis, M. L., Fisher, J. R., Becker, E. S., ... & Smits,

J. A. (2017). Reducing approach bias to achieve smoking cessation: a pilot randomized placebo-controlled trial. *Cognitive Therapy and Research*, *41*, 662-670.

Bamford, S., & Ward, R. (2008). Predispositions to approach and avoid are contextually sensitive and goal dependent. *Emotion, 8*(2), 174-183.

Banaji, M. R., & Greenwald, A. G. (2013). *Blindspot: Hidden biases of good people*. Bantam.

Barlett, C., Oliphant, H., Gregory, W., & Jones, D. (2016). Ego-depletion and aggressive behavior. *Aggressive behavior*, *42*(6), 533-541.

Barlett, C. P., & Rodeheffer, C. (2009). Effects of realism on extended violent and nonviolent video game play on aggressive thoughts, feelings, and physiological arousal. *Aggressive Behavior: Official Journal of the International Society for Research on Aggression*, *35*(3), 213-224.

Baumeister, R. (2019). Self-control, ego depletion, and social psychology's replication crisis.

Baumeister, R. F., Bratslavsky, E., Muraven, M., & Tice, D. M. (1998). Ego depletion: Is the active self a limited resource? *Journal of Personality and Social Psychology*, *74*(5), 1252–1265.

Baumeister, R. F., Heatherton, T., & Tice, D. M. (1994). *Losing control: How and why people fail at self-regulation.* San Diego, CA: Academic Press, Inc.

Baumeister, R. F., Vohs, K. D., & Tice, D. M. (2007). The strength model of self-control. *Current directions in psychological science*, *16*(6), 351-355.

Beck, A. T., Steer, R. A., & Brown, G. (1996). Beck depression inventory–II. *Psychological assessment*.

Begg, S., Vos, T., Barker, B., Stevenson, C., Stanley, L., & Lopez, A. D. (2007). The burden of disease and injury in Australia 2003.

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.

Benjamin, A.J., & McCarthy, R. (2019, February 18). A Note of Caution in Justifying the Use of the Aggressive Word Completion Task. *Social Psychology in the Information Age.* https://ajbenjaminjrbeta.blogspot.com/2019/02/a-note-of-caution-in-justifying-use-of_18.html

Blázquez, D., Botella, J., & Suero, M. (2017). The debate on the ego-depletion effect: Evidence from meta-analysis with the p-uniform method. *Frontiers in Psychology*, *8*, 197.

Bohn, M. J., Krahn, D. D., & Staehler, B. A. (1995). Development and initial validation of a measure of drinking urges in abstinent alcoholics. *Alcoholism: clinical and experimental research*, *19*(3), 600-606.

Brass, M., & Haggard, P. (2008). The what, when, whether model of intentional action. *The Neuroscientist*, *14*(4), 319-325.

Brett, M., Anton, J. L., Valabregue, R., & Poline, J. B. (2002, June). Region of interest analysis using an SPM toolbox. In *8th international conference on functional mapping of the human brain* (Vol. 16, No. 2, p. 497).

Brockmeyer, T., Hahn, C., Reetz, C., Schmidt, U., & Friederich, H. C. (2015). Approach bias modification in food craving—A proof-of-concept study. *European Eating Disorders Review*, *23*(5), 352-360.

Brownstein, M., Madva, A., & Gawronski, B. (2019). What do implicit measures measure?. *Wiley Interdisciplinary Reviews: Cognitive Science*, *10*(5), e1501.

Brunel, F. F., Tietje, B. C., & Greenwald, A. G. (2004). Is the implicit association test a valid and valuable measure of implicit consumer social cognition?. *Journal of consumer Psychology*, *14*(4), 385-404.

Bush, G., Shin, L. M., Holmes, J., Rosen, B. R., & Vogt, B. A. (2003). The Multi-Source Interference Task: validation study with fMRI in individual subjects. *Molecular psychiatry*, *8*(1), 60-70.

Buss, A. H., & Perry, M. (1992). The aggression questionnaire. *Journal of personality and social psychology*, *63*(3), 452-459.

Carnagey, N. L., & Anderson, C.A. (2005). The effects of reward and punishment in violent video games on aggressive affect, cognition, and behavior. *Psychological Science, 16,* 882-889.

Carter, E. C., Kofler, L. M., Forster, D. E., & McCullough, M. E. (2015). A series of meta-analytic tests of the depletion effect: Self-control does not seem to rely on a limited resource. *Journal of Experimental Psychology: General*, *144*(4), 796-815.

Carter, E. C., & McCullough, M. E. (2014). Publication bias and the limited strength model of self-control: has the evidence for ego depletion been overestimated?. *Frontiers in psychology*, *5*, 823-834.

Carver, C. S., & White, T. L. (1994). Behavioral inhibition, behavioral activation, and affective responses to impending reward and punishment: The BIS/BAS scales. *Journal of Personality and Social Psychology*, 67, 319-333.

Classen, C., Koopman, C., Angell, K., & Spiegel, D. (1996). Coping styles associated with psychological adjustment to advanced breast cancer. *Health Psychology*, *15*(6), 434-437.

Cousijn, J., Goudriaan, A. E., & Wiers, R. W. (2011). Reaching out towards cannabis: Approach-bias in heavy cannabis users predicts changes in cannabis use. *Addiction*, *106*(9), 1667-1674.

Cristea, I. A., Kok, R. N., & Cuijpers, P. (2016). The effectiveness of cognitive bias modification interventions for substance addictions: a meta-analysis. *PloS one*, *11*(9), e0162226.

Curtin, J. J., McCarthy, D. E., Piper, M. E., & Baker, T. B. (2006). Implicit and explicit drug motivational processes: A model of boundary conditions. *Handbook of implicit cognition and addiction*, 233-250.

Dai, H., Milkman, K. L., Hofmann, D. A., & Staats, B. R. (2015). The impact of time at work and time off from work on rule compliance: the case of hand hygiene in health care. *Journal of Applied Psychology*, *100*(3), 846-862.

Dang, J. (2016). Commentary: A multilab preregistered replication of the ego-depletion effect. *Frontiers in Psychology*, *7*, 1155.

Dang, J. (2018). An updated meta-analysis of the ego depletion effect. *Psychological research*, *82*(4), 645-651.

Dang, J., Barker, P., Baumert, A., Bentvelzen, M., Berkman, E., Buchholz, N., ... & Zinkernagel, A. (2021). A multilab replication of the ego depletion effect. *Social Psychological and Personality Science*, *12*(1), 14-24.

Dang, J., Liu, Y., Liu, X., & Mao, L. (2017). The ego could be depleted, providing initial exertion is depleting. *Social Psychology*, *48*(4), 242-245.

De Houwer, J. (2006). What are implicit measures and why are we using them. *The handbook of implicit cognition and addiction*, 11-28.

De Houwer, J., & Moors, A. (2007). How to define and examine the implicitness of implicit measures. *Implicit measures of attitudes: Procedures and controversies*, 179-194.

Donker, T., van Straten, A., Marks, I. M., & Cuijpers, P. (2009). A brief Web-based screening questionnaire for common mental disorders: development and validation. *Journal of Medical Internet Research*, *11*(3), e19.

Eberl, C., Wiers, R. W., Pawelczack, S., Rinck, M., Becker, E. S., & Lindenmeyer, J. (2013). Approach bias modification in alcohol dependence: do clinical effects replicate and for whom does it work best?. *Developmental cognitive neuroscience*, *4*, 38-51.

Eberl, C., Wiers, R. W., Pawelczack, S., Rinck, M., Becker, E. S., & Lindenmeyer, J. (2014). Implementation of approach bias re-training in alcoholism—how many sessions are needed?. *Alcoholism: Clinical and experimental research*, *38*(2), 587-594.

Eiler, T. J., Haßler, B., Grünewald, A., Brück, R., Machulska, A., Klucken, T., ... & Gethmann, C. F. (2020). Swipe up to smoke less cigarettes! Introducing a mobile Approach-Avoidance Task Application to fight Smoking. *Current Directions in Biomedical Engineering*, *6*(3), 567-570.

Enkavi, A. Z., Eisenberg, I. W., Bissett, P. G., Mazza, G. L., MacKinnon, D. P., Marsch, L. A.,

& Poldrack, R. A. (2019). Large-scale analysis of test–retest reliabilities of self-regulation measures. *Proceedings of the National Academy of Sciences*, *116*(12), 5472-5477.

Ernst, L. H., Plichta, M. M., Dresler, T., Zesewitz, A. K., Tupak, S. V., Haeussinger, F. B., ... & Ehlis, A. C. (2014). Prefrontal correlates of approach preferences for alcohol stimuli in alcohol dependence. *Addiction biology*, *19*(3), 497-508.

Etter, J. F., Le Houezec, J., & Perneger, T. V. (2003). A self-administered questionnaire to measure dependence on cigarettes: the cigarette dependence scale. *Neuropsychopharmacology*, *28*(2), 359-370.

Everitt, B. J., Cardinal, R. N., Parkinson, J. A., & Robbins, T. W. (2003). Appetitive behavior: impact of amygdala-dependent mechanisms of emotional learning. *Annals of the new York Academy of Sciences*, *985*(1), 233-250.

Fazio, R. H., & Olson, M. A. (2003). Implicit measures in social cognition research: Their meaning and use. *Annual review of psychology*, *54*(1), 297-327.

Freud, S. (1922). Group psychology and the analysis of the ego. New York: Boni & Liveright.

First, M. B., & Gibbon, M. (2004). The Structured Clinical Interview for DSM-IV Axis I Disorders (SCID-I) and the Structured Clinical Interview for DSM-IV Axis II Disorders (SCID-II).

Fricke, K., & Vogel, S. (2020). How interindividual differences shape approach-avoidance behavior: Relating self-report and diagnostic measures of interindividual differences to behavioral measurements of approach and avoidance. *Neuroscience & Biobehavioral Reviews*, *111*, 30-56.

Friese, M., Hofmann, W., & Wiers, R. W. (2011). On taming horses and strengthening riders: Recent developments in research on interventions to improve self-control in health behaviors. *Self and Identity, 10*(3), 336-351.

Gawronski, B., Deutsch, R., & Banse, R. (2011). Response interference tasks as indirect measures of automatic associations. In K. C. Klauer, A. Voss, & C. Stahl (Eds.), *Cognitive methods in social psychology* (pp. 78–123). The Guilford Press.

Gladwin, T. E., Rinck, M., Eberl, C., Becker, E. S., Lindenmeyer, J., & Wiers, R. W. (2015). Mediation of cognitive bias modification for alcohol addiction via stimulus-specific alcohol avoidance association. *Alcoholism: Clinical and Experimental Research, 39*(1), 101-107.

Golden, C. J. (1978). *Stroop Color and Word Test: A Manual for Clinical and Experimental Uses*. Chicago, IL: Stoelting Co.

Gračanin, A., Krahmer, E., Rinck, M., & Vingerhoets, A. J. (2018). The effects of tears on approach–avoidance tendencies in observers. *Evolutionary Psychology, 16*(3), 1-10. https://doi.org/10.1177/1474704918791058.

Greenwald, A. G., McGhee, D. E., & Schwartz, J. L. (1998). Measuring individual differences in implicit cognition: the implicit association test. *Journal of personality and social psychology, 74*(6), 1464-1480.

Greenwald, A. G., Nosek, B. A., & Banaji, M. R. (2003). Understanding and using the implicit association test: I. An improved scoring algorithm. *Journal of personality and social psychology, 85*(2), 197-216.

Hagger, M. S., Chatzisarantis, N. L., Alberts, H., Anggono, C. O., Batailler, C., Birt, A. R., ...

& Zwienenberg, M. (2016). A multilab preregistered replication of the ego-depletion effect. *Perspectives on Psychological Science*, *11*(4), 546-573.

Heatherton, T. F., Kozlowski, L. T., Frecker, R. C., & Fagerstrom, K. O. (1991). The Fagerström test for nicotine dependence: a revision of the Fagerstrom Tolerance Questionnaire. *British journal of addiction*, *86*(9), 1119-1127.

Heuer, K., Rinck, M., & Becker, E. S. (2007). Avoidance of emotional facial expressions in social anxiety: The approach–avoidance task. *Behaviour research and therapy*, *45*(12), 2990-3001.

Hoefling, A., & Strack, F. (2010). Hunger induced changes in food choice. When beggars cannot be choosers even if they are allowed to choose. *Appetite*, *54*(3), 603-606. https://doi.org/10.1016/j.appet.2010.02.016

Hofmann, W., Friese, M., & Wiers, R. W. (2008). Impulsive versus reflective influences on health behavior: A theoretical framework and empirical review. *Health psychology review*, *2*(2), 111-137.

Houben, K., Havermans, R. C., Nederkoorn, C., & Jansen, A. (2012). Beer à No-Go: Learning to stop responding to alcohol cues reduces alcohol intake via reduced affective associations rather than increased response inhibition. *Addiction*, *107*(7), 1280-1287.

Houben, K., Nederkoorn, C., Wiers, R. W., & Jansen, A. (2011). Resisting temptation: decreasing alcohol-related affect and drinking behavior by training response inhibition. *Drug and alcohol dependence*, *116*(1-3), 132-136.

Hyman, S. E. (2005). Neurotransmitters. *Current biology*, *15*(5), R154-R158.

Hyman, S. E., Malenka, R. C., & Nestler, E. J. (2006). Neural mechanisms of addiction: the

role of reward-related learning and memory. *Ann. Rev. Neurosci.*, *29*, 565-598.

Inzlicht, M., & Friese, M., (2019). The past, present, and future of ego depletion. *Social Psychology*, *50*(5-6), 370-378.

Izard, C. E. (1977). Differential emotions theory. *Human emotions*, 43-66.

Jacobus, J., Taylor, C. T., Gray, K. M., Meredith, L. R., Porter, A. M., Li, I., ... & Squeglia, L. M. (2018). A multi-site proof-of-concept investigation of computerized approach-avoidance training in adolescent cannabis users. *Drug and alcohol dependence*, *187*, 195-204.

Jocham, G., Neumann, J., Klein, T. A., Danielmeier, C., & Ullsperger, M. (2009). Adaptive coding of action values in the human rostral cingulate zone. *Journal of Neuroscience*, *29*(23), 7489-7496.

Jones, A., & Field, M. (2013). The effects of cue-specific inhibition training on alcohol consumption in heavy social drinkers. *Experimental and clinical psychopharmacology*, *21*(1), 8-16.

Juergensen, J., & Leckfor, C. (2019). Stop pushing me away: Relative level of Facebook addiction is associated with implicit approach motivation for Facebook stimuli. *Psychological Reports*, *122*(6), 2012-2025.

Kakoschke, N., Hawker, C., Castine, B., de Courten, B., & Verdejo-Garcia, A. (2018). Smartphone-based cognitive bias modification training improves healthy food choice in obesity: A pilot study. *European Eating Disorders Review*, *26*(5), 526-532.

Kemps, E., & Tiggemann, M. (2015). Approach bias for food cues in obese individuals. *Psychology & Health*, *30*(3), 370-380.

Kemps, E., Tiggemann, M., Cibich, M., & Cabala, A. (2019). Cognitive bias modification for energy drink cues. *PLoS One*, *14*(12), e0226387.

Klein, A. M., Becker, E. S., & Rinck, M. (2011). Approach and avoidance tendencies in spider fearful children: The approach-avoidance task. *Journal of Child and Family Studies*, *20*(2), 224-231.

Koch, S., Holland, R. W., & van Knippenberg, A. (2008). Regulating cognitive control through approach-avoidance motor actions. *Cognition*, *109*(1), 133-142.

Kong, G., Larsen, H., Cavallo, D. A., Becker, D., Cousijn, J., Salemink, E., ... & Krishnan-Sarin, S. (2015). Re-training automatic action tendencies to approach cigarettes among adolescent smokers: a pilot study. *The American journal of drug and alcohol abuse*, *41*(5), 425-432.

Kühn, S., Kugler, D. T., Schmalen, K., Weichenberger, M., Witt, C., & Gallinat, J. (2019). Does playing violent video games cause aggression? A longitudinal intervention study. *Molecular psychiatry*, *24*(8), 1220-1234.

Kühn, S., Lorenz, R. C., Weichenberger, M., Becker, M., Haesner, M., O'Sullivan, J., ... & Gallinat, J. (2017). Taking control! Structural and behavioural plasticity in response to game-based inhibition training in older adults. *Neuroimage*, *156*, 199-206.

Laurens, M. C., Pieterse, M. E., Brusse-Keizer, M., Salemink, E., Allouch, S. B., Bohlmeijer,E. T., & Postel, M. G. (2020). Alcohol Avoidance Training as a Mobile App for Problem Drinkers: Longitudinal Feasibility Study. *JMIR mHealth and uHealth*, *8*(4), e16217.

Lobbestael, J., Cousijn, J., Brugman, S., & Wiers, R. W. (2016). Approach and avoidance

towards aggressive stimuli and its relation to reactive and proactive aggression. *Psychiatry Research*, *240*, 196-201.

Loijen, A., Vrijsen, J. N., Egger, J. I., Becker, E. S., & Rinck, M. (2020). Biased approach-avoidance tendencies in psychopathology: A systematic review of their assessment and modification. *Clinical psychology review*, *77*, 101825.

Luethi, M. S., Friese, M., Binder, J., Boesiger, P., Luechinger, R., & Rasch, B. (2016). Motivational incentives lead to a strong increase in lateral prefrontal activity after self-control exertion. *Social cognitive and affective neuroscience*, *11*(10), 1618-1626.

Lurquin, J. H., & Miyake, A. (2017). Challenges to ego-depletion research go beyond the replication crisis: A need for tackling the conceptual crisis. *Frontiers in Psychology*, *8*, 568. https://doi.org/10.3389/fpsyg.2017.00568

Machulska, A., Zlomuzica, A., Adolph, D., Rinck, M., & Margraf, J. (2015). "A cigarette a day keeps the goodies away": smokers show automatic approach tendencies for smoking—but not for food-related stimuli. *PloS one*, *10*(2), e0116464.

Machulska, A., Zlomuzica, A., Rinck, M., Assion, H. J., & Margraf, J. (2016). Approach bias modification in inpatient psychiatric smokers. *Journal of psychiatric research*, *76*, 44-51.

Manning, V., Garfield, J. B., Staiger, P. K., Lubman, D. I., Lum, J. A., Reynolds, J., ... & Verdejo-Garcia, A. (2021). Effect of cognitive bias modification on early relapse among adults undergoing inpatient alcohol withdrawal treatment: a randomized clinical trial. *JAMA psychiatry*, *78*(2), 133-140.

Maren, S. (1999). Long-term potentiation in the amygdala: a mechanism for emotional learning

and memory. *Trends in neurosciences*, *22*(12), 561-567.

Mehl, N., Mueller-Wieland, L., Mathar, D., & Horstmann, A. (2018). Retraining automatic action tendencies in obesity. *Physiology & Behavior*, *192*, 50-58.

Meuldijk, D., Giltay, E. J., Carlier, I. V., Van Vliet, I. M., van Hemert, A. M., & Zitman, F. G. (2017). A validation study of the web screening questionnaire (WSQ) compared with the Mini-International neuropsychiatric Interview-Plus (MINI-Plus). *JMIR mental health*, *4*(3), e5453.

Meule, A., Richard, A., Lender, A., Dinic, R., Brockmeyer, T., Rinck, M., & Blechert, J. (2020). Measuring approach–avoidance tendencies towards food with touchscreen-based arm movements. *Psychological research*, *84*, 1789-1800.

Meyer, D. E. & Schvaneveldt, RW (1971). Facilitation in recognizing pairs of words: Evidence of a dependence between retrieval operations. *Journal of Experimental Psychology*, *90*(2), 227-234.

Mischel, W., Shoda, Y., & Rodriguez, M. L. (1989). Delay of gratification in children. *Science*, *244*(4907), 933-938.

Morris, J. S., Öhman, A., & Dolan, R. J. (1998). Conscious and unconscious emotional learning in the human amygdala. *Nature*, *393*(6684), 467-470.

Mueller, V. A., Brass, M., Waszak, F., & Prinz, W. (2007). The role of the preSMA and the rostral cingulate zone in internally selected actions. *Neuroimage*, *37*(4), 1354-1361.

Muraven, M., Gagné, M., & Rosman, H. (2008). Helpful self-control: Autonomy support, vitality, and depletion. *Journal of experimental social psychology*, *44*(3), 573-585.

Nosek, B. A., Greenwald, A. G., & Banaji, M. R. (2007). The Implicit Association Test at Age 7: A Methodological and Conceptual Review. In J. A. Bargh (Ed.), *Social psychology and the unconscious: The automaticity of higher mental processes* (pp. 265–292). Psychology Press.

Ogawa, S., Lee, T. M., Kay, A. R., & Tank, D. W. (1990). Brain magnetic resonance imaging with contrast dependent on blood oxygenation. *proceedings of the National Academy of Sciences*, *87*(24), 9868-9872. https://doi:10.1002/mrm.1910140108

Osgood, J. M., & Muraven, M. (2015). Self-control depletion does not diminish attitudes about being prosocial but does diminish prosocial behaviors. *Basic and Applied Social Psychology*, *37*(1), 68-80.

Osgood, J. M., & Muraven, M. (2016). Does counting to ten increase or decrease aggression? The role of state self-control (ego-depletion) and consequences. *Journal of Applied Social Psychology*, *46*(2), 105-113.

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.

Persson, E., Barrafrem, K., Meunier, A., & Tinghög, G. (2019). The effect of decision fatigue on surgeons' clinical decision making. *Health economics*, *28*(10), 1194-1203.

Piercy, H., Manning, V., & Staiger, P. K. (2021). Pushing or Pulling Your "Poison": Clinical Correlates of Alcohol Approach and Avoidance Bias Among Inpatients Undergoing Alcohol Withdrawal Treatment. *Frontiers in psychology*, 1580-1586.

Quintero, G. C. (2013). Role of nucleus accumbens glutamatergic plasticity in drug addiction. *Neuropsychiatric disease and treatment*, *9*, 1499-1512.

Radke, S., Güths, F., André, J. A., Müller, B. W., & de Bruijn, E. R. (2014). In action or inaction? Social approach–avoidance tendencies in major depression. *Psychiatry research*, *219*(3), 513-517.

Richardson, D. S., & Green, L. R. (2003). Defining direct and indirect aggression: The Richardson Conflict Response Questionnaire. *Revue Internationale de Psychologie Sociale*, *16*(3), 11–30.

Rinck, M., & Becker, E. S. (2007). Approach and avoidance in fear of spiders. *Journal of behavior therapy and experimental psychiatry*, *38*(2), 105-120.

Rinck, M., Wiers, R. W., Becker, E. S., & Lindenmeyer, J. (2018). Relapse prevention in abstinent alcoholics by cognitive bias modification: Clinical effects of combining approach bias modification and attention bias modification. *Journal of Consulting and Clinical Psychology*, *86*(12), 1005-1016.

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

Robinson, E., Haynes, A., Hardman, C. A., Kemps, E., Higgs, S., & Jones, A. (2017). The bogus taste test: Validity as a measure of laboratory food intake. *Appetite*, *116*, 223-231.

Rosenzweig, S. (1945). The picture-association method and its application in a study of reactions to frustration. *Journal of personality*, *14*, 1-23.

Saraiva, A. C., Schüür, F., & Bestmann, S. (2013). Emotional valence and contextual affordances flexibly shape approach-avoidance movements. *Frontiers in psychology*, *4*, 933.

Schimmack, U. (2021). The Implicit Association Test: A method in search of a construct.

Perspectives on Psychological Science, *16*(2), 396-414.

Schmidt, A. F., Zimmermann, P. S., Banse, R., & Imhoff, R. (2015). Ego depletion moderates the influence of automatic and controlled precursors of reactive aggression. *Social Psychology*, *46*, (3), 132-141.

Schumacher, S. E., Kemps, E., & Tiggemann, M. (2016). Bias modification training can alter approach bias and chocolate consumption. *Appetite*, *96*, 219-224.

Sharbanee, J. M., Hu, L., Stritzke, W. G., Wiers, R. W., Rinck, M., & MacLeod, C. (2014). The effect of approach/avoidance training on alcohol consumption is mediated by change in alcohol action tendency. *PloS one*, *9*(1), e85855.

Sheehan, D. V., Lecrubier, Y., Sheehan, K. H., Amorim, P., Janavs, J., Weiller, E., ... & Dunbar, G. C. (1998). The Mini-International Neuropsychiatric Interview (MINI): the development and validation of a structured diagnostic psychiatric interview for DSM-IV and ICD-10. *Journal of clinical psychiatry*, *59*(20), 22-33.

Sievertsen, H. H., Gino, F., & Piovesan, M. (2016). Cognitive fatigue influences students' performance on standardized tests. *Proceedings of the National Academy of Sciences*, *113*(10), 2621-2624.

Skinner, H. A., & Allen, B. A. (1982). Alcohol dependence syndrome: measurement and validation. *Journal of abnormal psychology*, *91*(3), 199-209.

Smith, E. R., & DeCoster, J. (2000). Dual-process models in social and cognitive psychology: Conceptual integration and links to underlying memory systems. *Personality and social psychology review*, *4*(2), 108-131.

Sobell, L. C., & Sobell, M. B. (1992). Timeline follow-back. *Measuring alcohol consumption*

(pp. 41-72). Humana Press, Totowa, NJ.

Solarz, A. K. (1960). Latency of instrumental responses as a function of compatibility with the meaning of eliciting verbal signs. *Journal of experimental psychology, 59*(4), 239-245.

Spielberger, C. D. (1983). State-trait anxiety inventory for adults.

Stacy, A. W., & Wiers, R. W. (2006). An implicit cognition, associative memory framework for addiction.

Stacy, A. W., & Wiers, R. W. (2010). Implicit cognition and addiction: a tool for explaining paradoxical behavior. *Annual review of clinical psychology, 6*, 551-575.

Strack, F., & Deutsch, R. (2004). Reflective and impulsive determinants of social behavior. *Personality and social psychology review, 8*(3), 220-247.

Struijs, S. Y., Lamers, F., Vroling, M. S., Roelofs, K., Spinhoven, P., & Penninx, B. W. (2017). Approach and avoidance tendencies in depression and anxiety disorders. *Psychiatry research, 256*, 475-481.

Stucke, T. S., & Baumeister, R. F. (2006). Ego depletion and aggressive behavior: Is the inhibition of aggression a limited resource?. *European journal of social psychology, 36*(1), 1-13.

Sullivan, J. T., Sykora, K., Schneiderman, J., Naranjo, C. A., & Sellers, E. M. (1989). Assessment of alcohol withdrawal: the revised clinical institute withdrawal assessment for alcohol scale (CIWA-Ar). *British journal of addiction, 84*(11), 1353-1357.

Tangney, J.P., Baumeister, R.F., & Boone, A.L., (2004). High self-control predicts good

adjustment, less pathology, better grades, and interpersonal success. *Journal of Personality, 72*, 271-322.

Tiffany, S. T. (1990). A cognitive model of drug urges and drug-use behavior: role of automatic and nonautomatic processes. *Psychological review*, *97*(2), 147-168.

Ullsperger, M., & von Cramon, D. Y. (2004). Decision making, performance and outcome monitoring in frontal cortical areas. *Nature neuroscience*, *7*(11), 1173-1174.

US Department of Health and Human Services, (2006). The Health Consequences of Involuntary Exposure to Tobacco Smoke: A Report of the Surgeon General. *Coordinating Center for Health Promotion, National Center for Chronic Disease Prevention and Health Promotion, Office on Smoking and Health*.

Van Dessel, P., De Houwer, J., & Gast, A. (2016). Approach–avoidance training effects are moderated by awareness of stimulus–action contingencies. *Personality and Social Psychology Bulletin*, *42*(1), 81-93.

Van Dessel, P. , De Houwer, J. , Gast, A. & Tucker Smith, C. (2015). Instruction-Based Approach-Avoidance Effects. Experimental Psychology, 62 (3), 161-169.

Van Dessel, P., Eder, A. B., & Hughes, S. (2018). Mechanisms underlying effects of approach-avoidance training on stimulus evaluation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *44*(8), 1224-1241.

Veenstra, E. M., & de Jong, P. J. (2011). Reduced automatic motivational orientation towards food in restricting anorexia nervosa. *Journal of Abnormal Psychology, 120*(3), 708-718.

Vohs, K. D., Glass, B. D., Maddox, W. T., & Markman, A. B. (2011). Ego depletion is not just

fatigue: Evidence from a total sleep deprivation experiment. *Social psychological and personality Science*, *2*(2), 166-173.

Watson, D., & Clark, L. A. (1999). The PANAS-X: Manual for the positive and negative affect schedule-expanded form.

Wiers, C. E., Kühn, S., Javadi, A. H., Korucuoglu, O., Wiers, R. W., Walter, H., ... & Bermpohl, F. (2013). Automatic approach bias towards smoking cues is present in smokers but not in ex-smokers. *Psychopharmacology*, *229*(1), 187-197.

Wiers, C. E., Ludwig, V. U., Gladwin, T. E., Park, S. Q., Heinz, A., Wiers, R. W., ... & Bermpohl, F. (2015a). Effects of cognitive bias modification training on neural signatures of alcohol approach tendencies in male alcohol-dependent patients. *Addiction Biology*, *20*(5), 990-999.

Wiers, C. E., Stelzel, C., Gladwin, T. E., Park, S. Q., Pawelczack, S., Gawron, C. K., ... & Bermpohl, F. (2015b). Effects of cognitive bias modification training on neural alcohol cue reactivity in alcohol dependence. *American Journal of Psychiatry*, *172*(4), 335-343.

Wiers, C. E., Stelzel, C., Park, S. Q., Gawron, C. K., Ludwig, V. U., Gutwinski, S., ... & Bermpohl, F. (2014). Neural correlates of alcohol-approach bias in alcohol addiction: the spirit is willing but the flesh is weak for spirits. *Neuropsychopharmacology*, *39*(3), 688-697.

Wiers, R. W., Bartholow, B. D., van den Wildenberg, E., Thush, C., Engels, R. C., Sher, K. J.,

... & Stacy, A. W. (2007). Automatic and controlled processes and the development of addictive behaviors in adolescents: a review and a model. *Pharmacology Biochemistry and Behavior*, *86*(2), 263-283.

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.

Wiers, R. W., Rinck, M., Kordts, R., Houben, K., & Strack, F. (2010). Retraining automatic action-tendencies to approach alcohol in hazardous drinkers. *Addiction*, *105*(2), 279-287.

Wiers, C. E., Stelzel, C., Park, S. Q., Gawron, C. K., Ludwig, V. U., Gutwinski, S., ... & Bermpohl, F. (2014). Neural correlates of alcohol-approach bias in alcohol addiction: the spirit is willing but the flesh is weak for spirits. *Neuropsychopharmacology*, *39*(3), 688-697.

Willett, W. C., Sampson, L., Stampfer, M. J., Rosner, B., Bain, C., Witschi, J., ... & Speizer, F. E. (1985). Reproducibility and validity of a semiquantitative food frequency questionnaire. *American journal of epidemiology*, *122*(1), 51-65.

Wimmer, M. C., Dome, L., Hancock, P. J., & Wennekers, T. (2019). Is the letter cancellation task a suitable index of ego depletion?. *Social Psychology*, *50*(5-6), 345-354.

Woud, M. L., Becker, E. S., & Rinck, M. (2008). Implicit evaluation bias induced by approach and avoidance. *Cognition and Emotion*, *22*(6), 1187-1197.

Wrase, J., Grüsser, S. M., Klein, S., Diener, C., Hermann, D., Flor, H., ... & Heinz, A. (2002).

Development of alcohol-associated cues and cue-induced brain activation in alcoholics. *European Psychiatry*, *17*(5), 287-291.

Zech, H. G., Rotteveel, M., van Dijk, W. W., & van Dillen, L. F. (2020). A mobile approach-avoidance task. *Behavior Research Methods*, *52*(5), 2085-2097.

Zech, H. G., Gable, P., van Dijk, W. W., & van Dillen, L. F. (2022). Test-retest reliability of a smartphone-based approach-avoidance task: Effects of retest period, stimulus type, and demographics. *Behavior Research Methods*, *55*, 2652–2668.

8. Danksagung

Ich danke zu allererst meiner Erstgutachterin, Simone Kühn, die mir fachlich und menschlich immer zur Seite stand. Eine Eigenschaft die ich an Dir sehr zu schätzen weiß ist, dass Du Menschen die Möglichkeit gibst ihre eigenen Wege zu finden und zu gehen, anstatt sie vorzugeben. Danke, dass Du mich hast suchen lassen, was ich suchen musste und in schwierigen Situationen immer mit den richtigen Worten für mich da warst!

Mein ausdrücklicher und herzlicher Dank gilt ebenfalls Bastian Cheng und Olaf von dem Knesebeck für ihre kurzfristige Bereitschaft als Prüfungskomissionsmitglieder an dieser Dissertation mitzuwirken!

Ich möchte darüber hinaus Jürgen Gallinat und Tilo Strobach für Ihr Engagement, und das fachlich, organisatorische Mitwirken an dieser Arbeit danken!

Ein großer Dank gilt auch allen Probandinnen und Probanden für Ihre Bereitschaft die Wissenschaft zu unterstützen und diese Dissertation dadurch möglich zu machen!

Allen Helferinnen und Helfern, Kolleginnen und Kollegen, die mich auf unterschiedlichste Art und Weise in Zeiten der roten Null unterstützt und begleitet haben, möchte ich ebenfalls danken!

Und last but not least: Ein herzlicher Dank an meine Familie und Freunde! Ich bin dankbar, dass es Euch gibt und Ihr mich auf meinem Weg immer unterstützt!

9. Eidesstattliche Erklärung

Ich versichere ausdrücklich, dass ich die Arbeit selbstständig und ohne fremde Hilfe verfasst, andere als die von mir angegebenen Quellen und Hilfsmittel nicht benutzt und die aus den benutzten Werken wörtlich oder inhaltlich entnommenen Stellen einzeln nach Ausgabe (Auflage und Jahr des Erscheinens), Band und Seite des benutzten Werkes kenntlich gemacht habe.

Ferner versichere ich, dass die Dissertation bisher nicht einem Fachvertreter an einer anderen Hochschule zur Überprüfung vorgelegt oder mich anderweitig um Zulassung zur Promotion beworben habe.

Ich erkläre mich einverstanden, dass meine Dissertation vom Dekanat der Medizinischen Fakultät mit einer gängigen Software zur Erkennung von Plagiaten überprüft werden kann.

Unterschrift