ATTENTIONAL BIAS FOR ALCOHOL RELATED CUES: AN EXPLORATION OF ITS STATE-MOTIVATIONAL NATURE IN THE REAL WORLD

Thesis submitted in accordance with the requirements of the University of Liverpool for the degree of Doctor' in Philosophy by Panagiotis Spanakis

June 2019

Table of Contents

List of figures

List of tables

ATTENTIONAL BIAS FOR ALCOHOL-RELATED CUES: AN EXPLORATION OF ITS STATE-MOTIVATIONAL NATURE IN THE REAL WORLD.

Panagiotis Spanakis

Abstract

Incentive motivational models of addiction suggest that alcohol-related cues would automatically attract and hold attention (attentional bias - AB). Although AB is expected to be associated with alcohol-use behaviours, results in the literature have been equivocal. One explanation might be that AB is a state variable and, therefore, its magnitude and its association with alcohol-use behaviour will depend on the motivational context in which measures are taken (the state-hypothesis of AB). The aim of the thesis was to examine the state-motivational nature of AB in the real world.

For that purpose, a smartphone application (app) with the alcohol Stroop-task was developed. First, the thesis examined the psychometric properties of AB measures taken on the app in the real world (chapter three), and compared them with measures taken on different devices (e.g. computers) and on different environments (e.g. neutral laboratories and a Bar laboratory). The app was then used in a series of ecological momentary assessments (EMA) to examine the effect of the motivational context on magnitude of AB and on the association between AB and alcohol consumption. In the studies described in chapter four, AB and craving were measured in close temporal proximity to alcohol consumption, as well as in non-proximal assessments. In chapter five, AB and craving were assessed in moments of temptation to drink alcohol, in moments when alcohol was present and readily available for consumption (total availability of alcohol), and in fixed control moments. Finally, a focus group was conducted to explore attitudes of prospective users of the app (chapter six).

Measures of AB taken on the app in the real world showed good internal reliability but poor predictive validity. All other administration modalities showed poor psychometric properties. Results from the EMAs showed that AB did not increase in magnitude when measured proximally to alcohol consumption, in moments of temptation, or in moments of total availability. In one study, AB predicted alcohol consumption when measured proximally to alcohol use but this was not replicated in a second study. AB did not predict alcohol consumption when measured in moments of temptation or of total availability. Craving increased in magnitude when measured prior to initiation of drinking, in moments

of temptation, and in moments of total availability of alcohol. Craving predicted alcohol consumption when measured proximally to alcohol use and in moments of total availability, but not in moments of temptation. However, across all EMAs, measures of craving across the day were not associated with measures of AB. Findings from the focus group revealed that attitudes of prospective users of the app could be organized in three themes; users as participants, users as clients, and users as researchers. Re-development of the app in the future should focus on promoting autonomy of use, effectiveness of the app, and motivation for engagement.

Taken together, results did not support the state-hypothesis of AB and suggested that AB might have limited utility in predicting alcohol consumption. Craving increased in all situations when motivation to drink was expected to be high, and it robustly predicted alcohol consumption in most occasions. Future programs aiming to identify in real time moments of increased likelihood to consume alcohol should probably focus more on craving, rather than AB. The thesis also demonstrated the importance of measuring AB and craving in the real world, and that smartphone applications can be effectively used for that purpose. However, app developers should take attitudes of prospective users into account and balance the scientific validity of the app with its usability, user-friendliness and attractiveness.

Declaration

No portion of this work has been submitted in support of any other application for degree or qualification at this or any other University or institute of learning.

This thesis is dedicated to my dad and my grandad, for they have given me the first inspiration for studying, learning, and seeking knowledge.

Acknowledgements

First, I would like to thank my supervisory team for their help, guidance and support during my PhD. I would like to thank my primary supervisor, Dr. Paul Christiansen, for extending his role beyond commenting on drafts to really caring about his students. Paul - thank you for reminding me to press the pause button when needed. I would also like to thank my second supervisors, Prof. Matt Field and Dr. Andrew Jones. Matt - thank you for helping me to start this PhD and Andy - thank you for your valuable input throughout my studies.

My gratitude also goes to staff in the School of Psychology and the Department of Psychological Sciences for their help and support, to the Institute of Psychology, Health and Society for funding my PhD, and to all the people who expressed an interest for my research and took part in my studies.

Big shout-out to the group of companions, both here and back home, who generously provided their friendship, love and support. Danni, Eugenia, Flo, Esmira, Katerina, Vicente, and Danai - thank you for your friendship and for making us feel like home, when home was miles away. Graeme - thank you for being a valuable colleague, a kind friend, and an awesome officemate, since day one. Dimitri - thank you for always being a wise source of advice. Eri, Maki, Maria, and Konstantine - thank you for just keep doing everything you have been doing throughout the so many years we are best friends and which has been so important to me in many moments in my life.

I am grateful to all my family back home. A special thank you shall go to my mom. Her support - financially, morally and emotionally - gave me the opportunity to take the first steps in this big academic journey, since when I left home to start my undergraduate studies in Psychology. Since then, her support has never lapsed.

Most importantly, I am thankful to my wife Stefania, who stood by me on every moment, every hour, every day, and every step of the way, not only during the PhD but since the first day we found ourselves abroad at the start of this academic journey. I shall also thank one more person, whose support was so "unintended" but so soul-filling. I am thankful to our little daughter, Ariadne, who had a big (and initially toothless) smile to give me, every morning before going to the University, and every evening when coming back. Thank God that this academic journey is completed and that the work of these last years has come to fruition with this thesis.

Chapter One

1. General Introduction

1.1 General background of alcohol-use behaviours

1.1.1 Prevalence of alcohol use[1]

The most recent global report on alcohol and health, issued by the World Health Organization (World Health Organization - WHO,2014), reported that 52% of adults (defined as 15+ years of age) worldwide have drank alcohol in their lifetime, consuming on average 17.2ml of pure alcohol per day, which is the equivalent of approximately two UK units per day. Heavy episodic drinking (defined as more than 47.3 ml or approximately five UK units of alcohol consumed on a single occasion at least once per month) in adults shows a worldwide prevalence of 7.5% among the overall population and 16% among alcohol drinkers.

Within Europe, 79.4% of adults have consumed alcohol in their lifetime, consuming on average 30.3 ml of pure alcohol per day (approx. three units). The European region has the highest annual per capital alcohol consumption worldwide, with UK being in the second highest band (10-12.4 litres of pure alcohol per year) behind Portugal, several East-European countries, and Russia (over 12.5 lit/year). Prevalence rate of heavy episodic drinking is 16.5% among the overall European population and 22.9% among alcohol drinkers, with the UK being in the highest band (over 30% among alcohol drinkers).

Finally, an interesting measure is the 'pattern of drinking score' (PDS) which evaluates how risky the drinking pattern of a given population is and ranges from one (least risky pattern) to five (most risky pattern). High PDS indicates greater alcohol-attributable burden of disease in populations with comparable levels of consumption. Europe area shows a high diversity of PDSs, with UK scoring on the middle of the scale (PDS = three).

---

[1] In this section, alcohol consumption will be expressed as volume of alcohol (litres - lit. and millilitres -ml) or standard UK units (one UK unit = 8 g of alcohol).

1.1.2 Levels of severity in excessive drinking

Excessive drinking refers to problematic alcohol use that can have harmful effects on the user and on others in their environment (e.g. family and friends). Patterns of excessive drinking can range from hazardous drinking, to harmful drinking, and to alcohol dependence (Babor, Higgins-Biddle, Saunders, & Monteiro 2001). Hazardous drinking refers to alcohol use that increases the risk for harmful consequences on the user. Harmful drinking (also referred to as alcohol abuse) refers to alcohol use that is causing damage to the user's physical and mental health, while alcohol dependence refers to compulsive alcohol use which is hard to control, and which continues despite the harmful consequences (National Collaborating Centre for Mental Health - NCCMH, 2011). Results from the most recent epidemiological survey in the UK general population indicated that 16.6% of respondents were hazardous drinkers, 1.9% were harmful, and 1.2% were dependent to alcohol (McManus, Bebbington, Jenkins, & Brugha 2016).

In the Diagnostic and Statistical Manual of Mental Disorders, 5th edition (DSM-5; American Psychiatric Association - APA 2013), alcohol abuse (the diagnostic equivalent of harmful drinking described above; see NCCMH, 2011) and alcohol dependence are both included in a single diagnostic entity called alcohol use disorder (AUD). Diagnosis is provided on a severity continuum, based on 11 diagnostic criteria within 12 months (see table 1.1). Mild AUD is determined by the presence of two to three symptoms, moderate by the presence of four to five, while the presence of more than six criteria indicate severe AUD. The prevalence of DSM-5 AUD in the National Epidemiologic Survey on Alcohol and Related Conditions III (NASERC III) in the USA in adult general population was 13.9% and 29.1% for current and life-time diagnosis, respectively (Grant, Goldstein, Saha, & et al. 2015).

Diagnosis of AUD on a severity spectrum constitutes a change from the previous edition of the manual (the DSM-IV), were alcohol abuse and alcohol dependence were considered two distinct diagnostic entities. The distinction was eliminated on DSM-5 on the basis of lack of evidence that alcohol abuse is a distinct intermediate state between hazardous alcohol use and alcohol dependence (O'Brien 2011). However, the terms alcohol abuse and alcohol dependence will be retained in this thesis when referring to earlier studies that use the old classification terminology.

Table 1.1. Diagnostic criteria for alcohol-use disorder in DSM-5.

Diagnostic Criteria
1. Alcohol is often taken in larger amounts or over a longer period than was intended.
2. There is persistent desire or unsuccessful efforts to cut down or control alcohol use.
3. A great deal of time is spent in activities necessary to obtain alcohol, use of alcohol, or recover from its effects.
4. Craving or strong desire or urge to use alcohol.
5. Recurrent alcohol use resulting in failure to fulfil major role obligations at work, school or home.
6. Continued alcohol use despite having persistent or recurrent social or interpersonal problems caused or exacerbated by the effects of alcohol.
7. Important social, occupational, or recreational activities are given up or reduced because of alcohol use.
8. Recurrent alcohol use in situations in which it is physically hazardous.
9. Alcohol use is continued despite knowledge of having a persistent or recurrent physical or psychological problem that is likely to have been caused or exacerbated by alcohol.
10. Tolerance as defined by either of the following:
a. A need for markedly increased amounts of alcohol to achieve intoxication or desired effect.
b. A markedly diminished effect with continued use of the same amount of alcohol.
11. Withdrawal, as manifested by either of the following:
a. The characteristic withdrawal syndrome for alcohol.
b. Alcohol is taken to relive or avoid withdrawal symptoms.

(American Psychological Association, 2013, pp.490-491)

1.1.3 Risk factors for excessive drinking

According to previous literature reviews (Keyes, Hatzenbuehler, & Hasin 2011; Stone, Becker, Huber, & Catalano 2012) factors associated with increased risk for excessive drinking could be broadly organized in individual characteristics (e.g. being male, being white, low religiosity, positive alcohol expectancies), parental characteristics (e.g. maternal pre-natal alcohol consumption, mental-health disorders of the mother, and high parental education), family characteristics (e.g. family history of AUD, non-functional families with disrupted relationships), social-context characteristics (e.g. living in a disorganized neighbourhood and associating with high drinking peers), and life stressors (e.g. abused or maltreated in childhood or being discriminated as an adult due to belonging in a minority group).

Regarding the role of gender specifically, women seem to amass more protective factors (e.g. lower social acceptability of drinking and lower tolerance of alcohol) and less risk factors (e.g. less behavioural traits associated with increased alcohol consumption) compared to men (Nolen-Hoeksema 2004). It is worth noting that many of these factors might be more relevant to gender-role differences rather than biological differences.

Development of AUD has been associated with having a diagnosis of mental-health disorders (e.g. social anxiety, bipolar disorder, and intermittent explosive disorder; Buckner, Schmidt, Lang, Small, Schlauch, & Lewinsohn 2008; Swendsen, Conway, Degenhardt, Glantz, Jin, Merikangas et al. 2010), and early onset of regular drinking (broadly defined as earlier than 14 years old; Dawson, Goldstein, Patricia Chou, June Ruan, & Grant 2008; Hingson, Heeren, & Winter 2006; Hingson & Zha 2009). However, speed of transition from hazardous alcohol use to AUD slows down with earlier onset of use (Behrendt, Wittchen, Höfler, Lieb, & Beesdo 2009). A substantial role is also played by genetic factors which contribute up to 40-60% in AUD development (for a review see Schuckit 2009). For instance, genes responsible for metabolism-related enzymes may protect against AUD via increased alcohol sensitivity (greater susceptibility to the effects of alcohol). On the other hand, genes associated with certain behavioural traits (e.g. impulsivity, inhibition, and sensation-seeking; Schuckit 2009) and genes related to low responsiveness to alcohol (higher amounts of alcohol required to experience intoxication; Hu, Oroszi, Chun, Smith, Goldman, & Schuckit 2005) can contribute to the development of AUD.

1.1.4 Consequences of excessive drinking

Examining the substance-related harm in the UK, the Independent Scientific Committee on Drugs (ISCD)[2] rated the 20 most common psychoactive substances in the UK (e.g. alcohol, heroin, crack cocaine, methamphetamine, etc.) on a 0-100 scale, based on 16 criteria concerning harm to self and others (higher scores indicated greater harm; Nutt, King, & Phillips 2010). Alcohol had the highest score for harm to others (46/100), and the highest combined score for harm to others and to self (72/100). Regarding harm to self, alcohol had the fourth highest score (26/100), while methamphetamine (32/100), heroin (34/100) and crack cocaine (37/100) were the three highest.

1.1.4.1 Harm to self

Excessive alcohol use has been linked with a number of diseases, such as neuropsychiatric conditions (AUD being the most prominent, but also amnesia, alcoholic polyneuropathy, etc), gastrointestinal diseases, cancers, cardiovascular diseases, diabetes mellitus and infectious diseases (WHO,Rehm 2011; 2014). However, low average alcohol intake has been associated with beneficial effects on cardiovascular diseases and type-2 diabetes (Roerecke & Rehm 2012). The International Classification of Diseases, 10th edition (ICD-10), lists 39 conditions and sub-conditions which can be fully attributed to alcohol, as for example a) Mental-health and behavioural disorders attributed to use of alcohol (e.g. dependence syndrome, psychotic disorder, amnesic disorder), b) Alcoholic liver diseases and c) Poisoning (for the complete list see Rehm 2011). Apart from diseases, excessive drinking has also been associated with intentional and unintentional injuries. Life-time risk of fatal injury has been found to increase exponentially with increased average daily alcohol intake (Rehm, Zatonksi, Taylor, & Anderson 2011).

Excessive drinking also plays an important role in global rates of mortality and disability-adjusted life years (DALYs – years of life lost due to premature death or due to time living with a disease or disability). The recent global report on alcohol and health published by WHO (WHO,2014) estimates that in 2012 alcohol was responsible for 3.3 million deaths (5.9% of all deaths) worldwide, mainly due to cardiovascular diseases and diabetes (33.4% of all alcohol-related deaths), and for 5.1% of all DALYs, mainly due to neuropsychiatric disorders (24.6% of all alcohol-related DALYs). In 2010, alcohol-attributable injuries were responsible for almost ten deaths per 100,000 people globally, (13.2% of all deaths caused

[2] ISCD is a UK independent organization of drug experts.

by injury), and 422.6 DALYs per 100,000 (12.6% of all DALYs caused by injury; Rehm & Shield 2013).

Excessive drinking also contributes to the mortality and DALY rates of other serious conditions with high global burden of disease such as cancer and liver cirrhosis (Rehm & Shield 2013). In 2010, 4.2% of all cancer deaths globally and 4.6% of all cancer DALYs globally was caused by alcohol-attributable cancer. For the same year, 47.9% of all liver cirrhosis deaths globally and 47.1% of all liver cirrhosis DALYs globally was caused by alcohol-attributable liver cirrhosis.

1.1.4.2 Harm to others

Excessive drinking has a harmful effect not only on the individual drinker but also on others in the drinker's social and family environment (Laslett, Catalano, Chikritzhs, Dale, Doran, Ferris et al. 2010). The secondary effects of alcohol (Giesbrecht, Cukier, & Steeves 2010) can be exerted through physical and psychological violence against others, disruption of family life, driving under intoxication, as well as alcohol-induced foetal abnormalities and complications. Three studies conducted in Canadian adults from 2004-2008 (for a review see Giesbrecht et al. 2010) identified several violence-related alcohol secondary effects. A rate of 19.3%-22.1% of participants reported having been insulted or humiliated by a drunk person. More specific examples were being verbally abused (14.1%-15.8%) or being pushed or shoved (8.8% - 10.5%; for a complete list see Giesbrecht et al. 2010). A form of alcohol driven violence that seriously affects family life is violence against intimate partners. A meta-analytic review (Foran & O'Leary 2008) found a small to medium effect size ($r = .23$) for the association between male harmful alcohol consumption and aggression against the female partner, although a small effect size ($r=.14$) was found for the opposite association.

Another example of secondary effects of alcohol is motor vehicle accidents (MVA). A metanalysis of eight studies examining the association between acute alcohol consumption and MVA injuries showed that for every ten grams of pure ethanol consumed, odds ratios for an MVA injury increase by 24% (Taylor, Irving, Kanteres, Room, Borges, Cherpitel et al. 2010). Laslett et al. (2010), through analysis of public records in Australia, identified that in 2005, 277 people aged 15 and over have been killed in an MVA caused by alcohol consumed by another person (20% of all MVA-related deaths). Also, 3,643 people of the same age have been hospitalized for the same reason (9% of all MVA-related hospitalizations).

Regarding alcohol induced foetal abnormalities, a metanalysis (Popova, Lange, Probst, Gmel, & Rehm 2017) estimated that alcohol has been consumed during pregnancy by 25.2% of the female general population within Europe, and by 41.3% of the female general population within the UK (which was among the top five prevalence rates worldwide). They also found that the prevalence rate of foetal alcohol syndrome (FAS) in Europe was the highest worldwide (34.7 per 10,000 people) and that the UK belonged to the highest band of countries globally (more than 50 per 10,000 people). Based on countries where prevalence data were available for both drinking during pregnancy and FAS, the meta-analysis estimated that one out of 67 women who drink during pregnancy will give birth to a child with FAS, leading to an estimated 119,000 infants been born with FAS every year.

1.1.4.3 Economic costs

A review on the economic costs of excessive drinking on 13 high-income countries, including the UK (Mohapatra, Patra, Popova, Duhig, & Rehm 2010), showed that alcohol cost an average of $97.7 billion[3] to those countries; 1.58% of their gross domestic product (GDP). Most costs were associated with productivity loss due to premature mortality, morbidity, and disability (an average of $70.2 billion, 1.13% of their GDP).

In the UK, excessive drinking cost $32 billion (1.55% of GDP), with most costs associated with loss of productivity ($15.6 billion, 0.82% of GDP). A UK study focusing on health-related costs (Scarborough, Bhatnagar, Wickramasinghe, Allender, Foster, & Rayner 2011) found that alcohol-related health conditions cost the National Health Service (NHS) £3.2 billion, representing 4% of the total NHS budget. The alcohol-related conditions that cost the most were epilepsy (£900 million, 27.9% of all alcohol-related health costs) and other cancers (excluding mouth, oropharynx, oesophagus and liver cancer –£415 million, 12.9% of all alcohol-related health costs).

1.1.5 Summary and concluding notes

Alcohol is widely consumed globally, and a sizable portion of society is affected by excessive drinking problems. Excessive drinking has harmful consequences for the user, both regarding their health as well as their social environments. Although the development of alcohol dependence is a prominent health hazard of excessive drinking, there also several other serious health conditions related to excessive drinking (i.e. cardiovascular disease, diabetes, and cancer). Apart from personal health, other aspects of life, such as family

[3] For comparison reasons, all costs were converted from the local currency to US dollars, according to the 2006 exchange rate.

relationships and performance at work, are also adversely affected. Finally, problems derived from excessive drinking have an impact to society as whole through adverse consequences in public health and economy.

As a result, it is important to explore and understand the psychological processes that contribute to alcohol-use behaviours in order to inform the relevant behaviour-change interventions and public-health policies. Part of this exploration is the examination of the cognitive-motivational processes that drive excessive alcohol consumption.

1.2 The cognitive-motivational nature of addiction

1.2.1 Liking and wanting: Two related but distinct components of motivation.

Motivation broadly refers to a psychological orientation away or towards certain targets, depending on whether individuals find those targets aversive or attractive respectively. The concepts of 'liking' and 'wanting' are central in the conceptualization of motivation-driven behaviour. Liking refers to the hedonic experience which is associated with the relevant hedonic centres in the brain affected by opioid, endocannabinoid and GABA-benzodiazepine neurotransmitters. Wanting is a motivational process that makes things desirable, controlled mainly by mesolimbic dopamine systems (Berridge & Kringelbach 2008). Under most circumstances, liking and wanting act together, with pleasure (liking) experienced by what is desired (wanting) and vice versa. For example, an individual who enjoys cycling (likes to cycle) would also have a desire to cycle whenever they have the opportunity (wants to cycle). However, these two motivational aspects have disparate behavioural and neurobiological substrates and can also act independently from one another (Berridge, Robinson, & Aldridge 2009). In compulsive behaviours (e.g. substance administration in addiction), although there might be a strong urge to exert the behaviour (presence of wanting), the behaviour might not necessarily be enjoyable (absence of liking).

In the context of addiction, the distinction between 'wanting' and 'liking' might be evident in the continued use of the addictive substance, despite its harmful effects on the user's life (a process also referred to as the "paradox of addiction"; see Stacy & Wiers 2010). Substance use might initially be both pleasurable and desirable for the user. However, as the user develops dependence on the substance, the experience becomes less pleasurable due to the harmful consequences but remains wanted due to the addictive properties of the substance. Several psychological theories for addiction, which will be discussed next, explore the cognitive-motivational underpinnings of addiction and the role of the 'wanting' aspect of motivation-driven behaviour in seeking and consuming the substance.

1.2.2 Incentive-motivational models of addiction

1.2.2.1 Incentive sensitization model

Incentive salience refers to the ability of a target to evoke the 'wanting' aspect of motivation-driven behaviour. The incentive sensitization model (ISM) of addiction (Robinson & Berridge 1993, 2001) proposes that substance users become highly sensitive to the incentive salience of the substance, due to neuroadaptations in the mesolimbic dopamine system (particularly the nucleus accumbens), caused by repeated administration of the addictive substance.

The mesolimbic dopamine system can be activated by any natural reward (i.e. any pleasurable activity) and its response is normally greater to novel experiences, compared to more familiar (habituation of responsiveness; De Luca 2014;). For example, the first time a child tastes chocolate, they might find the experience highly rewarding. Although eating chocolate will probably remain pleasant in the future, the rewarding effect will not be as high as the first times. However, addictive substances can evoke the same rewarding effect as the first time, even after repeated administration (lack of habituation of responsiveness; Di Chiara 2002). This can lead to the mesolimbic dopamine system being 'hijacked' by addictive substances. As a result, incentive learning for addictive substances is more powerful compared to natural rewards, leading to an amplified incentive salience of the substance.

Importantly, this process does not only increase the incentive salience of the substance itself but also of any cue that has been repeatedly paired with it. Following the principles of Pavlovian conditioning, the substance represents the unconditioned stimulus (US), which evokes the unconditioned response (UR) of wanting. An initially neutral cue would be the conditioned stimulus (CS), which after repeated co-occurrence with the substance evokes the same response as the substance would do (the conditioned response - CS). For example, a bottle of alcohol (US) would evoke an alcohol-wanting behaviour (UR) in an alcohol-dependent individual. A bottle opener (CS), after being repeatedly paired with episodes of drinking, would also evoke an alcohol-wanting behaviour (CS).

After conditioning has been established, the substance-related cue will be able to trigger a desire to approach and acquire the cue in the expectation that the substance will be available as well (conditioned approach behaviour), or trigger a desire to seek and consume the associated substance (instrumental transfer), or even act as a reinforcer itself in the absence of the substance (conditioned reinforcement; for a review of the conditioning processes between cues and substances see Berridge et al. 2009; Robinson & Berridge 2008)

1.2.2.2 The role of automatic selective attention: Attentional bias.

In a related model, Franken (2003) proposed that between the detection of a substance-related cue and the conditioned activation of wanting, there is an intermediate step where selective attention is involved. Selective attention is a cognitive mechanism for effectively distributing cognitive resources, so that important stimuli are identified and processed, and non-important stimuli can be filtered out. Consistent with incentive sensitization theory, Franken argues that substance-related cues activate the dopaminergic reward systems, which leads to these stimuli being perceived as highly salient and, therefore, automatically grabbing and holding selective attention (a phenomenon called attentional bias - AB). Several studies, predominantly tobacco studies (for a review see Field & Cox 2008), have manipulated the conditioning of stimuli that were previously unrelated to the substance, to either explicitly pair (CS+) or un-pair them (CS-) with the substance. Participants reported increased subjective craving, and showed increased attentional bias, in response to the CS+ but not the CS-.

The model theorizes that this process of automatically allocating attention to substance-related cues is the implicit cognitive pathway through which the cues evoke the wanting behaviour described in the incentive sensitization model (i.e. craving). For example, AB would drive an alcohol misuser to fixate their attention on alcohol-related cues in the environment, which would then induce a state of pre-occupation and general preparation to receive the substance, which would lead to an urge to seek and consume the substance (i.e. craving). Furthermore, it is argued that AB and craving form a positive-feedback loop. Craving for the substance would make the person to further fixate on anything in the environment related to the substance, which would induce further craving.

This process of automatically granting priority to substance-related cues and making them the epicentre of attentional focus would contribute to the maintenance of addiction in two ways (Robbins & Ehrman 2004). First, the substance user is kept focused on substance-related cues, which can then trigger withdrawal symptoms and activate associated memories. Second, this process might consume cognitive resources, which would otherwise be used to exert some cognitive control and employ some substance-avoidance techniques.

1.2.2.3. Cognitive process model

Consistent with Robinson and Berridge (1993) and Franken (2003), Tiffany's (1990) cognitive process model suggests that seeking and consuming the substance is an automatic process triggered by the presence of substance-related cues. However, in contrast to the previous models, the cognitive process model suggests that evoking of wanting/craving by these cues is a non-automatic process, which is not necessary for seeking and consuming the substance.

Automatic processes are involuntary, fast and effortless, and they require very little cognitive capacity as they have been practised numerous times before. Although they are performed automatically, the individual is not completely unaware of them. Typical examples of such automatic sequences of actions would be driving a car, riding a bike, or reading a text. Although initial attempts to perform these actions probably required effortful engagement, those behaviours became automatic after repeated performance.

On the other hand, non-automatic processes are controlled by the individual and they require more time to be completed and more cognitive resources. Non-automatic processes tend to disrupt each other and be mutually disrupted by automatic processes (Tiffany & Conklin 2000). A typical example of this is the classic Stroop-task (Stroop 1935) where the automatic tendency to read the word interferes with the controlled and effortful attempt to ignore the meaning of the word and focus on the font colour.

Tiffany (1990; 2000) proposes that substance use is controlled by automatic cognitive processes that are triggered by substance-related cues. When those cues are encountered, the individual proceeds to seek, acquire, and use the substance in a ritualistic manner via a sequence of actions that are quick, out of the person's control, and very difficult to stop once initiated. Crucially, Tiffany stresses that no pre-conscious wanting/craving is necessary at this stage. Every time the addicted individual engages with seeking and using the substance, this is the outcome of an automated process, practiced numerous times in their life, which is not necessarily associated with a strong, irresistible urge to use the substance.

However, craving is not rendered irrelevant. Instead, the model argues that craving is a non-automatic process. Two situations when non-automatic processes are evoked, and which have relevance to addictive behaviours, are when automatic processes cannot be completed due to an obstacle or when the person tries to resist and inhibit an automatic process. Therefore, craving would be evoked when the automated process of seeking and obtaining the substance is obstructed (e.g. an addicted individual not being able to acquire their substance of addiction) or inhibited (e.g. an addicted person trying to abstain from using the substance).

Contrary to Robinson and Berridge (1993) and Franken (2003), craving in Tiffany's model is not considered an automatic, pre-conscious, conditioned response to trigger-cues. It is rather seen as an effortful, post-conscious behaviour to find the substance, which is activated when the habitual behaviour does not produce substance use, either due to efforts to reduce substance use or because of substance unavailability.

1.2.3 Theoretical Synthesis

All the three models described in section 1.2.2 argue for the role of some automatic processes that drive addicted individuals to repeatedly and compulsively seek out and use addictive substances. The element of automaticity is central, as they all describe a chain of events that happen rapidly, are not controlled by the individual, and are very difficult to stop or inhibit. Importantly, they all recognize the importance of cue reactivity and the central role played by substance-related cues. Furthermore, they all make a case for the addicted individual not being merely attracted by the substance itself but by substance-related cues that gained incentive salience through classical condition.

A typical example could be an alcohol-dependent individual who faces an alcohol advertisement in social media. Incentive sensitization model argues that the presence of the advertisement will automatically trigger the wanting of alcohol (i.e. craving) and that this is a learned response to alcohol advertisements.

Franken (2003) model would propose an intermediate step where the person's attention will first be automatically attracted by the advertisement and then it will be difficult to disengage. This immediate fixation and the subsequent enhanced cognitive processing of the advertisement will eventually evoke alcohol wanting.

Finally, the cognitive process model would suggest that there is no need for an alcohol-wanting state to be automatically evoked. The presence of the advertisement will set in motion an automated process where the individual may go to the fridge, grab a beer, open it and serve it, following a well-practiced ritual. Craving will only be experienced if they try to resist going to the fridge or should the latter be empty of beers.

Despite the apparent differences across models, the importance of reactivity to substance-related cues is a common factor. Although all theories describe slightly different ways of how substance use occurs, all suggest that substance-related cues would act as triggers of behaviour, activating either a conditioned response of response, a biased attentional system, or an automated sequence of actions.

1.3 Operationalisations of automatic responses to substance-related cues

As discussed in the previous section, an underlining principle in all incentive-motivation theories is that substance-related cues will trigger automatic cognitive responses. Encountering a substance-related cue activates these cognitive responses, which then biases the person's cognitive functioning towards a single target (to seek and consume the substance).

1.3.1 Implicit associations in memory

1.3.1.1 Measuring implicit associations

As discussed earlier, it is suggested that substance-related cues activate circuits of the brain implicated in reward, and therefore they acquire increased incentive salience. As a result, it is expected that substance-using individuals will hold implicit attitudes (i.e. not completely aware of these attitudes, but not completely unaware either) that the substance is something "positive" and "rewarding". A large body of literature uses the Implicit Associations Task (IAT) (Greenwald, McGhee, & Schwartz 1998) to quantify implicit associations between substance-related cues (e.g. a picture of an alcoholic drink) and emotional attributes (i.e. the word "good"), without participants necessarily having a conscious awareness of these associations.

In the IAT, participants are presented with a set of target stimuli (e.g. pictures of alcoholic drinks or soft-drinks) and emotional-attribute stimuli (e.g. the words "good" and "bad"), and they are required to press a unique response key (e.g. left arrow or right arrow) to categorize the stimulus in terms of the target-related concept (e.g. left-arrow for alcohol pictures / right-arrow for soft-drinks pictures) or in terms of the emotional attribute (e.g. left arrow for the word "good" / right-arrow for the word "bad"). Allocation of response keys to target-related concepts and emotional attributes alternates over the course of the task, until each target-related concept has shared the same response key with each emotional attribute (e.g. left arrow for alcohol pictures and "good"/right arrow for soft-drinks pictures and "bad", and vice versa.)

It is expected that participants will react faster to combinations of target-related concepts and attributes that are already associated in their minds, compared to when they are not. For example, quicker responses in an alcohol + positive / soft-drinks + negative condition than in an alcohol + negative / soft-drinks + positive condition would be indicative of a stronger implicit association of alcohol, rather than soft drinks, with positive, rather than negative, attributes.

A limitation of the classic IAT (also called bipolar IAT) is that the revealed associations are relative in nature. For example, the classic IAT might reveal stronger associations when alcohol is paired with a positive word and soft-drinks with a negative, rather than the opposite. However, it remains unclear whether this is a function of more positive attitudes for alcohol or more negative attitudes for soft-drinks. Two variations of the classic IAT that aim to overcome this limitation, and have been commonly used in addiction research, is the unipolar IAT and the single-category (or single-target) IAT (SC-IAT; for a review see Roefs, Huijding, Smulders, MacLeod, de Jong, Wiers et al. 2011). The unipolar-IAT examines whether the target-related concepts (alcohol and soft-drinks) would hold implicit associations with a single emotional attribute (either positive or negative). For example, a positive unipolar-IAT might reveal stronger associations when alcohol is paired with positive words and soft-drinks with neutral words, rather than the opposite. The SC-IAT examines whether a single target-related concept (e.g. alcohol) would hold stronger implicit associations with positive or with negative attributes. For instance, an alcohol SC-IAT might reveal stronger associations when alcohol is paired with positive words rather when it is paired with negative.

1.3.1.2 Alcohol implicit associations in non-clinical and clinical samples of alcohol using participants

Previous studies found that associations were stronger when alcohol-related cues were paired with negative attributes and soft drinks with positive, in both alcohol-dependent (De Houwer, Crombez, Koster, & De Beul 2004) and non-dependent samples (Houben, Nosek, & Wiers 2010; Houben & Wiers 2007b; Houben & Wiers 2008b; Wiers, Van Woerden, Smulders, & De Jong 2002). However, one study in non-dependent drinkers used personalised emotional attributes (i.e. participants categorised individually selected stimuli as liked or disliked) and found the opposite pattern (Houben & Wiers 2007a). It should be noted though that this finding might have to be discarded or at least interpreted with caution, as the authors report that the association was only "borderline" significant (p value was .056). In all studies (apart from De Houwer et al. 2004 who did not examine alcohol-use behaviour) positive alcohol associations were related to increased alcohol-use indices and negative alcohol-associations with decreased alcohol-use indices.

Other studies have used arousal attributes (arousal or sedation), rather than emotional attributes, and they found that associations were stronger when alcohol-related cues were paired with arousal attributes and soft-drinks with sedation (alcohol-dependent sample: De Houwer et al. 2004; non-dependent sample: Wiers et al. 2002). It was also found that this association was stronger in heavy, rather than light drinkers, and that the association of alcohol with arousal attributes was related to increased alcohol-use indices (Wiers et al. 2002).

As discusses earlier, the unipolar IAT examines the implicit associations of alcohol-related cues and soft-drinks with a single emotional attribute. Using the unipolar IAT, non-dependent drinkers have been found to hold stronger implicit association of alcohol with both emotional attributes (positive and negative), compared to soft drinks (Houben et al. 2010; Houben & Wiers 2006). However, another study in non-dependent drinkers found that only negative implicit associations were stronger for alcohol-related cues compared to soft drinks (Houben & Wiers 2008a). Regarding the extent to which these implicit associations were related to alcohol-use behaviour, some findings suggest that positive, but not negative, implicit association of alcohol were related to alcohol-use indices (Houben et al. 2010; Houben & Wiers 2008a), while some others demonstrated no relations at all (Houben & Wiers 2006).

Regarding unipolar associations with arousal- or sedation-related attributes, not-dependent drinkers showed stronger associations of alcohol with both attributes, compared to soft-drinks (arousal: Houben & Wiers 2006, 2008b; sedation: Houben & Wiers 2006). Both arousal and sedation-related implicit associations of alcohol were related to increased alcohol-use indices (although in Houben & Wiers 2008b this association was not robust across different environments and computer-software).

Finally, some studies (Houben & Wiers 2008a; Houben & Wiers 2009) used a combination of the unipolar IAT and the SC-IAT to examine associations of alcohol-related cues with a single emotional attribute (either positive or negative), without using soft-drinks as a reference category. Consistent with findings in studies using the unipolar IAT, they found that non-dependent drinkers held implicit association of alcohol with both emotional attributes, but only positive implicit associations were related to alcohol-use indices.

1.3.1.3 Summary and concluding notes

Although there is a level of variability, the previous findings suggest that alcohol holds implicit associations with arousal, while its association with emotional attributes is more ambivalent, being associated with both positive and negative attributes. The association of alcohol with arousal is thought to represent the psychomotor arousal involved in the 'wanting' aspect of motivated behaviour. Implicit associations of alcohol with arousal, combined with its ambivalent positive/negative implicit associations, might represent the synergy of 'wanting' and 'liking'. As discusses earlier, although the substance may initially be both wanted (arousal associations) and liked (positive associations), eventually it can be wanted (arousal associations) without being liked (negative associations).

However, it is worth noting that alternative explanations have been suggested for the negative implicit associations of alcohol (not all of them related to the cognitive-motivational basis of addiction). For example, the negative associations of alcohol found in the classic IAT might be a function of the contrasting category (i.e. soft-drinks), revealing stronger positive associations for soft-drinks, rather than negative associations with alcohol. Also, implicitly associating alcohol with a negative attribute (e.g. "bad"), might not necessarily reveal a personal negative attitude against alcohol. Instead, it might reflect use of alcohol when experiencing negative emotions (e.g. individuals drinking to cope with negative affect), or social norms that alcohol is damaging.

1.3.2 Approach bias

1.3.2.1 Measuring approach bias

Incentive-motivation theories suggest that encountering substance-related ques can motivate the substance-user to seek and consume the substance, either by increasing craving (Robinson & Berridge 1993 and Franken 2003) or by activating an automatic sequence of actions (Tiffany & Conklin 2000). As a result, it is expected that substance-using individuals will show a tendency to automatically approach substance-related cues. The Stimulus Response Compatibility Task (SRC) and the Approach Avoidance Task (AAT) are the most commonly used measures of approach bias (Watson, De Wit, Hommel, & Wiers 2012).

In the SRC, participants control a digital manikin on a computer, and they are instructed to categorize stimuli appearing on the screen (e.g. pictures) by moving the manikin towards (approach response) or away (avoidance response) from them. Categorization can be based either on the content of the stimuli (e.g. approach alcohol/avoid soft-drinks, and vice versa - relevant-SRC) or a content-irrelevant feature (e.g. approach landscape pictures/avoid portrait pictures, and vice versa - non-relevant SRC). Approach bias is inferred by comparing reaction times between the "approaching alcohol/avoiding control stimuli" block of the task and the "avoiding alcohol/approaching control" block. Faster reaction times when approaching alcohol and avoiding control stimuli are indicative of alcohol approach bias, while the opposite is indicative of alcohol avoidance bias.

A modified version of the SRC (the side-ways SRC; Barkby, Dickson, Roper, & Field 2012) includes an additional stimulus-response assignment, where the manikin can also perform a neutral movement, neither approaching nor avoiding the cue (i.e. moving sideways above the cue). In the classic SRC, alcohol-approach speed is compared to alcohol-avoidance speed, offering a relative measure of alcohol approach bias. The difference in the side-ways SRC is that it can provide an absolute measure of alcohol approach or alcohol avoidance bias by comparing the alcohol approach or alcohol avoidance speed with the speed of the neutral movement (the side-ways movement).

In the AAT, participants use a joystick lever to categorize stimuli appearing on a computer screen (e.g. pictures) by either pulling the lever towards them (i.e. approaching behaviour) or push it away (i.e. avoidance). Similar to the SRC, categorization can be based either on the content (relevant-AAT) or a content-irrelevant feature of the stimuli (non-relevant AAT). Approach bias is inferred in a similar way to the SRC task (comparing reaction times of approaching alcohol to reaction times of avoiding alcohol). To avoid the ambiguity of whether flexing the arm to pull the joystick is an approach behaviour (pulling something closer to the body) or an avoidance (retracting the hand away from the cue)[4], a zooming in/out element is incorporated where the stimulus is magnified when pulling and minimized when pushing.

[4] The same ambiguity also applies to extending the arm, which can be interpreted as avoidance (pushing something away from the body) or approach (reaching to grab the cue on the screen).

1.3.2.2 Alcohol approach bias in non-clinical and clinical samples of alcohol using participants.

1.3.2.2.1 Studies with non-clinical samples

A great body of literature has used the relative-SRC task to examine approach bias for alcohol-related cues and its association with alcohol-use behaviour. Heavy non-dependent drinkers have been found to show significant approach bias for alcohol related-cues (Schoenmakers, Wiers, & Field 2008). Studies comparing groups of non-dependent drinkers have found a significant alcohol approach bias in heavy but not light drinkers (Field, Caren, Fernie, & De Houwer 2011; Field, Kiernan, Eastwood, & Child 2008), and in high-craving but not low-craving drinkers (Field, Mogg, & Bradley 2005). However, Gladwin, Mohr, and Wiers (2014) found that alcohol drinkers showed significant approach bias towards appetitive stimuli in general, not specific to alcohol. Several studies have found that approach bias for alcohol-related cues was associated with increased alcohol-use indices (e.g. typical alcohol consumption and/or hazardous drinking patterns; Christiansen, Cole, & Field 2012; Field, Caren, et al. 2011; Field et al. 2008), although these findings have not been replicated in Field et al. (2005).

Other studies have used the non-relevant AAT to examine approach bias for alcohol-related cues. Wiers, Rinck, Dictus, and Wildenberg (2009) found that heavy drinkers with the OPRM1 allele (a gene that has been associated with rewarding effects of alcohol and craving) showed a significant approach bias for appetitive stimuli in general, not specific to alcohol. Participants without the allele showed no approach bias at all. Peeters, Wiers, Monshouwer, Schoot, Janssen, and Vollebergh (2012), compared drinking and non-drinking adolescents and found no difference in magnitude of alcohol approach bias, although approach bias was associated with individual differences in typical alcohol consumption.

One study compared the relevant SRC, the non-relevant SRC, the relevant AAT, and the non-relevant AAT, in terms of their ability to predict alcohol-use behaviour (Kersbergen, Woud, & Field 2015). They found that approach bias measured with the relevant SRC and the relevant AAT was associated with typical weekly consumption and hazardous drinking habits, although approach bias measured with the non-relevant tasks was not associated with alcohol-use behaviour.

1.3.2.2.2 Studies with clinical samples

Previous studies using the relevant-SRC task showed that alcohol-dependent patients hold a significant avoidance bias for alcohol related cues (Snelleman, Schoenmakers, & Mheen 2015) and that the magnitude of avoidance bias is greater in alcohol-dependent patients compared to healthy controls (Spruyt, De Houwer, Tibboel, Verschuere, Crombez, Verbanck et al. 2013). It should be noted though that Snelleman et al. (2015) examined patients in two sessions with a four-weeks interval and significant avoidance bias was found only in the first session, while no bias at all (neither approach nor avoidance) was found in the second. Regarding the extent to which magnitude of avoidance bias would be associated with relapse after a period of abstinence, the studies provided contradicting results. Although Spruyt et al. (2013) found that magnitude of avoidance bias was associated with chances of relapse, Snelleman et al. (2015) found that magnitude of avoidance bias did not differ between participants who relapsed and those who did not, and it was not associated with time to relapse (in either of two sessions).

However, presence of avoidance bias in alcohol-dependent patients has not been replicated in studies using the sideways SRC. Field, Di Lemma, Christiansen, and Dickson (2017) found no significant bias in alcohol-dependent patients (neither approach nor avoidance), and Barkby et al. (2012) found no difference in the magnitude of bias (either approach or avoidance) between alcohol-dependent patients and healthy controls. However, although Field et al. (2017) found no evidence of bias in alcohol-dependent patients, they found that individual differences in magnitude of avoidance bias were associated with percentage of heavy-drinking days after relapse (although the association was found in the four- and six-months follow-up but not in the two-months follow up).

Studies using the AAT to measure approach bias in alcohol-dependent patients and healthy controls revealed no significant approach bias for alcohol-related cues in either of the two groups (Ernst, Plichta, Dresler, Zesewitz, Tupak, Haeussinger et al. 2014 - using a relevant-AAT) and no difference in magnitude of approach bias between the two groups (Wiers, Gladwin, Ludwig, Gröpper, Stuke, Gawron et al. 2017 - using a non-relevant AAT). Wiers et al. (2017) also showed that there was no association between magnitude of approach bias and alcohol-use behaviour (life-time drinking history or duration of abstinence) in either of the two groups.

1.3.2.3 Summary and concluding notes

The findings presented above suggest that non-dependent heavy drinkers show a strong automatic tendency to approach alcohol. This is in line with predictions of the incentive-motivation models of addiction that alcohol-related cues will automatically activate a behaviour of acquiring and consuming the substance. However, some inconsistencies are also evident. Some studies showed that alcohol drinking participants hold a generic approach bias for rewarding stimuli or no significant approach bias at all.

Regarding alcohol-dependent patients, results suggest that these samples hold an automatic tendency to avoid, rather than approach alcohol-related cues. One explanation could be that alcohol-dependent patients who follow treatment would be strongly motivated to abstain from alcohol and therefore they might have developed a motivation to avoid alcohol-related cues. However, these findings seem to be task-specific to the SRC as they did not generalize to the side-ways SRC or the AAT. Finally, results are also inconsistent regarding the association of avoidance bias with alcohol-use behaviour.

1.3.3 Attentional bias

Incentive-motivation models of addiction suggest that alcohol-related cues, after being repeatedly paired with alcohol administration, may acquire the same incentive salience as alcohol itself through Pavlovian conditioning. Specifically, Franken (2003) proposes that as a result of their increased incentive salience, alcohol-related cues will automatically grab and hold attention (a phenomenon called attentional bias - AB). While approach bias discussed earlier refers to a tendency to automatically physically approach and obtain the alcohol-related stimulus, attentional bias refers to a tendency to automatically allocate and fixate attention to alcohol-related stimuli. It is suggested that automatic allocation of attention to alcohol-related stimuli increases the cognitive processing of the stimuli, which can induce a strong urge to seek and consume the substance. As attentional bias is the central focus of this thesis, it will be discussed separately.

1.4 Attentional bias for alcohol-related cues and its role in alcohol-use behaviour.

1.4.1 Measuring attentional bias

Attentional bias (AB) can be quantified trough two procedures; direct measurements of eye-gaze that would reveal where participants fixate their eye-gaze and for how long, or indirect measurements comparing reaction times, where AB is inferred through task performance. It is notable that the majority of AB tasks used in the literature are indirect measures, based on reaction times. The most common measures used in the addiction field (which will be discussed below) are the addiction Stroop-task, the visual probe task, the attentional cuing task, and the flicker induced change-blindness task. A greater focus will be devoted on the addiction Stroop-task as it is the task used throughout this thesis.

1.4.1.1 The addiction Stroop task.

The addiction Stroop-task is a specific variant of the emotional Stroop-task, based on the classic Stroop-task developed by Stroop (1935). In the classic Stroop-task, participants are asked to name the colour of the ink that various colour-names are written in. Some colour-names are presented in a matching ink-colour (e.g. the word yellow written in yellow ink - congruent presentation) while some others are presented in a contrasting ink-colour (e.g. the word yellow written in blue ink - incongruent presentation). Individuals require more time to read the incongruent, rather than congruent presentations; a finding known as the Stroop effect. Slower reaction time to the incongruent presentation is attributed to the cognitive interference caused by trying to supress the automatic response of reading the colour name in order to focus on the ink colour.

The emotional Stroop-task is an adaptation of the classic Stroop-task where participants must name the ink colour of various words, some of which are emotionally salient while others are neutral. In the addiction Stroop-task (Cox, Fadardi, & Pothos 2006), the emotionally salient words are related to the substance (e.g. alcohol-related words such as beer, spirit, etc.) while the neutral words belong to a common semantic category which should not have any particular emotional valence (e.g. environmental features: valley, cloud, etc).

Slower colour naming of the substance-related words compared to the neutral (the addiction Stroop effect), is indicative of greater AB for substance-related cues. Slower responses to the substance-related words is interpreted as attentional priority granted to the substance-related content of the word rather than its colour (a process which is absent in the neural stimuli), which interferes with completion of the task (naming the colour). Theoretical explanations of the addiction Stroop effect (for a review see Cox et al. 2006; Field & Cox 2008) propose that substance-related features of the stimulus (e.g. alcohol-related content) gain this attentional priority when competing with other features (e.g. the colour) due to their increased incentive salience (e.g. being more closely related to the substance user's motivation and current concerns). Also, fixation of attention to the alcohol-related features increases the cognitive processing of these features, which might then activate other cognitive processes that would also contribute to the cognitive interference caused. For example, processing the alcohol-related feature may elicit certain emotions and craving, as well as trigger an automatic retrieval of positive or negative substance-use expectancies from memory. These processes might increase the cognitive load caused by substance-related stimuli, also leading to slower reaction times.

A particular adaptation of the addiction Stroop-task that has been used in the literature (e.g. Bruce & Jones 2004) is the pictorial Stroop-task, where substance-related and neutral pictures are used instead of words. The pictures are presented after a colour-filter has been applied on them to recolour the entire picture (in some other studies pictures simply had a coloured outline; see Hester, Dixon, & Garavan 2006). As with the word-based addiction Stroop-task, the task is to name the colour of the filter ignoring the content of the picture. A study that used both a pictorial and word-based addiction Stroop-task to examine AB for cocaine-related cues in cocaine users (Hester et al. 2006) found significant AB (significantly slower reaction times in cocaine-related stimuli compared to neutral) for both words and pictures but they did not compare the magnitude of AB (size of the difference between reaction times for cocaine-related stimuli and neutral) produced by pictorial and word stimuli.

Although the addiction Stroop-task is widely used in the literature as a measure of AB, there have been some methodological considerations (Cox et al. 2006) which would be worth discussing. First, although colour-naming speed is expected to reduce when there is greater attention towards and cognitive processing of alcohol-related cues, it would be reasonable to assume that the same might happen if participants automatically turn their attention away from the alcohol-related words and try to supress any alcohol-related thoughts. The implication is that the addiction Stroop-task cannot differentiate whether participants have an attentional bias towards the substance (e.g. current heavy drinkers motivated to seek alcohol) or attentional bias away from the substance (e.g. excessive drinkers following treatment who are motivated to abstain). For example, Klein (2007) assigned alcohol-dependent patients in either of two conditions where they had to complete an alcohol thoughts suppression exercise or they were freely expressing alcohol-related thoughts, and their AB was then assessed with the alcohol Stoop-task. Patients showed significant AB for alcohol-related cues only in the thoughts-suppression condition, showing that avoidance of alcohol may induce an alcohol Stroop effect.

Second, another important methodological consideration regards the match of control to substance-related words. As described by Cox et al. (2006) control words should resemble the substance-related ones in various linguistic dimensions (e.g. number of syllables, frequency of use, etc.), they should comprise a single semantic category (e.g. all control words belonging to the group "environmental features") and there should be an equal number of substance-related and control words presented.

Finally, "carry over effects" may impact performance in the addiction Stroop-task (for a review see Waters, Sayette, Franken, & Schwartz 2005). Carry over effects describe the continued cognitive processing of the substance-related stimulus, even after it has disappeared from the screen. This prolonged processing can inflate reaction times in the following stimulus too. The carry over effect is important for the choice between blocked or mixed formats of the task. In blocked formats, all substance-related words will be presented together in one block and all control words in another. In the substance-related block, the continued cognitive processing of each word will inflate reaction-times on the next, on top of the increased reaction-times caused by the attention-grabbing properties of each word. The overall Stroop effect will, thus, be inflated as reaction-times differences between the two blocks will be increased. In mixed formats, substance-related words and controls are presented together in a random order. Due to carry over effects, neutral words that follow substance-related ones will "borrow" some of the interference caused by the substance-related words, as the individual might still be processing the substance-related word when the control appears. In this case, overall Stroop effects will be deflated, as reaction-time differences between the two types of words will be decreased. For a study aiming to identify the attention-grabbing properties of addiction-related words, a mixed presentation format might lead to false negative results. A proposed solution (Waters et al. 2005) has been the use of "filler trials". A third type of words, not related to either the substance or the control semantic category, are presented between mixed trials to absorb any prolonged interference effects and prevent them from carrying over to the next trial.

1.4.1.2 The Visual probe task and the Attention cueing task.

The Visual probe task (VPT), first developed by MacLeod, Mathews, and Tata (1986), is a reaction-time task measuring attention allocation towards specific stimuli. Participants are presented with a pair of pictures on a computer screen, placed side by side, one target-related (e.g. for an addiction VPT, a substance-related picture) one control (e.g. stationery items). After a delay, both pictures disappear from screen and a probe (e.g. a dot or an arrow) appears on the position previously occupied by one of the two pictures. Participants have to respond to the probe (e.g. pressing a key on the keyboard) as rapidly as possible. In the context of the addiction VPT, faster responses when the probe replaces the substance-related picture, compared to when it replaces the control picture, would suggest that participants allocated their attention more to the side of the screen where the substance-related cue appeared, thus indicating AB for these cues.

An important feature of the VPT, is its ability to differentiate between two components of attention, initial orientation of attention and maintained attention (Field & Cox 2008). Initial orientation refers to the automatic allocation of attention to the substance related-stimulus the moment that the stimulus appears, while maintained attention refers to the difficulty to disengage from the substance-related stimulus and shift attention towards another part of the screen. The two components of attentions can be differentiated in the VPT by using different durations of stimuli presentation. Short stimuli presentations (between 50 and 200ms) do not allow for further shifts of attention after initial presentation of the stimuli, thus measuring bias in initial orientation. Presentation durations longer than 200ms provide room for subsequent shifts of attention between the pictures, thus measuring bias in maintenance of attention.

The VPT can also be combined with eye-tracking technology (see for example, Christiansen, Mansfield, Duckworth, Field, & Jones 2015), providing the most direct measure of AB (compared to simply utilising reaction times). This method provides multiple outcome variables, as for example initial fixation to alcohol (i.e. time lapse between presentation of the alcohol-related stimulus and eye-gaze fixation on it), dwell-time (i.e. total time spend on looking at the alcohol-related stimulus) and disengagement ability (i.e. time lapse between fixating on the alcohol-related stimulus and withdrawing eye-gaze away).

The VPT has some key differences from the addiction Stroop-task (Robbins & Ehrman 2004). First, when reacting to a substance-related stimulus, AB will facilitate performance in the VPT (faster reactions to the probes replacing the substance-related stimulus compared to neutral), while it will obscure performance in the addiction Stroop-task (slower reactions when naming the colour of the substance-related stimulus compared to neutral). Second, while the addiction Stroop effect is attributed to greater cognitive interference caused by the alcohol-related stimulus compared to neutral, the effects observed in the VPT are related to greater visual orientation towards the substance-related stimulus compared to neutral. As a result, the VPT can also be combined with eye-tracking to provide a direct measure of attention (eye-gaze rather than reaction times). Finally, the simultaneous presentation of both stimuli in the VPT on either side of the screen, eliminates any carry over effects from one trial to the other.

A slight variation of the VPT is the Attention Cueing task (ACT; Stormark, Field, Hugdahl, & Horowitz 1997). In contrast to the VPT, where the two stimuli (substance-related and control) appear together on screen, participants on the ACT see only one stimulus, occupying the right or left half of the screen (the other half is left empty). After a delay, a probe appears either in the empty part of the screen (invalidly cued probes) or it replaces the stimulus (validly cued probes). AB is inferred by faster reaction times when cues are validly cued by alcohol-related stimuli compared to neutral and by slower reaction times when cues are invalidly cued by alcohol-related stimuli compared to neutral.

1.4.1.3 Flicker-induced change blindness.

Change blindness refers to the failure of participants to identify changes in a scene that are otherwise obvious. Change blindness has been linked to attentional focus, as it is suggested that detection of a change, no matter how obvious, is based on the level of attention allocated to the part of the scene where the change occurs (Simons & Rensink 2005). A previous study (Rensink, O'Regan, & Clark 1997) found that changes in a scene were detected faster when they concerned parts of the scene that were considered important, demonstrating that change blindness is inversely related with the importance of the place where the change occurs. They also found that when hints regarding the location of the change were provided, participants could easily detect it, showing that change blindness cannot be attributed simply to reduced visibility of the change.

In the context of AB for substance-related cues, participants would be expected to more easily detect changes in the part of a scene where substance-related stimuli appear, compared to other parts. This assumption can be tested with the flicker-induced change blindness task (flicker ICB, for an example see Jones, Bruce, Livingstone, & Reed 2006). In the flicker ICB, there are two versions of a picture, the original stimulus (OS) and the changed stimulus (Change-S). The OS depicts both substance-related and control objects (e.g. alcohol bottles and bottles of cleaning products). The Change-S is identical to the OS apart from a changed item which is either one of the substance-related or the control objects (e.g. a particular alcohol bottle replaced by another alcohol bottle). The OS and the Change-S are presented on a computer screen one after the other with a mask stimulus (MS) in between (e.g. a blank screen, an array of letters, etc.). The OS - MS - Change-S succession happens very fast so that the picture appears to flicker from one version to the other. Participants are then asked to make a response (e.g. press a specific key) the moment they have detected the change. Faster detection of the change when it concerns the substance-related, rather than the control object, would be indicative of AB for substance-related cues.

1.4.1.4 Psychometric properties of tasks measuring AB and related criticism

1.4.1.4.1 AB measuring tasks show poor psychometric properties

Although the measures described above, particularly the addiction Stroop-task and the VPT, are the most commonly used for evaluating AB for substance-related stimuli, the psychometric properties of these tests have not been promising. In an important paper, Ataya, Adams, Mullings, Cooper, Attwood, and Munafò (2012) reviewed the internal reliability of the addiction Stroop-task and VPT from a series of studies exploring AB for alcohol- and tobacco-related cues.

Regarding the VPT, they found poor internal consistency, particularly for initial allocation of attention. None of the VPTs had acceptable internal reliability (defined as Cronbach's $\alpha > 0.70$, see Kline 1999), while six out of the ten VPTs included in the study had very low Cronbach's α (between 0 and 0.09). This is consistent with Jones, Christiansen, and Field (2018) who reanalysed VPT studies from their laboratory and they also conducted two new studies. They found that in all cases internal reliability did not reach acceptable levels and test-retest reliability was also poor. Comparable results for the VPT have also been reported in anxiety research (Rodebaugh, Scullin, Langer, Dixon, Huppert, Bernstein et al. 2016; Schmukle 2005; Staugaard 2009).

With regard to the Stroop task, Ataya et al. (2012) found greater variability in internal consistency (Cronbach's α between 0 and 0.98). Notably, the lowest internal reliability was found in the unblocked Stroop-tasks. Those having a blocked format had higher Cronbach's α (from 0.53 to 0.98), with two cases crossing the threshold for acceptable internal reliability (these two cases had also the largest number of trials). A reasonable assumption behind this finding would be that carry over effects in the unblocked format (see section 1.4.1.1) inflate reaction times in some of the neutral stimuli trials (those preceded by a substance-related stimulus trial), creating the impression of inconsistent responding (Field & Christiansen 2012).

When directly comparing the two tasks, Ataya et al. (2012) found that the Stroop task had significantly greater internal reliability compared to the VPT, although they suggest these differences should be interpreted with caution due to unbalanced number of trials in each task. Indeed, they also found that number of trials affected each task differently, increasing internal reliability in the addiction Stroop-task while decreasing it in the VPT.

1.4.1.4.2 Suggested solutions

Suggestions that have been put forward to overcome these reliability problems include the use of personalized stimuli (Field & Christiansen 2012), pictorial stimuli (Bruce & Jones 2004), or eye-tracking technology (Christiansen, Mansfield, et al. 2015). It has also been suggested to use computation methods of AB scores that take into account temporal dynamics of AB (Zvielli, Bernstein, & Koster 2015).

To begin with, personalized stimuli refer to stimuli that reflect participants' personal preferences regarding the substance, as opposed to stimuli generally related to the substance. For example, in a study examining alcohol AB, generic stimuli would represent random types/brands of alcohol-drinks (e.g. beer, wine, or even generic words like alcohol or liquor) while personalized stimuli would focus only on the specific alcoholic drink that participants previously identified as their drink of choice (e.g. Guinness, rose wine). As Field and Christiansen (2012) explain, alcohol-cues that correspond to participants' own conditioning history will have greater incentive salience compared to more generic alcohol-related cues. Moreover, in generic AB tasks, participants may even be exposed to types of alcohol they dislike and which they do not consume at all. As a result, in tasks where not all stimuli are personalised, participants might show weaker AB (or not AB at all) for substance-related stimuli that they do not relate to, which would lead to reduced internal consistency of the task. Indeed, Christiansen, Mansfield, et al. (2015) have showed that personalized, but not generic versions of the alcohol VPT task, had acceptable internal consistency, although this was not replicated in Jones, Christiansen, et al. (2018).

Field and Christiansen (2012) also suggest that pseudo-personalization of stimuli might explain the slightly better reliability of tobacco VPTs, compared to alcohol (see Ataya et al. 2012), despite the very low overall VPT reliability. As most cigarettes have a very similar appearance when taken out of their boxes, this might remove any differences in personal preferences for certain brands, thus leading to more consistent responding.

For traditionally word-based tasks (e.g. the addiction Stroop-task), internal consistency might also be improved if pictorial stimuli are utilized instead of words. The meta-analysis by Ataya et al. (2012) discussed above, found that internal reliability was higher in addiction Stroop-tasks that used pictures instead of words, although this was not found in the VPT. However, these findings should be interpreted with caution as the meta-analysis included only one study with a pictorial addiction Stroop-task and only one with a word-based VPT.

One interpretation of this finding might be that pictorial stimuli are more salient compared to words (presumably because they look more realistic). Indeed, a study from the anxiety literature found that arachnophobic participants rated pictures of spiders as more aversive compared to spider-related words (Kindt & Brosschot 1997). As a result, pictures might attract and hold attention more consistently compared to words. However, it should be noted that in Kindt and Brosschot (1997) spider-phobic participants showed AB (measured with a spider Stroop-task) of comparable magnitudes for both types of stimuli. In the addiction literature, a study that compared cocaine-related pictures and words in the Stroop-task found that cocaine-users showed AB for both types of stimuli, although there was no comparison regarding the magnitude of AB produced by each type of stimuli.

Furthermore, eye-tracking based measures of AB have been shown to have better psychometric properties compared to reaction-time paradigms. Christiansen, Mansfield, et al. (2015) showed that eye-gaze indices of AB for alcohol-related stimuli in the VPT had higher internal reliability compared to reaction-times, either with generic or personalized stimuli (although reliability was acceptable only with personalized stimuli). Similar patterns have been found in the VPT with cocaine-related stimuli, where eye-gaze measures had superior test-retest reliability compared to reaction times (Marks, Pike, Stoops, & Rush 2014), and participants showed significant AB when AB was inferred from eye-gaze indices rather than reaction times (Marks, Pike, et al. 2014; Marks, Roberts, Stoops, Pike, Fillmore, & Rush 2014). However, Jones, Christiansen, et al. (2018) found that both internal and test-retest reliability were poor in the alcohol VPT regardless of whether reaction-time or eye-gaze measures were used.

Finally, Zvielli et al. (2015) argue that the poor reliability (especially for the VPT) can be attributed to the way that the bias scores are calculated and the accompanying assumptions for the nature of AB. Traditional calculation of AB scores (average reaction time for neutral cues subtracted from average reaction time for substance-related cues) assumes that attentional bias is a "static signal", where participants' attentional bias remains stable across all trials and their attention is selectively focused on substance-related cues in a consistent way. Zvielli et al argue that AB is expressed dynamically in time, with episodic increases and decreases, and therefore they introduced the trial-level bias scores (TL-BS). TL-BS is based on the calculation of individual AB scores in the VPT for pairs of adjacent congruent (probe replaces substance picture) and incongruent trials (probe replaces control picture). These scores can be then plotted in a timeline and reveal fluctuations from early to late trials in the task. They can also be aggregated in average AB towards or away substance-related stimuli, average peak AB towards or away substance-related stimuli, and average variability among AB scores. The authors compared split-half reliability of a smoking VPT for the two computation methods and found that TL-BS scores had better reliability (Pearson's r between .31 and .67) than traditional AB calculation (r=.06). However, Jones, Christiansen, et al. (2018) found that the alcohol VPT had poor test-rest reliability when either the classic or TL-BS computation method was used.

1.4.1.4.3 Summary and concluding notes

The VPT and the addiction Stroop-task have been criticised for having poor psychometric properties, although the addiction Stroop-task might perform a bit better than the VPT. An important step forward would be for future studies to explore the suggested solutions to increase the psychometric properties of these tasks (although a recent attempt to improve the reliability of the VPT has not been successful; Jones, Christiansen, et al. 2018). Using tasks with good psychometric properties is imperative to accurately understand the phenomena under examination.

As discussed by Rodebaugh et al. (2016), unreliable measures hinder our ability to differentiate individuals based on their AB scores or examine mediation models that would reveal the role of AB as a mechanism behind various addiction-related phenomena. Lack of reliability is also important when assessing the association between AB and substance use. These limitations have implications for both our understanding of the role of AB in addiction and the translation of AB research into clinical practice. Rodebaugh et al. (2016) also note that the reliability crisis of AB measures might imply that the problem stems from flaws in the theory supporting the measures. One should entertain the idea that the poor reliability of the Stroop and the VPT might not be the outcome of only poor methodological choices (e.g. blocked vs unblocked format, personalized vs generic stimuli, etc.) but also of erroneous theoretical conceptualizations of the construct that is measured.

1.4.2 Evidence for AB for alcohol-related cues in non-clinical and clinical samples of alcohol-using participants

1.4.2.1 Studies with non-clinical samples.

Several studies have examined AB for alcohol related cues in non-dependent drinkers using the VPT. For example Harrison and McCann (2014) found that moderate social drinkers showed significant AB for alcohol related-cues. Miller and Fillmore (2010) and Christiansen, Mansfield, et al. (2015) combined the VPT with concurrent eye-tracking to obtain both reaction-time based and eye-gaze based indices of AB. Miller and Fillmore (2010) also used simple pictures (e.g. a single alcohol item on plain background) and complex scenes (e.g. real-life scenes of people consuming alcoholic drinks) and found that participants showed significant indices of AB (both reaction-time and eye-gaze) in simple pictures but not complex scenes. Christiansen, Mansfield, et al. (2015) used stimuli generally related to alcohol and personalised stimuli and found that participants showed significant AB for all stimuli when eye-gaze indices were used but they showed significant AB only for personalised stimuli when reaction-time indices were used.

Comparisons of groups of non-dependent drinkers revealed that significant AB was shown by low alcohol-sensitivity participants but not by those with high alcohol-sensitivity (Shin, Hopfinger, Lust, Henry, & Bartholow 2010)[5], and that magnitude of AB was higher in heavy compared to light drinkers (Townshend & Duka 2001). Regarding the difference between heavy and light drinkers, Field, Mogg, Zetteler, and Bradley (2004) showed that this might be dependent on the component of attention (initial orientation or maintenance) examined. Heavy drinkers had an AB of greater magnitude compared to light drinkers when a long stimuli presentation duration was used (500ms and 2000ms) but not when short duration when used (≤ 200ms).That would indicate that although the two groups did not differ in terms of the extent to which their attention was initially attracted by alcohol-related cues, heavy drinkers found it more difficult that light drinkers to disengage their attention.

Despite the encouraging findings described above, other studies have failed to replicate these results, finding no significant AB in samples of non-dependent alcohol users with the VPT (Field, Duka, Eastwood, Child, Santarcangelo, & Gayton 2007; Manchery, Yarmush, Luehring-Jones, & Erblich 2017). Also, Groefsema, Engels, Kuntsche, Smit, and Luijten

---

[5] Sensitivity to alcohol refers to the extent that drinkers are susceptible to the effects of consuming alcohol and it is thought that low alcohol sensitivity is a risk factor for harmful alcohol consumption.

(2016) assessed AB of social drinkers using both simple pictures and complex social scenes and, in contrast to the study discussed above (Miller & Fillmore 2010), they found no significant AB, regardless of the pictorial context.

Other studies have examined AB with the alcohol Stroop-task (for a review see Cox et al. 2006). Results indicated that heavy, but not light drinkers, showed significant AB for alcohol-related cues (Adams, Ataya, Attwood, & Munafo 2012; Field, Christiansen, Cole, & Goudie 2007), and that magnitude of AB was greater in heavy compared to light drinkers (Bruce & Jones 2004). Magnitude of AB has also been found to be greater in hazardous and harmful drinkers compared to social drinkers (Fadardi & Cox 2009). Finally, significant AB has been found in adolescent offspring of alcohol dependent parents (as an indication of familial risk for hazardous alcohol use) but not those of healthy parents (Zetteler, Stollery, Weinstein, & Lingford-Hughes 2006).

Again, there are also studies which contradict the findings presented above, as they did not find significant AB in non-dependent drinkers using the alcohol Stroop task (Cox, Brown, & Rowlands 2003; Moss, Albery, Siddiqui, & Rycroft 2013). Also, Christiansen and Bloor (2014) used both personalised and general alcohol-related stimuli and they did not find significant AB for either type of stimuli.

AB in non-dependent drinkers has also been measured with other paradigms, such as the flicker ICB and eye-gaze detection in a naturalistic setting. Jones, Jones, Blundell, and Bruce (2002) used the flicker ICB in heavy and light social drinkers and found that heavy drinkers were more likely to detect the alcohol-related change than the change in the neutral object, compared to light drinkers. Monem and Fillmore (2017) utilized mobile eye-tracking techniques to examine AB in a naturalistic setting. Social drinkers entered a recreational room where alcohol and control objects were present, while wearing glasses embedded with eye-tracking technology. To control for the effects of naturally exploring novel environments, participants entered the room twice. Although the first time there was no difference in alcohol- and control-objects fixation times, participants fixated more on alcohol-related object when they entered the room again, indicating significant AB.

1.4.2.2 Studies with clinical samples.

Similar to non-dependent samples, several studies have used the VPT to examine AB for alcohol-related cues in alcohol-dependent samples. For example, Vollstädt-Klein, Loeber, Richter, Kirsch, Bach, Goltz et al. (2012) found significant AB for alcohol-related cues in alcohol-dependent patients, although there was no healthy control group.

Noël, Colmant, Van Der Linden, Bechara, Bullens, Hanak et al. (2006) and Stormark et al. (1997) used the VPT and ACT, respectively, to examine alcohol-dependent patients and healthy controls. These studies also used different stimuli presentation durations to examine different components of attention (initial orientation of attention or maintenance of attention). Patients showed significant AB in the short (50ms and 100ms) but not long (500ms) stimuli presentation durations, while the opposite was observed for control participants. Noël et al. (2006) included also a condition where stimuli were presented for 1,250ms, in which case both groups showed no significant AB and there was no group difference in magnitude of AB. Overall, these results would suggest that alcohol-dependent patients have a bias for initially orienting their attention towards alcohol-related stimuli but not for maintaining their attention on these stimuli, while the opposite was true for the control participants.

However, there are numerous examples where these results have not been replicated. A number of studies using the VPT failed to detect a significant AB for alcohol cues in alcohol-dependent patients (Garland, Boettiger, Gaylord, Chanon, & Howard 2012; Garland, Franken, Sheetz, & Howard 2012; Loeber, Vollstädt-Klein, Goltz, Flor, Mann, & Kiefer 2009). Loeber et al. (2009) detected significant AB only in a sub-sample of alcohol-dependent patients with a shorter duration of dependence (less than eight and a half years), although no significant AB was found in the full sample of dependent patients. Interestingly, Townshend and Duka (2007) found that alcohol-dependent patients had a significant negative AB, which would indicate that patients were automatically allocating their attention on the control stimuli, avoiding the alcohol-related stimuli. Previous studies also indicate no difference in magnitude of AB between alcohol-dependent and control participants (Field, Mogg, Mann, Bennett, & Bradley 2013; Wiers et al. 2017). Field et al. (2013) even examined this for different components of attention (stimuli presented for 200ms, 500ms and 2,000ms) and found no significant differences between the two groups.

Other studies have used the alcohol Stroop-task to examine AB for alcohol-related cues in alcohol-dependent participants (for a review see Cox et al. 2006). For example, Snelleman et al. (2015) found significant AB for alcohol-related cues in a sample of alcohol-dependent patients, although there was no healthy control group. Several studies that examined groups of dependent patients and healthy controls found significant AB for alcohol-related cues in the patient group but not the controls (Field et al. 2013; Johnsen, Laberg, Cox, Vaksdal, & Hugdahl 1994; Sharma, Albery, & Cook 2001;Stetter, Ackermann, Bizer, Straube, & Mann 1995; Stormark, Laberg, Nordby, & Hugdahl 2000), as well as greater magnitude of AB in the patient group compared to controls (Lusher, Chandler, & Ball 2004; Stetter et al. 1995). Cox, Blount, and Rozak (2000) compared magnitude of AB for alcohol-related and concerns-related stimuli (stimuli related to individual concerns in areas of personal life, e.g. divorce) in alcohol-dependent patients and healthy controls. They found that magnitude of AB in the patients group was greater for alcohol-related cues rather than concerns-related cues, although no difference was found in the control group. This indicates that alcohol-related stimuli attracted attention of dependent patients to a greater extent compared to other stimuli related to patient's current concerns.

Despite the results discussed above, contradicting results have been revealed by other studies using the alcohol-Stroop task in alcohol-dependent samples. Studies examining alcohol-dependent patients and healthy controls found that both groups demonstrated significant AB for alcohol-related cues (Cox, Hogan, Kristian, & Race 2002; Ryan 2002) and that there was no difference in magnitude of AB between the two groups (Bauer & Cox 1998).

Finally, another paradigm that has been used to examine AB in alcohol-dependent patients is the flicker-ICB task. Jones et al. (2006) studied alcohol-dependent patients and healthy controls and found that dependent patients, unlike controls, demonstrated significant AB for alcohol-related cues.

1.4.2.3 Summary and concluding notes

The studies presented above demonstrate that a great body of literature has examined AB for alcohol-related cues in participants using alcohol either socially or at harmful levels using a variety of tasks (a number of reviews offer comprehensive summaries, e.g. Cox et al. 2006; Field & Cox 2008; Franken 2003; Robbins & Ehrman 2004). Taken together, findings are equivocal. Some of the results would suggest that level of alcohol involvement (heavy or light drinkers), as well as clinical status (non-dependent drinkers or alcohol-dependent patients) is associated with differences regarding both the presence and magnitude of AB. More specifically, alcohol-related cues attract the attention of heavy drinkers and dependent drinkers to a greater extent compared to light and non-dependent drinkers respectively. However, these patterns are inconsistent, as several studies have not replicated these results.

It is possible that the inconsistent results are attributable to methodological factors and the poor psychometric properties of the tasks. For example, splitting samples based on arbitrary cut-offs (e.g. shorter and longer duration of alcohol dependence) and introducing multiple variations of stimuli (e.g. simple or complex alcohol-related pictures, and control stimuli that are neutral or emotional or concerns-related) might further complicate interpretation of results and overall conclusions. Also, the documented poor psychometric properties of the tasks used to measure AB (particularly the VPT; see section 1.4.1.4) might reduce the ability of the measures to capture the associations of interest. However, these findings might also indicate erroneous conceptualizations in the incentive motivation models of addiction regarding the nature of AB.

1.4.3 Evidence for the association of AB with alcohol-use behaviours in non-clinical and clinical samples of alcohol-using participants.

In previous sections presentation of studies examining AB for alcohol-related cues was based on the task used to measure AB (e.g. VPT or alcohol Stroop-task). In this section, greater emphasis is given on the different alcohol-use indices and clinical outcomes (e.g. typical quantity of alcohol consumption, hazardous drinking habits, time until relapse, and chances of relapse) and the nature of the association between AB and alcohol use (e.g. cross-sectional associations as opposed to prospective), and therefore results from studies using different tasks are aggregated together.

1.4.3.1 Studies with non-clinical samples.

Several studies in non-dependent drinkers have found cross-sectional associations of reaction-time measures of AB with various indices of alcohol-use, such as increased typical quantity of alcohol consumption (Bruce & Jones 2004; Fadardi & Cox 2008; Fadardi & Cox 2009; Field, Christiansen, et al. 2007; for a review see Field & Cox 2008), decreased time since last drink and more hazardous drinking habits (Field, Christiansen, et al. 2007), and greater alcohol involvement (expressed as a combination of typical weekly quantity of consumption and hazardous drinking habits; Christiansen & Bloor 2014).

Comparable patterns have also been revealed in studies using eye-gaze measures of AB (which have been found to be more reliable compared to reaction-time measures; see section 1.4.1.4), where AB showed cross-sectional associations with increased quantity of annual alcohol intake (Roy-Charland, Plamondon, Homeniuk, Flesch, Klein, & Stewart 2017), increased number of drinking days (Miller & Fillmore 2010), increased number of drinks consumed and binge-drinking days (Miller & Fillmore 2010; Monem & Fillmore 2017), and more episodes of drunkenness (Monem & Fillmore 2017).

Cox, Pothos, and Hosier (2007) examined the prospective relationship of AB (reaction-time based measures) with alcohol-use behaviour. Measures of AB, number of drinking days, average weekly amount of consumption, and maximum weekly amount were taken at baseline, in 3-months follow-up, and in 6-months follow up. Results indicated that all alcohol-use indices reduced over the study period (naturally occurring reduction without intervention). Participants with low AB scores at baseline showed greater reduction in number of drinking days in 6-months follow up compared to those with high AB scores (based on midpoint split). The two groups did not differ in reductions in the rest of the alcohol-use indices. This finding suggests that magnitude of AB at baseline played a role in the number of times participants chose to initiate drinking but not in the quantity of alcohol consumed each time.

However, a large body of studies in non-dependent drinkers have not replicated these associations. Reaction-time based measures of AB showed no cross-sectional associations with typical quantity of alcohol consumption (Christiansen, Mansfield, et al. 2015; Field et al. 2005; Field et al. 2004; Jones et al. 2006; Miller & Fillmore 2010), general drinking patterns (Field et al. 2005; Field et al. 2004), time since last alcoholic drink (Cox et al. 2003; Field et al. 2005), hazardous drinking habits (Christiansen, Mansfield, et al. 2015; Sharma et al. 2001), and number of drinking days and binge-drinking days (Miller & Fillmore 2010). Even when using measurements with higher reliability than reaction-time tasks, as for example eye-tracking with personalized stimuli, no association has been found between AB and typical weekly quantity of consumption and hazardous drinking habits (Christiansen, Mansfield, et al. 2015).

There are also examples of studies which examining the prospective associations of reaction-time based measures of AB and alcohol-use behaviour, which contradict the positive results described above. Groefsema et al. (2016) measured AB at baseline and then lent smartphones to young adults for five weeks and asked them to report quantity of drinking every time they were prompted. Pieters, Burk, Vorst, Engels, and Wiers (2014) used a sample of drinking adolescents and measured AB at baseline and typical weekly alcohol consumption two years after baseline. Both studies found no associations between baseline measures of AB and prospective measures of alcohol consumption.

1.4.3.2 Studies with clinical samples.

Several studies have examined cross-sectional associations of AB and various alcohol-use indices or treatment-related variables in alcohol-dependent participants. For example, Garland, Boettiger, et al. (2012) found that AB was associated with number of drinks per day in the year before entering treatment, and Jones et al. (2006) and Noël et al. (2006) demonstrated that AB was associated with a larger number of previous treatment attempts (although Noël et al. 2006 examined this for different components of attention and a significant association was found between number of previous attempts and biases in initial orientation of attention, but not in maintenance of attention). However, other studies failed to find an association of AB with hazardous drinking patterns (Sharma et al. 2001), quantity of life-time alcohol intake (Wiers et al. 2017), duration of dependence (Noël et al. 2006), problem severity (Loeber et al. 2009), and time being in treatment (Garland, Boettiger, et al. 2012).

More importantly, studies have also examined the prospective association of AB with relapse to alcohol use after a period of abstinence. In Garland, Franken, and Howard (2012) participants received either of two types of alcohol intervention programs, and had their AB assessed at the end of the programs. Incidence of relapse (consuming any amount of alcohol in at least one occasion) was assessed in a six-month follow-up. Although average magnitude of AB did not differ between those who relapsed and those who did not, individual differences in magnitude of AB predicted chances and time of relapse (those with higher magnitudes of AB had more chances to relapse and to relapse sooner, compared to those with lower magnitudes). Cox et al. (2002) measured AB before and after an inpatient alcohol treatment program. Alcohol use was measured at a 3-months follow up. Patients were characterised as having a successful treatment if they retained abstinence or had only a brief episode of drinking[6] over the follow-up period, and as having a non-successful treatment in they relapsed or lost contact over the follow-up period. Findings demonstrated that average magnitude of AB did not differ between groups of patients. However, patients who had a non-successful treatment, but not those who had a successful treatment, showed a significant increase in AB from pre- to post-treatment.

Although these results would suggest that AB can be prospectively associated with alcohol treatment outcomes, other studies have disputed this. Field et al. (2013) assessed AB in alcohol-dependent patients during the first week of a six-week treatment program. Patients

[6] Definition of brief episodes of drinking is not clarified in the paper.

were required to remain abstinent over the duration of the program and their place was withdrawn if they missed two consecutive days of the program or if they consumed alcohol before completion of the program. AB did not predict whether patients were dropped out of the program, or for how long they stayed in the program. Snelleman et al. (2015) assessed alcohol-dependent patients undergoing a treatment program at baseline and in three follow-ups (one, two and three months later). Measures of AB were taken at baseline and in the first follow-up, while measures of relapse (defined as at least one episode of binge drinking) were taken in the follow-up sessions. Results revealed that those relapsing did not differ in magnitude of AB from those not relapsing, across both measurements of AB. Also, measures of AB were not associated with relapse latency (time interval between AB-assessment day and relapse day).

1.4.3.3 Summary and concluding notes

Taken together, findings are equivocal regarding associations of AB with alcohol-use behaviour in both non-clinical and clinical samples. Also, the wide variety of alcohol-related outcomes used in literature (e.g. quantity of alcohol consumption over different periods, number of drinking days, or number of binge-drinking days) and the unstandardized definitions of relapse (e.g. any amount of alcohol consumption after a period of abstinence, or an episode of binge drinking after a period of abstinence) make it more difficult to draw general conclusions.

More importantly, it remains unclear whether measures of AB can robustly predict treatment outcomes in groups of alcohol-dependent patients (e.g. relapsing after a period of abstinence). It is important to note that a recent review (Christiansen, Schoenmakers, & Field 2015) identified inconsistencies regarding the clinical relevance of AB (e.g. predicting relapse) across the substance-use literature, not only alcohol. More alarmingly, the review revealed methodological caveats in the studies providing positive results, as for example underpowered samples, unaccounted sources of type I error, and inconsistent indices of AB (e.g. reaction times from only a sub-section of the task instead of all trials, or use of error rates instead of reaction-times).

Overall, the inconclusive results regarding the association of AB with alcohol-use behaviour across all samples (clinical and non-clinical) may be attributed to methodological caveats in the studies and poor psychometric properties of the tasks used to measure AB (see section 1.4.1.4). However, they may also indicate that the nature of AB and its role in alcohol-use behaviour might have to be further explored under refined theoretical conceptualisations.

1.4.4 Reducing AB as an intervention for reducing alcohol-use behaviours

The main theoretical conceptualization of AB suggests that AB can play a causal role in alcohol-related behaviours, from social drinking to relapse after a period of abstinence (Field & Cox 2008). Experimentally, the most effective way to test this assumption is to manipulate AB and examine whether reductions in magnitude of AB would be associated with reductions in alcohol-use behaviour. ABM programs have been developed in an attempt to re-train automatic allocation of attention away from alcohol-related cues, often toward healthier alternatives. Two of the most common ABM programs are the ABM-Visual Probe Task (ABM-VPT, e.g. Schoenmakers, de Bruin, Lux, Goertz, Van Kerkhof, & Wiers 2010) and the Alcohol Attention-Control Training Program (AACTP, e.g. Fadardi & Cox 2009).

1.4.4.1 Attentional bias modification (ABM) training programs.

The ABM-VPT is a modified classic VPT (see section 1.4.1.2) where the probe replaces the control cues more frequently than the alcohol. In the classic VPT, where the purpose is to assess AB for alcohol-related cues, the probe appears 50% of the time after the alcohol picture and 50% after the control. In the ABM-VPT, the probe almost always replaces the control picture (100% control/0% alcohol, although 90% control/10% alcohol has also been used). Therefore, if the participants are to respond to the task accurately, they need to constantly shift their attention away from the alcohol cue and towards the control.

The Alcohol Attention-Control Training Program (AACTP; Fadardi & Cox 2009) is a colour-naming task, loosely based on the Stroop task, which requires participants to shift their attention away from the alcohol-related feature of the cue and towards a coloured element. The training includes three levels of increased difficulty. Initially, participants see an alcohol or control picture and they have to name the colour of the background, as fast and as accurately as possible. For the next stage, participants have to name the colour of a thin outline around the picture. Finally, participants are presented with a pair of an alcohol and control picture together and they have to name the colour of the outline of the control picture. Successful completion of each level of the AACTP requires constantly shifting attention away from the alcohol-related feature of the picture to the colour-related feature.

1.4.4.2 Evidence for ABM effectiveness in non-clinical samples of alcohol-using participants

Field and Eastwood (2005) and Field, Duka, et al. (2007) used ABM-VPT to manipulate AB in heavy drinkers in a single training session and examined the effect of training on alcohol-use behaviour. Field and Eastwood (2005) trained participants to avoid alcohol-related cues (probes mostly replacing control cues) or to attend to alcohol-related cues (probes mostly replacing alcohol-related cues; reversed ABM-VPT). AB (measured with the VPT) changed in the predicted direction, with participants showing reduced AB after avoid-alcohol training, and increased AB after attend-alcohol training. They also found that participants who have been trained to attend to alcohol showed higher *ad-libitum* alcohol consumption compared to those trained to avoid alcohol.

Field, Duka, et al. (2007) conducted a similar study but they used also a control-group which received no training, and they examined generalisability of training effects by measuring AB after training in a set of novel pictures (pictures not used in the training) and with tasks other than the VPT (the flicker-ICB and the alcohol Stroop-task). Their results were partially consistent with the study described above. They found that AB did not change in the control group after training, but it significantly increased in the attend-alcohol group. That effect generalized to novel stimuli but not to other tasks. However, AB was not reduced after training in the avoid-alcohol group[7], and there were no post-training differences among groups in *ad-libitum* alcohol consumption. A surprising result was that AB in the novel set of stimuli increased after training in the avoid-alcohol group.

Other studies used only the regular ABM-VPT (no reversed training to attend to alcohol) in non-dependent heavy drinkers (not necessarily seeking help for their drinking) to examine whether retraining AB away from alcohol-related cues would also lead to reduced alcohol-use behaviours. Schoenmakers, Wiers, Jones, Bruce, and Jansen (2007) assigned heavy drinkers to a single-session ABM-VPT training group or control group (completing a AB assessment with no training elements) and found that AB was reduced in the ABM group but not the control group. However, training effects did not generalize to novel alcohol stimuli or to AB measured with the flicker ICB, suggesting only task-specific effects. They also found no effect of training on choice preference for alcohol over soft-drinks. Comparable results were also found in a study (Boendermaker, Sanchez Maceiras, Boffo, & Wiers 2016) where participants completed four sessions of either a regular ABM-VPT or a

[7] Authors report a trend for reduced AB after training when AB was measured with the VPT. However, the reduction did not reach statistical significance.

gamified ABM-VPT (the task contained gamification elements like attractive graphics and winning points), or no-training at all (control group). Results showed that AB decreased after training in the regular ABM-VPT, but not in the other two conditions. However, reductions in AB were not accompanied by reductions in typical quantity of alcohol consumption or number of binge-drinking days.

In another study (McGeary, Meadows, Amir, & Gibb 2014), heavy drinkers not looking to reduce their drinking completed the ABM-VPT training or control sessions (no-training element) outside the laboratory in a naturalistic environment (training completed in their home computers), twice per week for four weeks. Results showed that frequency of drinking was reduced in the ABM group but not in the control. However, the study does not offer a comparison of AB reduction between the ABM and the control group. Moreover, the outcome variable refers only to frequency of alcohol consumption, providing no information about quantity. Although reduction of frequency is a promising result, without any quantify information it is not possible to evaluate whether participants consume less alcohol after the training. Taking an extreme case, participants could consume the same volume of alcohol over fewer drinking episodes, implying that heavier drinking is taking place, which would not be considered as a reduction in heavy drinking.

Finally, Cox, Fadardi, Hosier, and Pothos (2015) used the AACTP to examine the effectiveness of ABM training in reducing alcohol consumption. Heavy drinkers motivated to reduce their drinking completed four sessions of AACTP or no training (waiting control group), and their typical quantity of alcohol consumption was measured pre-training, post-training, and in three- and six-months follow-ups. Effect of training on reducing alcohol consumption was found in the three-month follow-up but not in post-training or in six-month follow up, showing that the training effect varied across time points.

1.4.4.3 Evidence for ABM effectiveness in clinical samples of alcohol-using participants

Some studies have used the ABM-VPT to examine ABM effects on AB and alcohol-use behaviours in alcohol-dependent and alcohol-abusing participants. For example, in Schoenmakers et al. (2010) alcohol-dependent participants completed five sessions of either an ABM training or a control-training (regular AB assessments with not training element). AB was assessed before and after the training with the VPT using different stimuli presentation durations (200ms and 500ms) to examine both initial orientation of attention and maintenance of attention. The ABM group showed a greater reduction in maintenance of attention bias compared to the control group, although no difference was found for bias in initial orientation. Training effects also generalised to a set of novel stimuli. Participants who received the ABM training were discharged earlier from the program and avoided relapse for a longer period compared to control participants. However, this finding should be interpreted with caution as it was found in exploratory analysis based on a very small sub-sample (N=17).

However, these results were not replicated by another study using the ABM-VPT (Clerkin, Magee, Wells, Beard, & Barnett 2016), where alcohol-dependent participants completed eight sessions of either an ABM-VPT training or control-sessions with not training element. AB scores were calculated both traditionally and with trial level – based scores (TL-BS; see section 1.4.1.4), and various alcohol-related outcomes were measured (quantity of alcohol consumption and withdrawal symptoms). Results demonstrated that ABM had no effect on either of the AB scores. Alcohol-related outcomes reduced in both conditions, revealing no training effects too.

Other studies have used the AACTP to examine the effect of ABM on AB and alcohol-use behaviour in participants abusing alcohol. Fadardi and Cox (2009) used AACTP to train hazardous (two sessions) and harmful drinkers (four sessions). AB was measured with the alcohol-Stroop task pre- and post-training in both groups, and in three-months follow up in the harmful drinkers' group. Alcohol consumption was measured only in the harmful drinkers' group at all time points. AB reduced after training in both groups, although this reduction in harmful drinkers was not sustained after three months. The same pattern was found for alcohol consumption in the harmful drinkers. It should be noted that this study included no control group and therefore reduction in alcohol consumption could be attributed to factors not specific to training (e.g. decision to take part in the study and self-monitoring alcohol use to complete the diaries). However, the authors argue that the fact that they included a waiting phase before initiation of training, during which no reduction in alcohol consumption was detected, would increase confidence in the training effects.

However, the encouraging results regarding the effect of AACTP on alcohol consumption were not replicated in a naturalistic study (Wiers, Houben, Fadardi, van Beek, Rhemtulla, & Cox 2015) where patients abusing alcohol completed four sessions of AACTP or control-sessions (cognitive bias assessment with no training-element) in their home computers (training was provided online). Alcohol consumption did not reduce after training in either group. Notably, the study did not examine the effect of training on AB, so it was not evaluated whether AB reduced after training.

1.4.4.4 Summary and concluding notes

Overall, results regarding the effectiveness of ABM in both non-clinical and clinical samples are at best equivocal. The literature indicates inconsistencies in the extent to which ABM reduces AB, and in the extent to which training effects would generalise to novel stimuli and other AB-measuring tasks. The latter is important in order to determine if any observed effects can be attributed to cognitive retraining or to mere response-contingency learning. More importantly, results are also inconsistent regarding the effects of ABM on alcohol-use behaviour and relapse after a period of abstinence, which would cast doubts on the theoretical suggestions that AB may play a casual role in alcohol-use behaviours.

Comparable conclusions have been drawn by recent reviews and meta-analyses which did not support the overall utility of ABM in reducing substance-use behaviour (mainly focused on alcohol and tobacco use; Christiansen, Schoenmakers, et al. 2015; Cristea, Kok, & Cuijpers 2016). Pooling results from several different studies, Cristea et al. (2016) found that although ABM had a significant effect on AB for alcohol/tobacco-related cues, this did not led to reductions in the corresponding seeking and consuming behaviours. Christiansen, Schoenmakers, et al. (2015) outlined several factors that are important to consider when interpreting the inconsistent findings in the literature. First, training effectiveness may differ between single-sessions (e.g. Field, Duka, et al. 2007) and prolonged administrations (e.g. Boendermaker et al. 2016), as well as between change-oriented participants (e.g. heavy drinkers seeking to reduce consumption - Cox, Fadardi, et al. 2015) and non-concerned drinkers (university students used in many studies – e.g. Schoenmakers et al. 2007). Second, ABM might exert an indirect effect on substance-use behaviour, via changes in AB, even in the absence of a direct effect. Third, the task-specificity of the effects demonstrated in some studies (e.g. Schoenmakers et al. 2007) might suggest that reductions in AB are due to practice (participants become better at completing the specific AB-measuring task on hand), rather than cognitive modification. Finally, lack of proper control groups in some studies (e.g. Fadardi & Cox 2009) warrants some caution when interpreting the results.

Lack of robust effects of AB modification on alcohol-use behaviour adds to the equivocal findings regarding the association of AB measures with alcohol-use behaviour discussed in section 1.4.3. Taken together, these findings might suggest that the current conceptualisations regarding the nature of AB and its role in alcohol-use behaviour might have to be refined.

1.5 The state-motivational nature of AB.

1.5.1 The state hypothesis of AB.

The studies discussed above measured AB in laboratory/clinical settings and examined whether AB will be associated with alcohol-use behaviours which were measured retrospectively (e.g. time-line follow back questionnaires), in the form of general behaviour patterns (e.g. total scores in questionnaires like the AUDIT), or prospectively (e.g. two years later in a longitudinal study; Pieters et al. 2014) and in different contexts (e.g. alcohol-use behaviour after discharge from a rehabilitation setting). This approach inherently frames AB as a static signal, related to the general motivation of the person to seek and consume alcohol. Under this conceptualization, AB is expected to be associated with alcohol-use behaviour across time and location. However, given the inconsistent results and the lack of robust associations, AB may have to be conceptualised in a different way.

Recent work suggests that AB is a state-characteristic which fluctuates under the influence of momentary changes in the incentive value of alcohol and will be associated with alcohol-use behaviour when both AB and alcohol-use are measured within the same motivational context (Field, Marhe, & Franken 2014; Field, Werthmann, Franken, Hofmann, Hogarth, & Roefs 2016). In the context of incentive-motivational models of addiction, AB is described as an automatic response to the increased incentive value of alcohol-related cues (Franken 2003). Therefore, situational factors that would affect the incentive value of alcohol would also be expected to affect AB. For example, in a situation when temptation to consume alcohol is high and alcohol is expected to be readily available (e.g. getting the bus to a pub on Saturday afternoon to meet friends), magnitude of AB will be higher and it will be more strongly associated with alcohol consumption in the short-term, compared to a situation when there is little temptation to drink alcohol and alcohol is considered unavailable (e.g. driving to work on Monday morning).

These predictions are consistent with work regarding the value-driven nature of attention (Anderson 2013, 2015; Anderson, Laurent, & Yantis 2011) which demonstrated that involuntary allocation of attention to a given stimulus can be driven by the perceived value of that stimulus, even if it is irrelevant to the individual's current goals. Crucially, Anderson (2015) also showed that this value-driven allocation of attention is context specific. When the perceived value of the stimulus changes from one context to the other due to the different stimulus-reward associations activated in different environments, the ability of the stimulus to automatically attract attention will also change. For example, a driver might pay little attention to road-signs when following a familiar route (e.g. driving to work) compared to when driving around unfamiliar streets (e.g. during holidays in another country).

Taken together, the state hypothesis of AB suggests that current inconsistencies in the literature might be explained by the state-motivational nature of AB. If AB constantly changes under the influence of the motivational context, its association with alcohol-use behaviour may only be temporal and transient. As a result, there are little chances that measures of AB at a given time and context will be robustly associated with alcohol-use behaviour taking place at a different time and in different contexts. For example, take an alcohol-dependent participant in inpatient treatment. When still in the program, that person is highly motivated to abstain from alcohol, there is no availability of the substance and presence of alcohol-related cues is restricted. Upon discharge, the person will find themselves back in their old environments which are rich in alcohol-related cues, associating with old drinking peers, expecting alcohol to be readily available and probably experiencing elevated craving and stress, trying to cope with the transition. It would be reasonable to expect that AB measures in the former occasion will be a poor indicator of any alcohol-related behaviour expressed in the latter.

1.5.2 Evidence for the state-motivational nature of AB.

The state-hypothesis of AB would predict that magnitude of AB and its association with substance use will be influenced by the motivational context of where measures are taken. Also, if AB is a state that constantly fluctuates under the influence of the motivational context, it would be expected to vary in time and across different environment. The current evidence supporting these predictions are summarised below.

1.5.2.1 Effect of motivational and contextual factors

1.5.2.1.1 Direct manipulation of the substance incentive value.

Rose, Brown, Field, and Hogarth (2013) and Rose, Brown, MacKillop, Field, and Hogarth (2018) directly manipulated the perceived value of alcohol by asking participants to consume a normal alcoholic drink or an alcoholic drink tasting aversively bitter. In both studies, devaluation of alcohol reduced AB for alcohol-relate cues. Reductions in AB mediated the effect of alcohol devaluation on chances of choosing alcohol in a choice task. In Rose et al. (2018) reducing AB through alcohol devaluation was also found to reduce *ad-lib* alcohol consumption. These findings demonstrate that when participants perceived alcohol as less valuable, they showed a smaller magnitude of AB for alcohol-related stimuli and this also reduced their alcohol-use behaviour.

1.5.2.1.2 Craving and temptations to consume alcohol

Several studies have examined the association between craving and AB. Field, Munafò, and Franken (2009) conducted a meta-analysis, including both dependent and non-dependent samples on a variety of substances and found an overall significant, albeit weak, positive association between AB and craving. More specifically, they found the association was stronger for illicit drugs, direct measures of AB (e.g. eye-tracking), and when craving is high at the time of measurement.

Comparable results have also been found in individual alcohol studies. Field et al. (2013) administered the alcohol-Stroop and the VPT in alcohol-dependent participants and found an association between AB and craving, although only for VPT measures. Manchery et al. (2017) exposed non-dependent participants to alcohol-related cues and found that craving was higher after the exposure and that AB was associated only with post-exposure craving measures. This is consistent with Field et al. (2009) conclusion that associations between AB and craving will be stronger when craving is higher at the time of measurement. However, Wiers et al. (2017) measured AB with the VPT in alcohol-dependent patients and controls and found no associations with craving for either group.

Finally, Waters, Marhe, and Franken (2012) lent PDAs to drug-dependent patients and measured their AB (with a drug Stroop-task) repeatedly over the course of a week. Measures were taken at random times as well as when participants experienced temptation to use drugs (temptation episodes). Results showed that AB on temptation episodes was higher compared to random measures. Interestingly, within random measures, AB was higher when these measures were temporally proximal to a temptation episode. Results from the same study published elsewhere (Marhe, Waters, van de Wetering, & Franken, 2013) showed that AB in temptation episodes, but not in random measures, was associated with increased chances of relapse over the period that participants were using the PDA. Nevertheless, neither random nor temptation AB measures were associated with chances of relapse over the two weeks follow-up period. Taken together, these two studies demonstrate that the motivational context within which AB measures were taken (experiencing temptation or not) influenced magnitude of AB and its association with drug use in the short-term. Overall, the results discussed in this section would suggest that AB might be affected by the underlying motivation to drink alcohol.

1.5.2.1.3 Mood and stress

AB has been found to elevate when participants emotional state was aligned with their primary motivation to consume alcohol. Specifically, participants that consume alcohol as a coping mechanism showed elevated AB when exposed to a stressor compared to either a control situation (Field & Powell 2007; Field & Quigley 2009) or a positive situation (Grant, Stewart, & Birch 2007). Similarly, participants whose drinking was driven primarily by the ability of alcohol to enhance positive experiences, showed increased AB when positive affect was induced compared to stressful affect (Grant et al. 2007). These findings suggest that magnitude of AB is affected by participants' emotional state.

1.5.2.1.4 Initial doses of alcohol

Schoenmakers et al. (2008) provided heavy social drinkers with a placebo-alcohol drink or a low dose of alcohol (0.3 g/kg) to consume and after that they measured their AB with the VPT and concurrent eye-tracking. Participants showed AB for alcohol-related cues after low doses of alcohol but not after a placebo-alcohol drink (for both eye-gaze and reaction-time indices of AB).

Other studies have shown that these effects might be task specific. Duka and Townshend (2004) replicated the previous study but they also added a high-dose of alcohol (0.6 g/kg) and an additional measure of AB with the alcohol Stroop-task. They found a significant AB for alcohol-related cues only after the low-dose of alcohol (0.3 g/kg) and only in the VPT. Adams et al. (2012) used a placebo, a 0.13 g/kg dose, or a 0.4 g/kg dose of alcohol in heavy and light social drinkers, and they measured AB with the VPT and with the alcohol Stroop-task. In the VPT, a significant AB was found only after participants consumed the 0.4 g/kg-dose, regardless of drinking status (heavy or light drinkers). In the alcohol Stroop-task, heavy drinkers showed AB for alcohol-related cues regardless of dose condition, while light drinkers showed no AB at all.

However, Miller and Fillmore (2011) contradicted the previous results by showing that AB persisted after different doses of alcohol. Social drinkers were provided with a placebo, a low (0.32 g/kg) or high (0.64 g/kg) dose of alcohol and had their AB measured using the VPT with concurrent eye-tracking. Participant showed significant eye-gaze indices of AB across all dose conditions, although no reaction-time indices of AB were found.

In a naturalistic study, Schoenmakers and Wiers (2010) examined AB of customers in a pub with the flicker-ICB task. Participants were asked to report how many units of alcohol they typically consume per week, as well as how many drinks they have consumed that day. Although dose of alcohol consumed was not standardized in the same way as in laboratory studies, participants were split in low and high dose of alcohol, according to units of alcohol consumed that day (cut-off score was four units for males and three for females). Results showed that those who had consumed low doses of alcohol (less than three or four units), showed no association between the number of drinks consumed so far that day and AB. However, for high-dose customers (more than three or four units), the number of drinks they have consumed so far was negatively associated with AB. For both groups, typical alcohol consumption over the previous weeks was positively associated with AB.

Taken together, these findings suggest that consumption of alcohol might prime the motivational features of alcohol-related cues which would enhance the attention grabbing properties of these stimuli (for a review see Field, Wiers, Christiansen, Fillmore, & Verster 2010). It is notable that the effects did not generalize across different measured of AB. Some of the studies also suggest that these effects are dose specific. Very low doses of alcohol (e.g. 14g/kg) might not reach the threshold for activation of alcohol-induced changes in automatic allocation of attention. Also, although middle doses might induce AB, under larger doses individuals might feel satiated and less motivated to seek and find alcohol, which might lead to alcohol-related cues momentarily losing their incentive value and their attention-grabbing properties (although impairments in task performance due to intoxication might also account for this finding).

1.5.2.1.5 Exposure to alcohol-related cues and alcohol expectancy

Exposure to alcohol-related stimuli might affect AB even in the absence of actual consumption. For example, participants showed increased AB when they saw or smelled alcohol, compared to when seeing or smelling water (Ramirez, Monti, & Colwill 2015). Also, cross-sectional associations between typical alcohol consumption and AB have become apparent after people were exposed to alcohol-related cues but not when exposed to control cues (Cox et al. 2003), although this was only seen in heavier drinkers.

It is also argued that individuals will show higher reactivity to alcohol-related cues when they perceive the substance to be readily available for consumption (for a review see Jędras, Jones, & Field 2013). For example, Jędras, Jones, Stancak, and Field (2018) measured participants AB for alcohol-related stimuli with an eye-tracking task where information on the probability of receiving a reward (beer or a chocolate-bar) was provided at the beginning of each trial (i.e. 100%, 50% or 0% chances of receiving the reward).

Magnitude of AB for alcohol-related cues was greater when participants were certain they would receive beer (100% chances trials) compared to when they expected no beer (0% chances trials). However, there was no difference between being certain and being unsure (50% chances trials). Notably, expectation of receiving the chocolate-related reward did not affect AB for alcohol-related cues, showing that the effect of expectation on AB was reward specific.

However, these results were partially replicated by another study (Jones, Hogarth, Christiansen, Rose, Martinovic, & Field 2012) which also measured eye-gaze indices of AB for alcohol-related cues after manipulating expectancy of alcohol- or chocolate-related rewards. Although they found that magnitude of AB was greater when participants were certain that would receive beer (100% chance trials) compared to when they expected no beer (0% chances trials), they also found that anticipation of either reward (chocolate or beer) increased AB for both alcohol- and chocolate-related cues (non-specific expectation effects). Regardless of expectancy rate, AB for alcohol or chocolate was not associated with indices of alcohol or chocolate consumption.

Notably, Field, Hogarth, Bleasdale, Wright, Fernie, and Christiansen (2011) demonstrated that alcohol-expectancy effects on AB might depend on the level of alcohol involvement (i.e. heavy as opposed to light drinkers). Using a design similar with the studies described above (but only with alcohol-related rewards and no chocolate), they found that heavy drinkers showed significant eye-gaze indices of AB for alcohol-related rewards, regardless of rate of expectancy. However, light drinkers showed significant AB when it was certain that their will receive alcohol (100% trials), but not when they expected no alcohol (0% trials) or when unsure (50% trials). Taken together, the findings discussed in this section would suggest that AB might be influenced by the presence of alcohol cues in the environment, as well as by the expectancy that alcohol will become available for consumption.

1.5.2.2 Inter-trial variability

Conceptualisation of AB as a motivational state would imply that magnitude of AB would fluctuate in time, showing momentary increases and decreases. Indeed, calculation of AB scores with the trial-level based scores (TL-BS see section 1.4.1.4) showed that within a single VPT task AB exhibits momentary increases and dips, rather than behaving like a static signal (Zvielli et al. 2015). It was also shown that TL-BS scores had better reliability than traditional scores[8] (although this was not replicated by Jones, Christiansen, et al. 2018).

Evidence has also been found that indices of AB variability might be better predictors of substance-use behaviour, compared to traditional scores. Zvielli et al. (2015) assessed the AB of smokers with the VPT and they calculated both traditional and TL-BS scores. They found that that TL-BS measures explained additional variance in daily number of cigarettes smoked, over traditional AB scoring. Another study (Gladwin 2017) administered the VPT in a sample of non-dependent drinkers and calculated the traditional reaction-time based AB score as well as a measure of AB variability across trials. For the latter, an individual AB score was calculated for each pair of alcohol and control stimuli (e.g. reaction time in the fourth trial where the probe replaced the neutral picture subtracted by reaction time in the fourth trial where the probe replaced the alcohol picture) and the standard deviation of these scores was used as a measure of AB variability across trials. Results showed that AB variability score, but not traditional score, was associated with measures of risky drinking. These results indicate that AB can change rapidly in short periods of time and that individual differences in temporal AB variability might be important indicators of substance use behaviour.

1.5.2.3 Discrepancies across testing environments

Many of the contextual and motivational factors that have been shown to affect AB (e.g. alcohol-related cues and perceived availability of alcohol; see section 1.5.2.1) would be expected to be more prominent in the natural environment where participants use the substance (i.e. real-world conditions), rather than in neutral laboratories and clinical setting (unless specific experimental manipulations are introduced – e.g. alcohol exposure protocols). As a result, measures of AB in neutral laboratories will expect to differ from measures taken in the real world.

Results from previous studies support this assumption. In a sample of smokers attempting to quit, Begh et al. (2016) examined the association between clinic-based and real-world measures of automatic selective attention. Participants had their AB assessed in the clinic with a tobacco Stroop-task and a VPT. Following that, they were asked to report over seven weeks how many smoking-related cues they have spotted, and to what extent they felt their attention was attracted by smoking-related cues. Results showed no association

[8] Average reaction time in trials where the probe replaced a neutral picture subtracted from average reaction time in trials where the probe replaced a substance-related picture.

between measures of AB in the clinic and real-world self-reported indices of attentional focus. Although the different types of measurements (implicit vs self-reported) warrant some caution, these are preliminary findings suggesting that measures of AB in a clinical environment were not indicative of how attention would function in the presence of substance-related cues in the real world.

In another example, smokers not seeking to reduce their smoking completed ABM-VPT or control-training sessions on personal digital assistants (PDAs) for one week, as they went about their lives (Kerst & Waters 2014). AB was assessed with the VPT on the PDAs over the training week, as well as in the lab at the end of the study. Although AB was reduced over the training week in the training group but not the control, when AB was measured in the lab at the end of the study no between-group differences were detected. This suggests that day to day changes in AB that were evident when AB was measured in the natural environment were not detected when AB was measured in the laboratory (but see also Robinson, Muench, Brede, Endrighi, Szeto, Sells et al. 2017 for a similar study where effects of training on AB were evident in both environments). Results from these two studies suggest that measures of AB in neutral laboratories might not provide an accurate indication of how AB is expressed in real-world conditions.

1.5.2.4 Summary and concluding notes

Overall, previous results suggest that AB might fluctuate across time and space, under the influence of the motivational context in which measures are taken. Direct manipulations of the incentive value of the substance, as well as changes in variables associated with the individual's internal state (e.g. craving, mood and stress) and substance-related contextual factors (e.g. exposure to substance-related cues and perceived availability of the substance) might influence automatic attraction of attention by substance-related cues. These factors might also influence the extent to which AB is associated with substance-use behaviour.

It would be reasonable to expect that these factors would constantly change in the real world causing constant fluctuations in AB. For instance, a person might start their day at work, with a very low urge to drink alcohol, being in an environment devoid of alcohol cues. The same person, joining some colleagues for drinks after work, might find their urges becoming stronger, expecting alcohol to be available very soon, and entering an environment with an abundance of alcohol-related cues. AB in the first time-point would be expected to be lower and less able to predict alcohol-use behaviour, compared to the latter.

1.6 Recapitulation of results from AB research and future directions

Incentive-motivation models of addiction (see section 1.2.2) suggest that cues consistently paired with alcohol consumption will, through classical conditioning, acquire the same rewarding properties for the user as alcohol itself. Through this mechanism, alcohol-related cues acquire heightened incentive value and therefore get attentional priority over other stimuli, automatically grabbing and holding attention (Attentional Bias; Franken 2003). It is expected that AB for alcohol-related cues will be able to differentiate among distinct levels of alcohol involvement (from light drinkers, to heavier and dependent drinkers), and it will be associated (either cross-sectional or prospective associations) with drinking habits or chances of relapse after abstinence (Field & Cox 2008). AB is also expected to play a causal role in alcohol-use behaviours and, therefore, active manipulations of attention towards or away alcohol (via ABM programs) are expected to increase or decrease, respectively, alcohol consumption and rates of relapse (see section 1.8).

So far, the reviewed empirical evidence is inconclusive and the clinical relevance of AB has been hotly debated (for a comprehensive discussion on the topic see Christiansen, Schoenmakers, et al. 2015). This thesis proposes that there are two possible explanations for the mixed findings. First, it has been shown that the most popular tasks for measuring AB suffer from low reliability and methodological flaws (see section 1.4.1.4). Although direct measures of attention with eye-tracking seem more promising in that aspect, reaction-time indices of AB still dominate the literature. Also, it has been repeatedly shown that results regarding AB do not generalize from one measuring paradigm to another (e.g. from the VPT to Stroop), indicating that convergent validity is another weak spot of the current measurement tools. Given these psychometric problems, it is unsurprising that it has been proven difficult to robustly show that AB is associated with the behaviours of interest.

Second, most of the studies examining the association between AB and alcohol behaviour, tend to measure AB in a given environment at a certain point in time. They then relate these AB measures with alcohol behaviours that take place in a difference environment and at a different time, either prospectively (e.g. treatment outcome), retrospectively (e.g. consumption over the past weeks) or in a cross-sectional manner (e.g. global questionnaires). This approach inherently frames AB as a static signal, related to the general motivation of the person to seek and consume alcohol. Under this conceptualization, AB is expected to be associated with alcohol use behaviour across time and location.

However, the state hypothesis of AB (Field et al. 2014; Field et al. 2016) suggests that AB might best be conceptualised as a state-like phenomenon, which would fluctuate under the influence of the current motivational context in which measures are taken. Therefore, its association with alcohol use behaviour may only be temporal and transient. Several studies have demonstrated that magnitude of AB, as well as its association with substance-use behaviour, might be affected by various motivational and contextual factors (see section 1.5.2). As a result, it is not surprising that measures of AB taken in laboratory or clinical settings show little concordance with alcohol-use which takes place at a different time (e.g. retrospective recall) and within different contexts (e.g. personal environments after discharge from clinical settings).

Discussing the state of the evidence from AB research, Christiansen, Schoenmakers, et al. (2015) and Field et al. (2016) suggest that future research on AB should take on a different approach. First, they propose that future studies should focus on the examination of AB shortly before the substance-use behaviour occurs and within the same contexts where the substance is used. Second, a suitable methodological approach to that end would be the examination of AB outside the laboratories and in the natural environments where the substance is consumed, taking multiple measures of AB in different points in time (e.g. temporally proximal to or distant from the substance-use behaviour) and within different motivational circumstances (e.g. in an episode of temptation or in random control moments). Ecological momentary assessment (repeated measures of a phenomenon in the natural environment where the phenomenon occurs; Shiffman, Stone, & Hufford 2008) would, therefore, provide a suitable methodological framework for future AB research.

1.7 Ecological Momentary Assessment (EMA)

1.7.1 EMA and its suitability for substance-use research.

Ecological momentary assessment (EMA), refers to the repeated assessment of phenomena in the natural environment where they occur (Shiffman et al. 2008). Participants involved in EMA studies provide data regarding a specific psychological construct (e.g. behaviours, emotions, thoughts, etc.) under real-life circumstances (the ecological aspect) and their recordings represent real-time, current manifestations of the construct (the momentary aspect). Ecological Momentary Intervention (EMI) follows the same principles as EMA with the additional target of delivering real-time interventions, rather than just assessments (Heron & Smyth, 2010).

EMA can be organized either around specific events of interest (event-related design), certain time-points (time-related design), or a combination of the two. In the former, participants make a data entry every time a certain events occurs (e.g. strong temptation to use the substance), while in the latter an assessment may be prompted any time, either randomly (e.g. text-messages or a device beeping at random points) or at pre-defined time-points (e.g. specific times over the day). Each of these designs provides different research possibilities.

Event-related designs are useful at identifying the frequency in which the events of interest occur, as well as the situational context within which they tend to occur (e.g. participants' location and surroundings, as well as their emotions, motives, etc.). Time-related designs, on the other hand, are particularly good at identifying how a psychological phenomenon unfolds in time, measuring peaks, dips and overall fluctuation. Lastly, combining the two designs provides two advantages. First, random-time measures can serve as a comparison level for event-related measures, considered as "control-moments" when the event is not present. Second, random-time measures that closely precede or follow the event-related measures can reveal triggers or immediate consequences of the event. For example, if participants' emotional state is evaluated when they use the substance, as well as in random moments when the substance is not used, the random measures may offer a reference level for the measures obtained when using the substance. Also, random measures that happen to be temporally proximal to substance-use (either before or after), may provide useful information on participants' emotional state prior to using the substance and shortly after.

The time-bound and context-related nature of EMA makes this methodology suitable for the study of substance use (Shiffman 2009). This is because incidents of substance use usually come and go in an episodic manner during the day and, as discussed in previous sections (1.5.2.1), motivation to use the substance is closely related to various internal and contextual factors (e.g. mood, availability of the substance, substance-related cues, etc.). EMA is increasingly used in the addiction literature to examine associations between substance-related behaviours and mood (e.g. Vinci, Li, Wu, Lam, Guo, Correa-Fernández et al. 2017; for a review see Shiffman 2009), craving (e.g. Fatseas, Serre, Alexandre, Debrabant, Auriacombe, & Swendsen 2015; for a review see Serre, Fatseas, Swendsen, &

Auriacombe 2015), fluctuations in inhibitory control (e.g. Jones, Tiplady, Houben, Nederkoorn, & Field 2018), social context (e.g. Phillips, Phillips, Lalonde, & Prince 2018), and advertising (e.g. Martino, Kovalchik, Collins, Becker, Shadel, & D'Amico 2016). EMA has also been used to understand how initial lapses escalate into a full relapse and which factors might be driving an abstinent individual back to substance use (for a review see Shiffman 2009).

Although most of the EMA studies in addiction have relied on repeated administration of self-report questionnaires, Waters and Li (2008) has also shown that administration of cognitive tasks, sometimes measuring implicit constructs not accessible through self-report, is also feasible (e.g. high completion rates and prompt reaction to assessment notifications). For example, attempts have already been made to measure AB in the real world (e.g. Marhe, Waters, van de Wetering, & Franken 2013), which will be discussed later.

1.7.2 EMA methodological strengths and considerations

1.7.2.1 Methodological advantages

There are numerous papers providing extensive discussions and reviews on the methodological advantages and considerations of EMA, which will be summarised here (Shiffman 2009; Shiffman et al. 2008; Moskowitz & Young 2006; Heron & Smyth 2010). First, EMA can overcome the common pitfalls of retrospective measures. When participants are asked to recall experiences, their reports are often inaccurate and biased, based more on heuristics rather than actual recollections (e.g. events that are more vivid in memory are also perceived as more frequent). Evidently, although there seems to be good agreement between EMA and consumption-recall diaries regarding averaged substance consumption across days, there are only weak correlations regarding day to day changes in patterns of consumption (for a review see Shiffman 2009). Nevertheless, it is worth mentioning that several EMA designs also use retrospection to some extent (e.g. a diary-like entry at the end of the day), although the time-intervals are usually quite short.

Second, EMA studies have challenged findings derived from global questionnaires (questionnaires assessing general patterns of behaviour). For example, Shiffman (2009) reviewed EMA studies attempting to replicate the seemingly robust finding that negative reinforcement (e.g. avoiding negative affect) is one of the primary motives for smoking and drinking behaviour. Regarding smoking, they found that measures of negative-reinforcement motives derived from global questionnaires were unrelated to EMA measures of smoking motives. Results from the EMA studies revealed that negative reinforcement was the least frequently reported motive for smoking. Regarding alcohol consumption, results from the EMA studies demonstrated that in participants who describe their drinking behaviour as a coping mechanism, measures of negative emotions were not associated with alcohol use behaviours. Moreover, in some EMA studies where drinking to cope was found to moderate the association between negative emotions and alcohol use, this was true only for a subset of negative emotions and on certain days during the week.

Finally, EMA can facilitate research participation of hard to reach populations with restricted access to traditional research or treatment settings. Access barriers might be either pragmatic (e.g. participants residing in rural, remote areas), health-related (e.g. older individuals or individuals with mobility difficulties) or due to the exceptional circumstances in participants' way of living (e.g. substance-dependent individuals, homelessness, etc.).

1.7.2.2 Methodological challenges

Despite the important advantages discussed above, there are also some considerations (Shiffman 2009; Shiffman et al. 2008; Moskowitz & Young 2006; Heron & Smyth 2010). First, when participants are required to complete multiple assessments over long periods of time, compliance rates and adherence to the study protocol is an important challenge. Lack of compliance can increase missing data and induce bias in the data, especially when data-entries are skipped in a non-random manner. A recent meta-analysis found that the pooled compliance rate in EMA studies with substance-using populations was 75.06% (Jones, Remmerswaal, Verveer, Robinson, Franken, Wen et al. 2019), which was lower than the 80% threshold suggested by the authors (although they also note that there is currently no consensus regarding a gold-standard compliance rate threshold). However, a study conducted in a challenging population (crack-cocaine addicted homeless participants) provided encouraging results (Freedman, Lester, McNamara, Milby, & Schumacher 2006). Participants were lent mobile phones to complete EMA assessments for two weeks. Eighty percent of participants completed the full two-week protocol and authors reported a very low number of incidents of study equipment going missing (only one phone and two battery chargers were missing at the end of the study).

Participants in EMA are acting independently as they go about their lives, without close researcher supervision, and as a result, monitoring of compliance is challenging. This might apply to a greater extent to event-related designs, where it is difficult to verify to what extent the event of interest occurred, but participants omitted to complete the assessment. A possible solution could be the use of wearable devices that can provide objective measures. For example, a plethora of such devices are available to obtain biomarkers of substance use (e.g. sweat-based wearable detectors and biosensor tattoos), although cost and data accuracy could be an issue (Mahmud, Fang, Carreiro, Wang, & Boyer 2018).

However, time-related designs might also be affected, as participants may forward- or back-file paper diaries and then falsely report times and dates of completion. Researches may be able to overcome this issue by using electronic devices that can timestamp the assessments. These considerations have been closely captured in a study by Stone, Shiffman, Schwartz, Broderick, and Hufford (2003) where they asked participants to record levels of pain using either a paper-based or electronic diary. Paper-diaries were equipped with light-sensors detecting when the diary was opened (a feature concealed from the participants), and electronic diaries digitally timestamped the assessments. A striking finding was that although compliance rate according to dates and times written by participants on the paper diary was 90%, timestamping from the light detector revealed a compliance rate of just 11%, suggesting that participants were falsifying compliance data to high rates. Compliance in electronic diaries was 94%, which could be explained by an increased sense of supervision (e.g. it is self-explanatory that a digital device would provide digital timestamps) and other facilitating factors (e.g. frequent reminders).

Another critical issue is reactivity to the measures, which refers to changes in participants' behaviour caused by the close self-monitoring of the behaviour (e.g. daily recording quantity of alcohol consumption). The idea that self-monitoring one's behaviour may act as a behaviour-change mechanism is also reflected in the inclusion of self-monitoring as an integral part of serval behaviour-change treatment protocols (for example see Boutelle & Kirschenbaum 1998). Any deviations from how participants would normally behave in their natural environment would undermine the ecological aspect of EMA. However, empirical data do not support the presence of reactivity effect in EMA studies in substance-using populations (Hufford, Shields, Shiffman, Paty, & Balabanis 2002; Simpson, Kivlahan, Bush, & McFall 2005). One possible explanation is that the continues repetition of assessments brings an element of habituation, which may attenuate any reactivity effects (Shiffman 2009).

1.8 Ecological studies of AB

A central argument in this thesis is that attentional bias is another substance-use related phenomenon that could be better understood if examined with EMA. As described earlier (see section 1.5) AB may be a cognitive marker of the current motivation status and, as such, would be susceptible to rapid changes in contextual and internal factors. Moreover, any prospective associations with substance-use behaviour might be expressed only if AB is measured under the appropriate motivational circumstances (e.g. when the substance's incentive value is high) and in close temporal proximity to the behaviour.

This section will focus primarily on real world studies which utilized portable electronic devices (i.e. smartphones or personal digital assistants - PDAs) to administer a series of either AB assessments (EMA studies) or attentional bias modification (ABM) training sessions (EMI studies). Although not strictly defined as EMI, studies where participants completed ABM programs on their desktop-computers at home will also be discussed briefly. Although these studies undeniably entail an element of ecological validity, testing is confined at specific times when participants are at home and in the specific room where the computer is based. Another reason for focusing primarily on portable devices is to provide previous examples of the main methodology used in the thesis (AB assessments completed in smartphones).

1.8.1 Re-training AB in the real world: EMI studies

Two studies (Kerst & Waters 2014; Robinson et al. 2017) have delivered ABM in smokers not attempting to quit in participants' natural environments utilizing portable electronic devices. In Kerst and Waters (2014) , participants completed three training sessions (ABM with the ABM-VPT or control training) and one assessment session each day for one week. The assessment session included measures of AB (with the VPT), craving after exposure to smoking-related cues (cued-craving) or without exposure (non-cued craving), likelihood of smoking, and time-intervals between smoking episodes. At the end of each day participants also reported how many cigarettes they have consumed that day in a paper diary. Before and after the EMI phase of the study, participants visited a lab where AB, craving, and biomarkers of smoking were assessed.

Findings showed that over the EMA period, magnitude of AB decreased in the ABM group but not the control, although the difference between groups became significant only after the fifth day of training. Notably, the groups did not differ when AB was assessed in the lab after the end of the EMI phase. Regarding craving, there was no significant day to day decrease over the EMI period in either group. However, when craving scores were averaged across all days in the EMI period, the ABM group showed lower overall craving compared to control, although this was only for cued craving. Groups did not differ when craving was measured in the lab after the EMI phase. Finally, the effect of ABM in AB and craving over the EMI phase was not accompanied by reductions in likelihood or quantity of smoking. Groups also showed no difference in biomarkers of smoking assessed in the lab after the EMI phase.

In Robinson et al. (2017) participants completed four PDA sessions each day for four weeks. Three of the sessions included ABM training (with the ABM-VPT) for the training group or a control procedure for the control group, while the fourth session included an AB assessment (with the VPT) for all participants. To test for generalizability of training effects, one of the AB assessments each week contained novel stimuli (stimuli not used in the ABM training). Every PDA session included also measures of the extent to which participants felt their attention was attracted by smoking-related cues (self-reported AB), cued and non-cued craving, and quantity of cigarettes smoked since the last session. At the end of each day participants also reported how many cigarettes they have consumed that day in a paper diary. Participants also visited a laboratory twice, once during the first EMI week and once during the second. The lab visits included measures of AB, self-reported AB, and biomarkers of smoking. Again, to test for generalisability, AB in the lab was assessed with novel stimuli and with a smoking Stroop-task.

Results demonstrated that AB measured with the VPT was lower in the training group compared to the control, regardless of whether it was measured in the natural environment or in the laboratory. Training effects were generalised to novel stimuli but not to different tasks. However, self-reported AB did not differ between the groups, neither in the natural environment nor in the laboratory. Despite the effect of training on AB, there was no training effect on craving, regardless of the testing environment. No training effect was also found on biomarkers of smoking measured in the lab and quantity of smoking recorded in the diaries over the EMI period. However, the smoking rate declined over days in the training group but not the control.

Results from these two studies suggest that EMI with ABM training programs might effectively reduce AB for substance-related cues. However, the studies provide contradicting results regarding the effect of training on substance use in participants' natural environments. Notably, the studies also found different results regarding the extent to which the effects of ABM delivered in the natural environment would generalise to AB measured in the laboratory. Although both studies found that AB in the natural environment was lower in the training group compared to the control, one study detected this difference also in the laboratory while the other found no difference in the laboratory. The state hypothesis of AB would suggest that measures of AB will differ across contexts with different motivational properties (e.g. neutral labs as opposed to environments where substance use naturally occurs), which might provide one explanation for the discrepancy in the two studies.

Interestingly, over the past few years, some ABM smartphone applications (apps) have become commercially available. For example, on an opinion paper, Cox, Intrilligator, and Hillier (2015) describe Chimpshop, a gamified alcohol-ABM app where participants control a chimp avatar as it strolls over a supermarket aisle. Alcohol-related and neutral products simultaneously appear on the screen and participants need to select only the neutral product to collect points.

Although Cox, Intrilligator, et al. (2015) provide some narrative information about the app's efficacy (e.g. 60% reduction in problematic drinking, as well as reduction in other alcohol side-effects, like physical side-effects, missed days at work and relational problems), no specific methodological information, statistical results or comparisons with control groups are provided. At the time this thesis is written, no additional publications regarding the efficacy of the Chimpshop app have been identified and a recent review on ABM smartphone apps (Zhang, Ying, Song, Fung, & Smith 2018) cites the 2015 opinion paper but no further relevant publications. Therefore, these efficacy claims should for the time being be discarded until specific findings have been published. Nevertheless, the app is presented here as an interesting example of the potentials of ABM programs to be adapted into a gamified applications and become available for download in commercial platforms (e.g. Apple iTune and Google Play - although see also Boendermaker et al. (2016) for a suggestion of caution when gamifying ABM).

Finally, although not strictly falling within the remit of EMI, some studies have delivered ABM on participants' home computers over multiple training sessions (Elfeddali, Bolman, Pronk, & Wiers 2016; McGeary et al. 2014; Wiers et al. 2015). In samples of hazardous non-dependent drinkers, McGeary et al. (2014) showed that home-delivered VPT-ABM reduced frequency of alcohol consumption (although see critiques in section 1.4.4.2) , while Wiers et al. (2015) found that ACCTP had no effect on typical volume of alcohol consumption. Notably, both studies did not report on the effects of training on alcohol AB. In a sample of heavy smokers attempting to quit, Elfeddali et al. (2016) found that VPT-ABM increased the chances of retaining abstinence for six months after completion of the training, although training had no effect on smoking AB. Taken together, these results indicate that ABM delivered at home might help reduce substance-use behaviour, although failures to replicate this are also reported.

1.8.2 Assessing AB in the real world: EMA studies

So far there have been few studies assessing AB for substance-related stimuli with EMA (Marhe et al. 2013; Waters, Marhe, & Franken 2012; Waters, Szeto, Wetter, Cinciripini, Robinson, & Li 2014). In Waters et al. (2014), smokers attempting to quit were lent a PDA for one week and completed assessments at random points (up to four per day), as well as every time they experienced temptation to smoke and every time they smoked (episode of relapse). Every PDA assessment included measures of craving and whether participants had smoked so far that day (measure of abstinence/relapse), while every other assessment included measures of AB (with a smoking Stroop-task) and self-reported attentional bias (participants self-reporting how strongly they felt their attention was drawn to cigarettes). At the end of the day participants recorded number of cigarettes smoked, in a diary. Results showed that across all assessments craving was associated with AB in a between-subjects manner but not within. This means that participants with generally higher scores of craving over the week also had generally higher scores of AB, while increases in craving during the day within the same individual where not associated with AB. The same pattern was found for the association between AB and self-reported AB. None of the reported associations was moderated by days in the study or as a function of abstinence. Also, the study did not report any comparisons between random assessment and temptation or relapse assessments.

Waters et al. (2012) and Marhe et al. (2013) report results from the same study on a sample of heroin and cocaine dependent inpatients receiving treatment. Participants were lent a PDA for one week[9] and they completed assessments at random points (up to four per day), as well as every time they experienced temptation to use drugs. Each PDA assessment

[9] Patients were residing in the treatment centre and they carried the PDA on them all day as they went about their normal daily routines. Although this setting does not represent an environment where drug use naturally occurs, the study may still be considered as EMA based on the fact that AB was assessed repeatedly within the current leaving environment of participants.

included measures of AB (with a drugs Stroop-task) and craving. Relapse (at least one episode of using drugs or dropping out from the program) was assessed with self-reports during the EMA phase of the study, and with self-reports and urine screens over the two-week follow-up period.

Waters et al. (2012) found that although participants showed significant AB at both random and temptation assessments, magnitude of AB was greater in the temptation assessments. Also, within random assessments, magnitude of AB was higher in assessment that were temporally proximal to a temptation episode (within one hour), compared to more distant assessments. Regarding craving, scores were higher when participants felt tempted to use drugs.

Marhe et al. (2013) presented comparisons among sub-groups of patients based on relapse status. Participants were coded as early-relapses (relapsed during the EMA period), non-EMA relapses (no relapse until the end of the EMA period), late-relapses (relapsed during follow up), and never-relapses (not relapsed by the end of the study). Results showed that there was no difference in the number of temptation episodes among relapse groups. Magnitude of AB when participants felt tempted to use drugs was greater in the early-relapse group compared to non-EMA relapse. There was no difference between the late-relapse and never-relapse groups. Also, AB when participants experienced a temptation episode was prospectively associated with higher chances of early relapse but not with chances of late relapse. Magnitude of AB measured in random assessments did not differ among groups and was not associated with chances of relapse. Craving did not differ among groups and it was not associated with relapse, regardless of whether it was assessed on temptation or random assessments.

The findings presented above demonstrate that measures of AB taken with EMA were associated with the underlying motivation to consume the substance (i.e. craving) and magnitude of AB was fluctuating under the influence of the motivation to consume the substance (e.g. greater AB in moments of temptation). More importantly, they reveal that AB can be prospectively associated with clinical outcomes (i.e relapse) only when measured in the appropriate motivational context (i.e. during episodes of temptation but not in random measures). Lab-based measures of AB in previous studies that were not robustly associated with relapse could be likened to the random AB assessments of EMA studies. In both cases, AB was measured under non-appropriate circumstances to reveal any strong associations with substance use. Crucially, the EMA studies discussed here also indicated that even when AB was measured under the appropriate circumstances, it showed short-term relationship with substance use (i.e. chances of relapse within the EMA period but not in the follow-up weeks).

Consistent with the state hypothesis of AB, these findings would suggest that AB is a short-living signal of motivation, which is stronger at certain points in time (e.g. during an episode of temptation or shortly before). The signal will then attenuate, providing no information for future behaviour, until the right circumstances are created again (e.g. another temptation episode) for the signal to be generated again.

1.9 Aims and hypotheses of the thesis

1.9.1 Literature gap

There are currently only a small number of EMA studies examining the state-motivational nature of AB and its association with substance-use, all of which focus on substances other than alcohol (e.g. smoking, heroin, and cocaine). Moreover, the psychometric properties of AB measures administered on portable digital devices in the real world have not been thoroughly examined. Therefore, this thesis seeks to develop an EMA program (based on a smartphone application) that measures AB for alcohol-related cues and craving. The psychometric properties of that program will be examined, and the program will then be used to examine variations in AB and craving in the real world and their association with alcohol consumption.

1.9.2 Aims

A novel smartphone application (app) was developed for the purpose of this thesis, which includes measures of AB (with the alcohol Stroop-task) and craving. The first aim of the thesis is to explore the psychometric properties of AB measures taken on the app in the real world. To that end, the thesis examined the internal reliability and predictive validity of the AB measures taken on the app in the real world and compared them with measures taken on different devices (e.g. laptop computers) and different environments (e.g. neutral laboratories, semi-naturalistic laboratories, and natural environments). Two variations of the program were examined, one assessing AB for alcohol-related words and the other assessing AB for alcohol-pictures that matched participants' personal preferences for alcoholic drinks.

The second aim is to explore the state-motivational nature of AB for alcohol-related cues and its transient association with alcohol consumption. Three EMA studies were conducted using a combination of time-related (i.e. measures obtained three times over the course of the day) and event-related designs (i.e. measures obtained in moments of temptation to drink alcohol, moments when alcohol was present and available for consumption, and control moments when none of these factors was present). The studies examined whether AB was associated with craving in the real world. They also examined whether magnitude of AB and craving, as well as their association with alcohol consumption, was affected by the context within which the measures were taken (e.g. temporal proximity to initiation of drinking, moments of temptation, and moments of alcohol availability).

Finally, the third aim was to explore attitudes of users of the app. Focus groups were organized to qualitatively explore the experiences of participants using the app and their opinions regarding usability and use-friendliness of the app. Providing that the previous studies would establish the scientific value of the app, the findings of the focus group could be used to inform a re-development of the app in the future into a commercially available digital behaviour-assessment tool.

1.9.3 Core hypotheses

The thesis posed three core hypotheses:

1. Magnitude of AB will fluctuate in the real world and it will co-fluctuate with craving.

2. AB and craving will be elevated shortly before participants consume alcohol. Also, AB and craving measured proximally to alcohol consumption, but not in distant assessments, will predict alcohol consumption.

4. AB and craving will be elevated when participants feel tempted to drink alcohol, and when alcohol is present and available for consumption. Also, AB and craving measured in these moments, but not in control assessments, will predict alcohol consumption.

Chapter Two

2. General Methods

Across all studies in the thesis, a smartphone application (app) was used to measure attentional bias (AB) for alcohol-related cues in the real world. The app was also used to measure intention to drink alcohol (chapter four) and craving (chapter four and five). Although the app was refined throughout the thesis, the basic structure and features remained the same. A description of the basic app parameters is provided below and then further modifications for each study are explained in the relevant chapter. Finally, typical alcohol weekly consumption and hazardous alcohol consumption were evaluated in all studies with the Time Line Follow Back (TLFB) and the Alcohol Use Disorder Identification Test (AUDIT), respectively. Analysis and data reduction strategies (e.g. scoring of AB, multilevel modelling) are also described in detail in this chapter.

<u>2.1 Questionnaires</u>

2.1.1 Time line follow back (TLFB)

The TLFB (Sobell & Sobell 1990) was used to assess typical alcohol consumption. Participants were required to estimate the number of UK alcohol units consumed over a given period (one UK unit = 8g of alcohol). In all studies in the thesis the TLFB was administered at baseline and covered the previous 14 days. To aid participants in estimation of UK units, the TLFB was accompanied by a table presenting servings and for the most common types of alcoholic drinks (e.g. pint of beer, small glass of wine, etc.), and the average number of units contained in each.

Daily estimation methods like the TLFB have advantages such as great face validity, provision of rich data, and time-bound observations, although accuracy of estimation of alcohol consumption in units is a concern (for a review see Del Boca & Darkes 2003). Also, retrospective recall of quantity of alcohol consumption may be affected by biases in memory, especially after episodes of heavier drinking (Monk, Heim, Qureshi, & Price 2015). However, retrospective time-line recall of alcohol consumption has shown to have good test-retest reliability ($r = .92$) in college students (as was many participants in the thesis), across various outcome variables (e.g. total number of drinks, number of binge drinking days, etc.) and time intervals (30-90 days; Sobell, Sobell, Klajner, Pavan, & Basian 1986). Nevertheless, shorter recall periods (e.g. 7 days) have been found to be more reliable and more accurate compared to longer (e.g. 30 days) periods (Hoeppner, Stout, Jackson, & Barnett 2010).

2.1.2 Alcohol use diary

In the EMA studies in the thesis (chapter four and five), participants were required to estimate the number of UK alcohol units consumed every day for seven days. To do that, they were provided with an alcohol-use diary containing a TLFB (see above) covering seven days. Participants were instructed to register each day's overall consumption at the end of the day or first thing the next morning. Although a recollection process was still involved, participant had to recall their consumption for a few hours ago rather than days or weeks ago. Paper diaries were printed out for one study (study one, chapter four). For the rest, a virtual-notepad app (ColorNote) was used to incorporate digital diaries on the smartphones. Diaries were accompanied by tables with units for the most common drinks as described above.

2.1.3 Alcohol Use Disorder Identification Test (AUDIT)

The AUDIT (Saunders, Aasland, Babor, De la Fuente, & Grant 1993) was used in all studies to assess hazardous drinking at baseline. The AUDIT consists of ten fixed-response questions regarding alcohol consumption and consequences of drinking. Scores ranged between 0 and 40, with scores between 8 and 15 indicating hazardous drinking, scores between 16 and 19 indicating harmful drinking, and scores of 20 or more indicating possible alcohol dependence. de Meneses-Gaya, Zuardi, Loureiro, and Crippa (2009) reviewed the psychometric properties of the AUDIT across several countries. They found a test-retest reliability for total scores that ranged from 0.84 to 0.95, and an average internal consistency of $\alpha = .80$. Furthermore, in a sample of UK university students, the AUDIT also exhibited high internal reliability ($\alpha = .83$; Atwell, Abraham, & Duka 2011).

2.2 The Alcohol Stroop Application (AlSAp; figure 2.2)

2.2.1 General description

AlSAp is a novel digital cognitive-experiment program that was developed exclusively for the purpose of this thesis to evaluate AB and craving in the real world. It included an alcohol Stroop-task (based on Marhe, Waters, van de Wetering, & Franken 2013), and two single-item questionnaires assessing craving and intention to drink.

AlSAp was compatible with Windows and Android devices (standard PCs, smart-phones or tablets) and was programed with OpenSesame (Mathôt, Schreij, & Theeuwes 2012), which is a Python-based software for building cognitive experiments. It was first programmed on a standard PC and then installed on smartphones. Information on the device models used in each study are provided in the relevant chapters.

AlSAp can be broadly divided in two parts. The first part was preparation of the experiment. The second part was the assessments, which included three stages: the intention to drink question, the alcohol Stroop-task, and the craving question.

2.2.2 Part one – Experiment preparation

Participants launched OpenSesame on the smartphone, selected the folder where AlSAp was saved and then selected AlSAp from a list of available experiments. After that, participants inserted their subject number and selected the folder where their output files should be saved. When all preparation steps were completed, the logo of the AlSAp appeared on screen for 2,500 milliseconds (ms) and the assessment part commenced.

2.2.3 Part two – Assessments

2.2.3.1 Stage one – Intention to drink

Intention to drink alcohol was assessed with a question reading “Do you have any plans for today that involve alcohol”. Response options were “yes” or “no”. The question remained on screen until a response was given.

2.2.3.2 Stage two – Alcohol Stroop-task

After the intention to drink item, AlSAp continued to the alcohol Stroop-task. Studies in the thesis used different versions of the task (e.g. a Stroop-task with words generally related to alcohol or with personalised pictures of alcohol). This section outlines the standard specifications of the alcohol Stroop-task and any variations are detailed in the relevant chapters.

At first, participants saw the following instructions, which stayed on screen until a response was made (phrases in brackets indicate variations of the instructions for the pictorial version of the task):

"You are about to see various words written (or various pictures presented) in different colours. Your task is to indicate the colour the word is written in (or the colour the picture is presented) as fast and as accurately as possible. Ignore the meaning of the word itself (or ignore the content of the picture) and only identify the colour. To do that, touch the appropriate colour name on the screen. To start touch ok"

Each trial of the alcohol Stroop-task commenced with a fixation dot, presented for 500ms. Following this, a stimulus was presented on the centre of the screen, which was alcohol-related (e.g. an alcohol-related word or picture) or control (e.g. a word related to the environment or a soft-drinks picture) and was presented in red, blue, or green colour (words were written in different colour fonts, while picture were recoloured applying colour filters). Below the stimulus, three buttons with the three different colour names appeared on the screen (their positions changed in each trial).

Participants were required to indicate the colour of the stimulus by selecting the corresponding button, as fast and as accurately as possible. The stimulus remained on the screen until a response was given or until a timeout period of 3,000ms. If an incorrect response was made or response time exceeded 3,000ms, a red "X" appeared on the screen for 500ms and the task moved on to the next trial. The inter-trial interval was 500ms.

The alcohol Stroop-task included eleven alcohol-related stimuli and eleven control stimuli which were matched in terms of their features (e.g. number of syllables if they were word stimuli, complexity and shapes for pictorial stimuli). Each stimulus was presented in the three different colours (red, blue or green) giving a total of 66 trials, organized into two blocks (33 trials per block) with a five seconds break between blocks. At the end of the Stroop-task, a feedback screen was displayed for 3,000 ms informing participants for their average reaction times and accuracy rate.

2.2.3.3 Stage three – Craving assessment

After the last trial of the alcohol Stroop task, an 11-point Likert scale appeared on screen (1-no urge, 11-very strong urge) with a prompt above the scale reading "Please touch a number to indicate how strong your urge to drink alcohol is right now". The question remained on screen until a response was given. The task ended with a final exit screen where participants were informed that they have completed the task and thanked for their participation.

2.3 Data preparation and analysis

2.3.1 Calculation of attentional bias scores

AB scores in the thesis represent the difference between the average response time (RT) for indicating the colour of an alcohol-related stimulus, and the average response time for indicating the colour of a control stimulus (see figure 2.3). Larger positive scores are indicative of increased AB for alcohol-related cues.

AB scores were calculated only for alcohol Stroop-task assessments where all 66 trials were completed. Before calculating AB scores, responses that were incorrect, faster than 200ms, slower than 2000ms, or three standard deviations above the average response time, were disregarded (see Christiansen, Mansfield, et al. 2015; Field, Duka, et al. 2007; Schoenmakers et al. 2008). Average RT for control stimuli were then subtracted from average RT for alcohol-related stimuli.

2.3.2 Hierarchical data clustered within multiple levels

In the EMA studies in the thesis (chapter four and five), participants were tested repeatedly over the course of the day (three times per day in chapter four and up to five times per day in chapter five) for seven days. This creates a hierarchical data-structure where units of analysis are clustered within different levels. For instance, sessions (every time a participant used the AlSAp) are clustered within each day of the week, which are further clustered within each participant (see table 2.1). Consequently, this creates a 3-level structure, where session is the first level of the analysis, day is the second, and participant is the third.

Each level contains a different sample of units of analysis. For example, if 62 participants are assessed three times per day for a week (see for example study one in chapter four), the first level (sessions) would include 1,302 sessions, the second level would include 434 days, and the third level would include 62 participants.

Each level of analysis has a unique contribution to the overall variance of the outcome variable. For example, participants might consume different quantities of alcohol due to individual differences (e.g. heavy and light drinkers). However, a given participant may consume different quantities of alcohol from one day to the other due to differences in motivation and opportunity to drink (e.g. one might consume more alcohol in the weekends, compared to weekdays). Also, a given participant may consume different quantities of alcohol in different times of a given day due to different circumstances at each time of the day (e.g. early in the morning compared to late in the evening). Finally, the magnitude of variance may differ across levels. For example, in a sample of first year university students, alcohol consumption might show less variation across individuals (i.e. similar drinking norms across first year students), although it can differ considerably across days, mainly concentrated in the weekends.

In hierarchical data-structures, measures clustered within the same level would be expected to be strongly correlated (intra-class correlation). That implies that observations in hierarchical data-structures cannot be considered independent from each other. For example, it is reasonable to expect that measures of AB taken from the same individual on the same day would be highly correlated to each other.

2.3.3 Analysing hierarchical data

2.3.3.1 Unsuitability of ordinary least square (OLS) regression models for hierarchical data

The EMA studies in the thesis explore the predictive relationship among the examined variables, for example AB, craving, and alcohol consumption. A common approach to explore these relationships would be to use ordinary least square (OLS) regression models (e.g. multiple linear regression models). OLS regression models can predict outcome variables within an error margin, which is defined by the residual terms. Residual terms express the distance between the predicted value of the outcome variable (e.g. the predicted quantity of alcohol consumption at the end of the day) and the observed value of the outcome variable (the actual quantity of alcohol consumed at the end of the day). A cardinal assumption of OLS regression analysis is that residuals are independent among observations. For example, the residual term for predicted quantity of alcohol consumption for one day should be independent from the residual term for predicted quantity of alcohol consumption for another day.

However, hierarchical data would violate this assumption due to intra-class correlations (see section 2.3.2). For instance, it is reasonable to expect that the residual terms for quantity of alcohol consumption on different days will not be independent if all measures are derived from the same individual. Ignoring the hierarchy of the data and fitting an OLS regression model leads to an underestimation of the standard errors of the regression parameters which consequently increases the risk of type I errors (Goldstein 2011).

2.3.3.2 Multilevel regression models

Multilevel (ML) regression models would be more suitable for analysing hierarchical data, as opposed to OLS regression models (Hayes 2006; Quené & van den Bergh 2004). ML regression models allow for the correlation in the residuals (see section 2.3.3.1) to be factored in the model. They also allow for unequal numbers of data points (e.g. a participant missing some AB assessments on some of the days), and for the examination of how the variance in the outcome variable is partitioned among the different levels of the model.

In ML regression models, the regression lines describing the association between the predictor and the outcome vary among units of analysis in higher levels of analysis. For example, in a "days within participants" 2-level ML regression model predicting alcohol consumption by AB, each participant will have a different regression line describing the association between AB and alcohol consumption. These lines may vary in terms of their intercept coefficient (random-intercept model), their slope-coefficient (random-slope model), or both. A random-intercept model would represent a different starting point of the association between AB and alcohol consumption for each participant (i.e. in the absence of any AB each participant would be predicted to consume a different amount of alcohol), while a random-slopes model would represent an association of different strength or direction for each participant (e.g. AB would be predicted to have a positive association with alcohol consumption for some participants but a negative association for some others).

The multilevel regression models presented in the thesis are random-intercept models. However, the nature of the examined associations (e.g. positive or negative) was not expected to change between participants or days and, therefore, no random-slope models were used.

2.3.3.3 Fitting a multilevel regression model

A prerequisite of ML modelling is that the outcome variable shows significant variability in the highest level (Peugh 2010). Otherwise, data can be considered to be non-hierarchical and they can be analysed with a single level OLS regression.

In studies with hierarchical data in the thesis (chapter four and five), an ML model with the highest possible number of levels was fitted first. Using the likelihood ratio statistic, this model was then compared to models with a smaller number of levels until no significant difference was detected in how well the model fitted the data (Rasbash, Steele, Browne, & Goldstein 2012). For example, in chapters four and five, alcohol consumption was measured once per day, so data for alcohol consumption could have a 2-level hierarchy, with days clustered within participants. A 2-level model was fitted first and it was then compared to a single-level model. Significant differences between the two models would indicate significant variability of alcohol consumption between participants, and the 2-level model would be retained as best fitting the data. Non-significant difference between the two models would indicate that alcohol consumption did not significantly vary between participants. In that case, the data would be considered as having a single-level, non-hierarchical structure, and the analysis would be performed with an OLS regression using days as the unit of analysis.

2.3.4 Ridge Regression

In cases when data did not meet assumptions for ML analysis, single-level regression models were used. However, standard linear or logistic regressions are based on the assumption that predictors entered in the model are not highly correlated with each other (lack of multicollinearity). Violation of the low multicollinearity assumption can lead to inflation of standard errors of the unstandardized slope coefficients of the model (*b* values; Field 2013), which in turn leads to reduced *t*-statistics and increased risk of type II error.

In several occasions in this thesis, the predictors included in the model represented measures of the same variable taken on different times during the day. For example, studies in chapter four examined whether alcohol consumption could be predicted by measures of AB in the morning, the early afternoon, and the late afternoon. On these occasions, it is reasonable to expect high multicollinearity among the predictors, as they represent the same variable being measured repeatedly within the same day.

When high multicollinearity is expected, ridge regression is a better alternative to traditional multiple or logistic regression. Ridge regression is a penalised regression model which corrects for the effects of multicollinearity by shrinking the regression coefficients and their SE's by a shrinkage parameter that is a function of the collinearity (Cule & De Iorio 2013).

Table 2.1. Example of hierarchical data structure

LEVEL 3	LEVEL 2	LEVEL 1	LEVEL 3 VARIABLE	LEVEL 2 VARIABLE	LEVEL 1 VARIABLES	
Participant	Day	Assessment Point	Gender	Alcohol consumption	AB	Craving
1	1	1	*female*	*a*	*a*	*a*
		2			*b*	*b*
		3			*c*	*c*
	2	1		*b*	*d*	*d*
		2			*e*	*e*
		3			*f*	*f*
	3	1		*c*	*g*	*g*
		2			*h*	*h*
		3			*i*	*i*
	4	1		*d*	*j*	*j*
		2			*k*	*k*
		3			*l*	*l*
	5	1		*e*	*m*	*m*
		2			*n*	*n*
		3			*o*	*o*
	6	1		*f*	*p*	*p*
		2			*q*	*q*
		3			*r*	*r*
	7	1		*g*	*s*	*s*
		2			*t*	*t*
		3			*u*	*u*

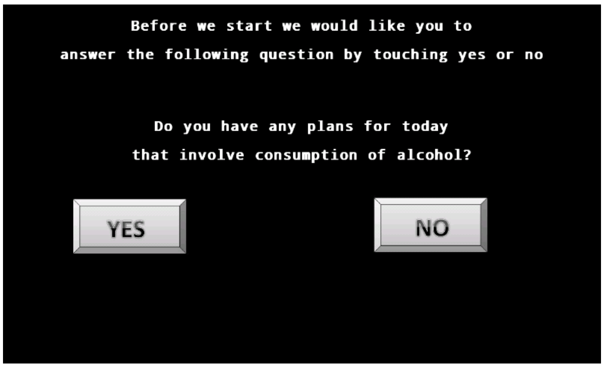
Example of a participant tested three times per day for seven days. In this example, participant 1 was *a female who* on day 4 consumed a *d* amount of alcohol and on the first assessment of the 4th day they had a score of *j* for AB and *j* for craving.

Figure 2.1. Parts and stages of AISAp

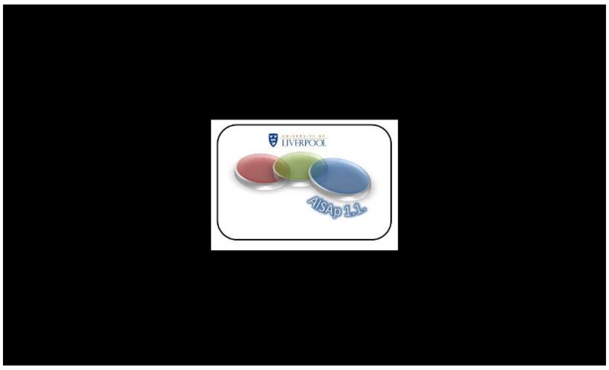
Part one – Experiment preparation

Figure 2.1: Parts and stages of AlSAp – continued

Part two/Stage one – Intention to drink



Part two/Stage two – Alcohol Stroop-task



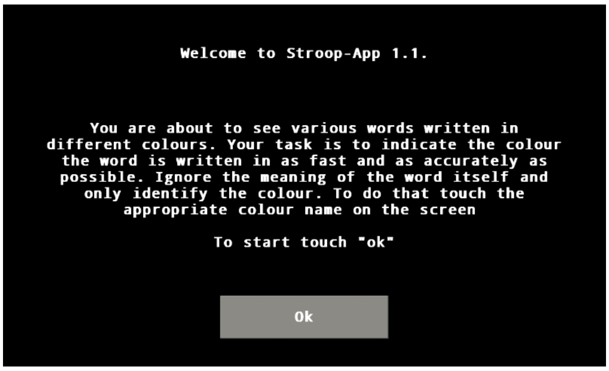
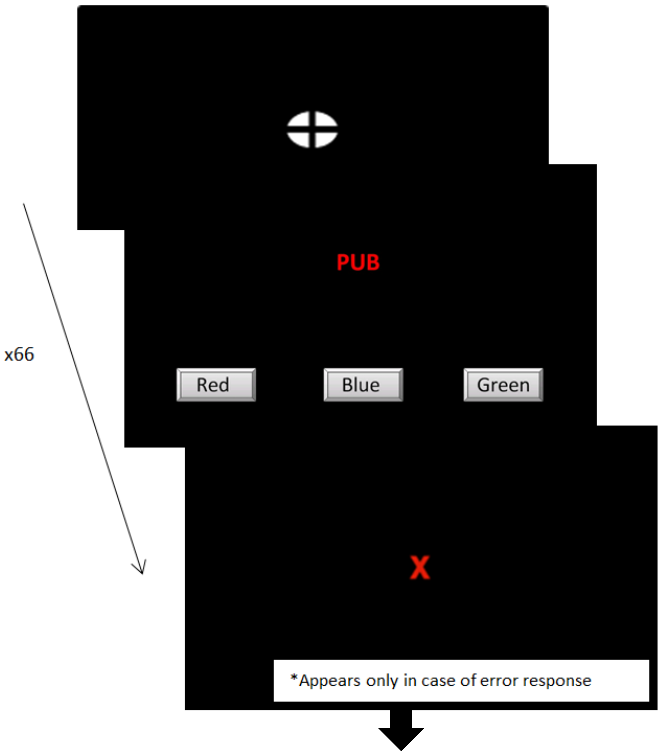


Figure 2.1: Parts and stages of AlSAp – continued



Part two/Stage three – Craving assessment

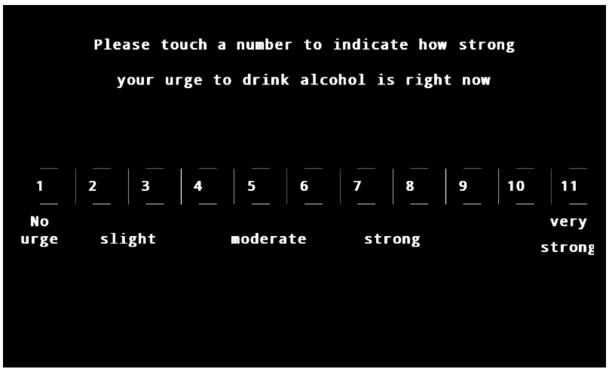


Figure 2.2. Calculation formula for AB scores

$$AB = \text{Mean RT}_{\text{alcohol}} - \text{Mean RT}_{\text{control}}$$

AB = Attentional Bias; RT = reaction time

Chapter three: Experiments one and two.

3. Go Natural: Measurements of AB with the alcohol Stroop-task are more reliable in the real world than in laboratories.

A briefer version of experiment one in this chapter has been published as Spanakis, P., Jones, A., Field, M., and Christiansen, P. (2018). A Stroop in the hand is worth two on the laptop: Superior reliability of a smartphone based alcohol Stroop in the real world. *Substance use and Misuse*, doi.org/10.1080/10826084.2018.1536716. See Appendix 12.

3.1 Abstract

The alcohol Stroop-task is widely used in the laboratory on computers for measuring attentional bias (AB) for alcohol cues, but lately there has been an increased use of portable electronic devices in the real world (e.g. ecological momentary assessments - EMA). In general, the substance Stroop-task has been shown to have poor reliability and inconsistent predictive validity (association with substance-use behaviour) in the laboratory, while very little is known for the task's psychometric properties in the real world. This chapter aims to assess conditions under which the alcohol-Stroop task has improved internal reliability and predictive validity. Study one examined two suggested solutions for the improvement of the psychometric properties of the task; the use of pictorial and personalised stimuli. A basic alcohol Stroop-task including general alcohol-related words (e.g. liquer) was compared with an upgraded version including personalized pictures of alcohol. Tasks were delivered in a neutral room on a laptop computer (traditional administration modality) or in participants' homes on a smartphone (EMA modality). Acceptable internal reliability was found for both types of the alcohol Stroop-task when administered on smartphones in participants' homes, but not when administered on laptops in university neutral rooms. However, no measure of AB predicted alcohol use. Study two investigated whether the finding that alcohol Stroop-task reliability increased to acceptable levels if assessment took place on a smartphone in the real world could be attributed to the ubiquitous nature of smartphones or the natural environment in which the testing took place. Combining data from larger studies in this thesis (study one in chapter four, and chapter five) the study examined the psychometric properties of an alcohol Stroop-task containing alcohol-related words administered on a smartphone in a neutral laboratory, a semi-natural environment (i.e. the Bar-lab), or in the real world. The task showed acceptable internal reliability, only when administered in the real world. Predictive validity remained poor. The two studies together demonstrate that participants' responses to the alcohol Stroop-task are more consistent in the real world and this is not an artefact of the ubiquitous nature of smartphones. Despite this improved internal reliability caution should still be exerted making inferences regarding predictive validity.

3.2 General Introduction

Incentive-motivational models of addiction (e.g. Franken 2003) suggest that a central factor in the aetiology of substance use is the ability of substance-related environmental cues to grab and hold attention (known as attentional bias; AB). Several paradigms have been developed to measure AB for substance-related stimuli, with the substance Stroop-task being one of the most commonly used (for a review see Cox et al. 2006). In the substance Stroop-task, participants are presented with substance-related and control stimuli in different colours and are required to name the colour the stimulus is presented in. It is argued that slower colour-naming of substance-related stimuli, compared to control, is indicative of AB, as the substance-related stimuli are grabbing and/or holding attention.

A large number of studies have revealed a positive association between AB, measured with the substance Stroop-task, and substance-use behaviour (for a comprehensive review, see Field & Cox 2008). However, numerous studies have also failed to show this (e.g. Ryan 2002; Sharma et al. 2001; Snelleman et al. 2015) and the clinical utility of these measures has also been argued to be limited (for a review see Christiansen, Schoenmakers, et al. 2015). One explanation for these inconsistencies might be the poor psychometric properties of the Stroop-task (see chapter one, section 1.6.2). For example, Ataya, Adams, Mulings, Cooper, Attwood and Munafo (2012) analysed the internal reliability of substance Stroop-tasks used in nine studies from their laboratory and found that only two had acceptable internal reliability (Cronbach's $\alpha > .70$).

It is notable that recent years have seen increased uptake of ecological momentary assessment (EMA) methods to assess AB in the real world using smartphones and similar devices (see chapter one, section 1.12). Field et al. (2016) argued that repeatedly measuring AB in the real world is critical because AB is likely to be a marker of the current motivational state to drink alcohol, which would be determined by a range of factors (e.g. alcohol availability, mood, and craving) rather than being a stable trait-like measure. Notably, measures of AB taken with the substance Stroop-task in EMA studies have shown some predictive utility (e.g. Marhe et al. 2013), and smartphone-based AB modification applications to help users cut down on drinking are already commercially available (Cox, Intrilligator, et al. 2015). However, the psychometric properties of smartphone-based substance Stroop-tasks outside the laboratory have yet to be investigated.

The two studies presented in this chapter examine the psychometric properties (i.e. internal reliability and predictive validity) of the alcohol Stroop-task when administered in laboratories and naturalistic environments, on laptop-computers and on smartphones. The overarching purpose is to examine whether EMA modality of administration (testing in the natural world on portable electronic devices) would show superior psychometric properties compared to more traditional modalities (laboratory measures).

3.3 Study one

3.3.1 Introduction

Psychometric properties of the substance Stroop-task have been found to be poor (Ataya et al. 2012). One solution to improve them would be to use personalized stimuli (see Field & Christiansen 2012). The term "personalized" refers to the use of stimuli that reflect participants' drinking preferences. As AB is argued to develop through repeated pairings between environmental cues and substance use (Franken 2003), it would be unlikely for an individual to show AB for a type of alcohol they rarely consume or find aversive (e.g. people who only consume beer should not show AB towards whiskey-related stimuli).

It would be reasonable to expect that responses to tailored stimuli would be more consistent and more closely related to the individual's alcohol use. Indeed, when measuring AB using a visual probe task, Christiansen, Mansfield, et al. (2015) found that personalised stimuli produced acceptable levels of internal reliability, compared to generalised alcohol stimuli (although predictive validity did not improve). In another study (Christiansen & Bloor 2014), a personalised alcohol Stroop-task (but not a Stroop-task using stimuli generally related to alcohol) predicted drinking behaviour in undergraduate drinkers, thus showing better predictive validity.

Another potential improvement would be the use of pictorial stimuli instead of words. Ataya et al. (2012) meta-analysed the internal reliability of the substance Stroop-task in six studies from their lab and found that reliability was higher when using pictorial stimuli compared to words. However, since the comparison was based on only one study using pictorial stimuli, further research is necessary. In real life, alcohol-cues are usually perceived as images (e.g. seeing an alcohol advertisement, a wine glass, etc.), rather than verbally (reading an alcohol-related word).

As a result, participants might find pictorial representations of alcohol more realistic and more salient. It might be reasonable to expect that participants will react more consistently to a stimulus they perceive as salient (pictorial stimuli), compared to a less vivid stimulus (word stimuli). A study exploring attentional bias in anxiety (Kindt & Brosschot 1997) showed that pictures might be perceived as more salient than words, as they found that arachnophobic participants rated pictures of spiders as more aversive than spider-related words (although there was no difference in AB for pictures and words). Although there is also a study which found that cocaine users showed significant AB for both cocaine-related words and pictures, they did not compare between magnitudes of AB produced by words and pictures (Hester et al. 2006).

The current study compared a basic Stroop-task (including general alcohol word) and an upgraded Stroop-task (including personalised pictures of alcohol) to investigate whether using personalized pictorial stimuli would increase the reliability and validity of the alcohol Stroop-task. Given the theoretical reasons for using EMA, as well as arguments that EMA methods would be more robust measures of AB (see section 1.11 and 1.12 in chapter one), reliability of the tasks was also compared on a between subjects basis, when completed in a neutral university room on a computer (traditional modality of administration) or in participants' own homes on a smartphone (EMA modality of administration).

It was hypothesized that the upgraded Stroop-task would show better reliability and predictive validity compared to the basic Stroop-task in both administration modalities. It was also predicted that participants would demonstrate AB for both alcohol words and personalized pictures in both administration modalities.

3.3.2 Methods

3.3.2.1 Participants

One hundred and twenty participants (61 female, average age 23.10 (±8.42) years) were recruited from the University of Liverpool and local community via advertisements and word of mouth. Inclusion criteria were; fluent English speaking, regular social drinker (consuming at least one alcoholic drink in a typical week), and beer being their preferred alcoholic drink. Participants were excluded if they had any current or previous alcohol abuse or dependence diagnosis, were pregnant or breastfeeding, or colour-blind. Participant characteristics are shown in table 3.1. All participants provided informed consent prior to participation and the study was approved by the University of Liverpool Ethics committee.

3.3.2.2 Questionnaires

To screen participants for drink preference, they were asked to list their four most preferred alcoholic drinks (either brands or types of alcohol), from the most preferred to the least preferred. Participants were included in the study only if they listed beer on the first place.

The Time Line Follow Back (TLFB; Sobell & Sobell 1990) and the Alcohol Use Disorders Identification Test (AUDIT; Saunders et al. 1993) were used to assess typical alcohol consumption and hazardous drinking habits, respectively. For a detailed description see chapter 2, sections 2.1.1 and 2.1.3.

3.3.2.3 Stimuli

3.3.2.3.1 General words.

Eleven general alcohol-related words (e.g. pub, liquer, cocktail, etc.) and 11 control words (e.g. bog, ravine, bridges, etc.) were presented in three different font-colours (blue, green, red). Words were matched in terms of word length and word frequency in the English language (see Sharma et al. 2001). Stimuli were organized in two blocks with each block containing either alcohol or control stimuli.

3.3.2.3.2 Personalized pictures.

Eleven beer pictures (bottles or cans of beer) and 11 control pictures (bottles or cans of soft drinks) were recoloured in three different colours (blue, green, red). As preference for beer was an inclusion criterion, beer pictures represent a personalised stimulus. Pictures of known beer brands (e.g. Heineken) and soft-drink brands (e.g. Seven-Up) were used as they are easily and widely recognizable as an alcoholic drink and as a soft-drink respectively, even after recolouring. Pictures selected for this study were matched in terms of presentation features (e.g. size and packaging – pictures were not matched for colour as they were then re-coloured). Stimuli were organized in two blocks with each block containing either alcohol or control stimuli.

It should be noted that it was necessary for a beer pictorial Stroop to use soft-drinks pictures as control stimuli, rather than environmental pictures. Environmental pictures cannot be matched to the beer pictures on perceptual qualities. Neutral, non-appetitive words were used in the word Stroop task to be consistent with previous research (particularly those using EMA; see Marhe et al. 2013), and also due to practical difficulties matching soft-drinks words (avoiding brand names) to alcohol words.

3.3.2.4 Alcohol Stroop task

See alcohol Stroop-task description in chapter 2, section 2.2.3.2

3.3.2.5 Procedure

After providing informed consent, participants provided demographic information and they self-completed the drink-preference questionnaire, the TLFB, and the AUDIT. They then completed two versions of the alcohol Stroop-task, a basic (with words generally related to alcohol) and an upgraded (with beer pictures), in a counterbalanced order. In both tasks, blocks were presented in a fixed order, with alcohol blocks presented first. Participants were recruited in two waves; the first wave (n=60) completed the Stroop-tasks on a standard laptop-computer in a neutral room in the university campus (i.e. empty class rooms, library rooms, etc.) and the second (n = 60) were loaned a smartphone (Samsung Galaxy Ace GT-S5830i) to complete the task in their homes (in parts of the house where alcohol is usually consumed, e.g. living rooms). At the end of the study, participants were debriefed.

3.3.2.6 Data reduction and analysis

AB scores were calculated by subtracting reaction times (RTs) for control stimuli from RTs for alcohol stimuli. For a detailed description of AB scores calculation see chapter 2, section 2.3.1. Positive AB scores (slower RTs for alcohol compared to control stimuli) are indicative of automatic selective attention towards alcohol-related stimuli.

Internal reliability of the tasks was calculated with Cronbach's α, where α ≥ .70 was considered acceptable (Kline 1999). For each pair of stimuli (alcohol stimulus and its matched control) a separate AB score was calculated. First, the average RT for each individual alcohol and control stimulus was calculated across the three coloured presentation (e.g., average RT for the word pub in red blue and green, and average RT for the word bog in red, blue, and green). Then, for each pair of stimuli, we subtracted the average RT for the control stimulus from average RT for the alcohol stimulus. Given that each Stroop task included 11 pairs of alcohol and control stimuli, Cronbach's α reflected the internal consistency among those 11 stimuli-specific AB scores.

For predictive validity, an alcohol use involvement index was calculated as the sum of the standardised TLFB and AUDIT scores (see Christiansen & Bloor 2014). Combining these correlated measures (TLFB and AUDIT, r_s = .34) into a single variable allows for the performance of a single analysis, which reduces chances of false positive results. A two-steps hierarchical regression was run, with age and gender being entered at the first step, and AB scores at the second. To examine the predictive validity of AB scores within each condition (neutral room-computer and home-smartphone) the regression was run for each condition separately. Although examining the interaction between condition and AB scores before splitting the sample could be argued to be a more appropriate approach, multicollinearity problems in the interaction term (VIFs > 10) did not allow for this. Due to this, p values were corrected for multiple comparisons by multiplying them by two.

Presence of AB was examined with a 2x2x2 mixed ANOVA on reaction times, with Stroop type (basic, upgraded) and stimulus type (alcohol, control) as within-subjects factors, and condition (neutral room-computer, home-smartphone) as the between-subjects factor. Differences in magnitude of AB were examined with a 2x2 mixed ANOVA on AB scores, with Stroop type (basic, upgraded) as within-subjects factor and condition (neutral room-computer, home-smartphone) as between. Due to Box's test suggesting covariance matrices of AB scores being heterogeneous, the main effect of Stroop type and the interaction between Stroop type and condition where analysed using Pillai's Trace multivariate test.

3.3.3 Results

3.3.3.1 Participants' characteristics

Conditions (neutral-room computer and home-smartphone) were compared in terms of gender ratio, age, TLFB and AUDIT scores. Regarding gender, there was no difference between conditions *($\chi^2(1) = .300$, $p = .584$, Cramer's V = .05).*

Due to age being skewed even after transformation, a non-parametric test was used (Mann-Whitney U). There was no difference between conditions in terms of age *(U (120) = 2003, $p = .268$).*

A one-way MANOVA was used to compare conditions in terms of AUDIT and TLFB scores. Due to TLFB scores being skewed, they were square-root transformed before entering into the analysis. There was no effect of condition on the dependent variables *($F(2,117) = .01$, $p = .99$, $\eta_p^2 < .001$).*

3.3.3.2 Internal reliability

An acceptable level of reliability was achieved for home-smartphone Stroop-tasks both in the basic *(alcohol words, α=.70)* and upgraded type *(beer pictures, α=.74).* The neutral room-computer Stroop task did not have acceptable reliability, neither in the basic *(α=.49)* or upgraded type *(α=.58)*.

3.3.3.3 Predictive validity: Regression between AB, condition, and alcohol involvement (Table 3.2).

Regardless of condition (neutral room-computer or home-smartphone), neither AB for words (basic type) nor AB for beer pictures (upgraded type) predicted individual differences in alcohol involvement *(ps>.05)*.

3.3.3.4 Presence of AB

A mixed design ANOVA revealed a main effect of Stroop type *(basic or upgraded; $F (1, 118) = 4.61$, $p =.034$, $\eta_p^2 = .04$)* on RTs, with participants responding slower to pictures compared to words. This demonstrated that participants showed a general AB for pictures, compared to words.

There was also a main effect of stimulus type *(alcohol or control; $F (1, 118) = 23.22$, $p <.001$, $\eta_p^2 = .16$)* with slower responses to alcohol stimuli compared to control. This reveals a significant AB for alcohol stimuli, compared to control.

Finally, there was also a main effect of condition (neutral room-computer or home-smartphone; *$F (1, 118) = 15.34$, $p <.001$, $\eta_p^2 = .12$)* with slower responses in the neutral room-computer condition, compared to the home-smartphone. This shows that participants had generally slower reactions when completing the task on a laptop in a neutral room, compared to when completing the task on a smartphone in their homes.

A significant three-way interaction was found between stimulus type (alcohol or control), Stroop type (basic or upgraded) and condition (neutral room-computer or home-smartphone*; $F (1, 118) = 5.43$, $p = .021$, $\eta_p^2 = .04$)*. Pairwise comparisons with t-tests revealed the following:

Within the neutral room-computer condition, participants responded slower to alcohol words compared to control in the basic Stroop *($t(59) = 7.97$, $p <.001$, $d=.52$)*, but there was no difference between beer and control pictures in the upgraded Stroop *($t(59) = -0.39$, $p = .700$, $d = - .03$)*. This demonstrates that participants in the neutral room-computer condition showed AB for alcohol-related words but not for beer pictures.

The same pattern was found within the home-smartphone condition. Participants responded slower to alcohol words compared to control in the basic Stroop *($t(59) = 2.93$, p*

=.005, d=.22) , but showed no difference between beer and control pictures in the upgraded Stroop *(t(59) = -0.40, p = .69, d = -.02).* This reveals that participants in the home-smartphone condition showed AB for alcohol-related words but not for beer picture.

3.3.3.5 Magnitude of AB

A mixed design ANOVA revealed a main effect of Stroop type (basic or upgraded) on AB scores *(F(1, 118) = 24.14, p < .001,* η_p^2 *= .17)*, with magnitude of AB being larger in the basic Stroop. There was no main effect of condition (neutral room-computer or home-smartphone; *F(1, 118) = 3.29, p = .072,* η_p^2 *= .03*) or an interaction between condition and Stroop type *(F(1, 118) = 2.54, p = .114,* η_p^2 *= .02)*. This shows that AB for alcohol-related words was stronger than AB for beer pictures, regardless of condition.

3.3.3.6 Post-hoc analysis: Comparing reaction times within alcohol and control stimuli.

The main analysis revealed that participants in both conditions showed AB for alcohol-related words but not for beer pictures. Contrasting the alcohol-related worlds with a set of neutral stimuli (i.e. environmental words) but the beer pictures with a set of appetitive stimuli (i.e. pictures of soft-drinks pics), might explain the absence of AB in the latter case. Due to their appetitive properties, soft-drinks might also automatically attract participants' attention and, therefore, they might mask any AB for alcohol-related cues.

To test this assumption, reaction times were compared between the alcohol stimuli used in the study (alcohol words and beer pictures), and between the control stimuli (environmental words and soft-drinks pictures). Participants were significantly slower when reacting to soft-drinks pictures compared to environmental words *(t(119) = -5.36, p < .001, d = -.30)*. However, no difference was found between alcohol words and beer pictures *(t(119) = 1.28, p = .202, d = .08).* These results suggest that participants showed AB for pictures of soft-drinks, compared to words related to environmental features. This is probably due to the content of the stimuli (appetitive vs neutral) rather than the format (pictorial vs word), as there were no differences in AB when a pictorial appetitive stimulus (pictures of beer) was contrasted to an appetitive word (words related to alcohol).

Table 3.1. Participants characteristics - Study one

	Neutral room-computer condition (n = 60)	Home-smartphone condition (n = 60)	Sample (N = 120)
Gender (F:M)	32:28	29:31	61:59
Age	22.28 (±7.28)	23.92 (±9.42)	23.10 (±8.42)
AUDIT[a]	15.27 (±5.49)	15.33(±6.77)	15.30 (±6.14)
Weekly consumption[b] (units)	24.41 (±18.83)	24.82 (±19.69)	24.61 (±19.19)
Reaction time (RT) for alcohol worlds (ms[c])	1325.56 (±188.42)	1167.26 (±174.98)	1246.41 (±197.74)
RT for control words (ms)	1230.48 (±174.13)	1127.83 (±180.44)	1179.16 (±183.94)
RT for beer pictures (ms)	1282.81 (±164.67)	1178.83 (±179.21)	1230.82 (±179.14)
RT for soft-drinks pictures (ms)	1287.96 (±176.55)	1183.38 (±190.70)	1235.67 (±190.37)

Values are means ±SD, apart from gender for which they represent frequencies. a. AUDIT = Alcohol use disorder identification test. Scores range between 0 (minimum) and 40 (maximum). A score of 8 or above indicates at least hazardous drinking. b. one UK unit = 8g of alcohol. c. ms = milliseconds

Table 3.2. Regression showing the association between AB and alcohol involvement in the neutral-room computer condition, adjusting for age and gender - Study one

Steps	Model fit		Predictors	Individual coefficients
	F	R^2		β
Step 1	1.04	.04	Gender	-.07
			Age	-.17
Step 2	.91	.06	Gender	-.08
			Age	-.18
			AB basic Stroop (ms[a])	-.10
			AB upgraded Stroop (ms)	.12

a. ms = milliseconds; *p<.05; Table reveals that AB was not associated with alcohol use involvement, regardless of whether it was assessed with alcohol words or beer pictures.

Table 3.3. Regression showing the association between AB and alcohol involvement in the home-smartphone condition, adjusting for age and gender - Study one

Steps	Model fit		Predictors	Individual coefficients
	F	R^2		β
Step 1	.08	.05	Gender	-.01
			Age	-.05
Step 2	.10	.09	Gender	-.03
			Age	-.01
			AB basic Stroop (ms)	-.07
			AB upgraded Stroop (ms)	.04

a. ms = milliseconds; *p<.05; Table reveals that AB was not associated with alcohol use involvement, regardless of whether it was assessed with alcohol words or beer pictures.

Figure 3.1. Mean reaction times in the alcohol Stroop-task, split per condition and type of stimuli - Study one

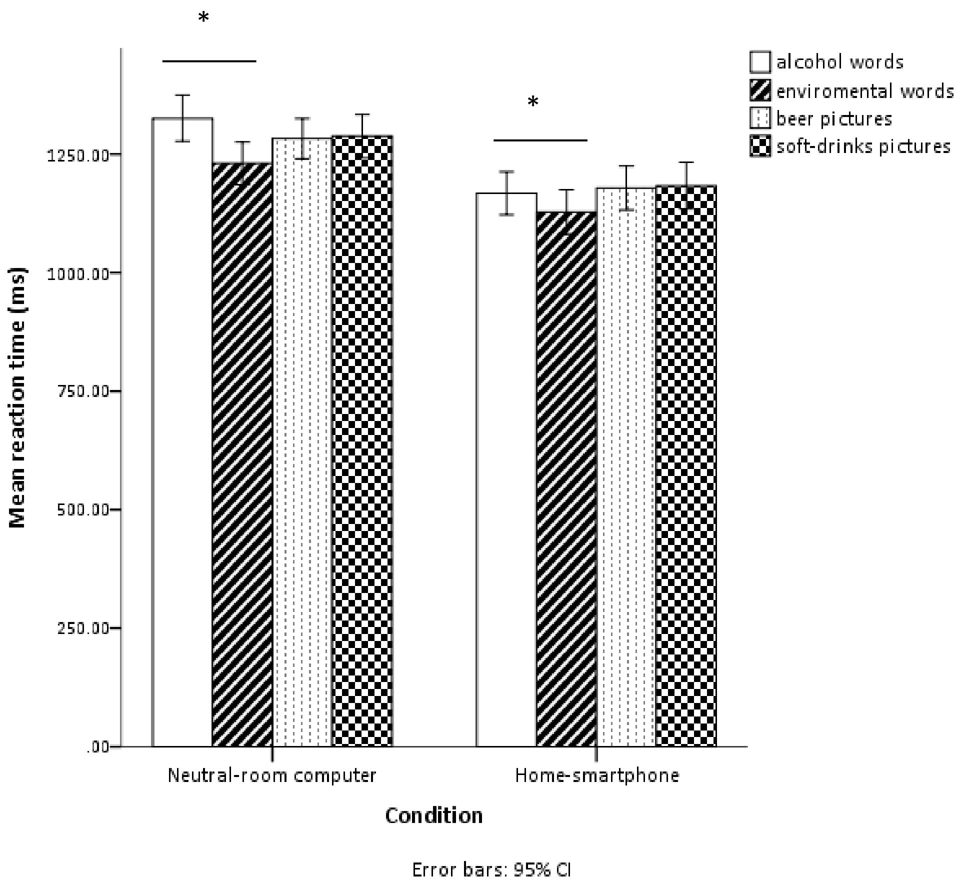
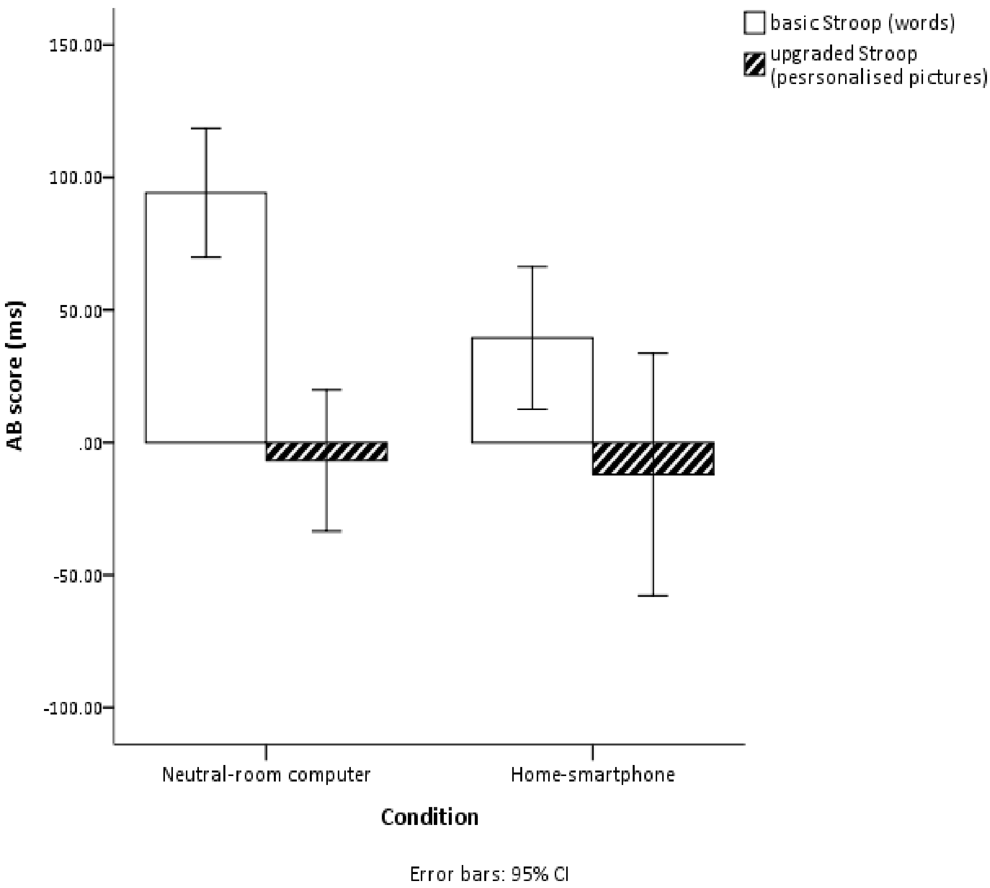


Figure 3.2. AB scores, split per condition and type of alcohol Stroop-task - Study one



3.3.4 Discussion

The current study contrasted the psychometric properties of a basic (general alcohol words) and an upgraded (personalized pictures) Stroop task, administered on a standard computer in a neutral room (traditional administration modality) and on a smartphone in participants' homes (EMA administration modality). The hypothesis that the upgraded Stroop-task would show acceptable reliability in both administration modalities was partially accepted. The task showed acceptable reliability in the EMA modality, but not the traditional administration modality. However, the hypothesis that the upgraded Stroop-task would show predictive validity was rejected, as neither of the measures were associated with alcohol use. Notably, the same pattern was found also in the basic Stroop. Reliability was acceptable in the EMA administration modality but not the traditional, although predictive validity was poor. Finally, the hypothesis that participants will show AB for alcohol-related cues was partially accepted. Participants showed AB for alcohol words but not for beer pictures, regardless of modality of administration.

Utilizing pictorial and personalized stimuli instead of general alcohol-related words did not robustly lead to more consistent responding in the Stroop-task. Responding was consistent only when the task was completed on smartphones in participants' homes, instead of in the laboratory on laptop computers, regardless of the type of stimuli used. This finding suggests that the modality of administration might be more important than the format of the stimuli in improving the reliability of the measures. However, the EMA modality of administration has two components; the means of administration (smartphones) and the environment where testing took place (real world). Results from this study do not show whether the acceptable reliability is attributable to the ubiquitous nature of smartphones, the ecologically valid environment of homes, or a combination of both. This question will be addressed in study two.

A surprising result was that participants showed AB for alcohol words but not for beer pictures. There are three possible explanations for this finding. First, the lack of AB for beer pictures might be explained by the use of appetitive control stimuli (pictures of soft-drinks). As shown in the post-hoc analysis, participants attention was more strongly attracted by soft-drinks pictures, compared to environmental words, revealing that pictures of soft-drinks are not as “neutral” as environmental words, probably due to their appetitive nature. The analysis also showed that this finding cannot be solely explained by pictures simply attracting more attention than words, as alcohol pictures did not attract attention more than alcohol words. In the Stroop-task which included pictures of beer and soft-drinks, the attention-grabbing properties of soft-drinks may have reduced the magnitude of difference in reaction times between beer-related stimuli and soft-drinks stimuli, masking any AB for beer stimuli. This is in line with Christiansen and Bloor (2014), who also failed to detect any AB for alcohol stimuli, compared with soft-drinks stimuli. However, despite the loss of magnitude in Stroop interference, it could be argued that appetitive control stimuli (e.g. soft-drinks) offer some methodological advantages too. When comparing alcohol with non-alcoholic appetitive stimuli, the appetitive content is kept constant, ensuring that any observed AB for alcohol can be attributed to the alcohol-specific connotations of the alcohol stimuli, rather than their appetitive appeal in general.

Second, stimuli with strong pre-existing associations with a colours-name in the response buttons (in this case blue, green, or red) could cause greater interference when completing the Stroop-task and, therefore, inflate reaction times (Cox et al. 2006). As several popular brands were used in the Stroop-task in this study, it might be possible that participants held pre-existing colour association for some of them (e.g. Heineken might bring green in mind, and Dr.Pepper might bring red in mind). Although this could have been true for both beer stimuli and soft-drinks stimuli alike, stimuli were not matched in terms of pre-existing colour associations. Participants in this study might have held pre-existing colour association for more of the soft-drinks, rather than the beers, which might have masked reaction time differences between the two sets of stimuli.

Third, colour filters applied on the pictures may mean that the stimuli were less like those previously experienced by participants in the real world, thus reducing their incentive salience. Future research could explore using either coloured outlines around pictures (see for example cocaine-related Stroop task by Hester et al. 2006), in order to explore optimum formats of Stroop-task stimuli.

This study had a number of limitations. First, for practical reasons, participants were allocated to conditions in waves of 60, rather than randomly. However, the same recruitment procedures were followed in both waves and participants were tested during the same period of the year. Also, results showed that conditions did not differ in terms of age, gender ratio, hazardous drinking habits, or typical alcohol consumption.

In addition, the study design was not fully counterbalanced. Although the order of the tasks (basic or upgraded alcohol Stroop-task) was counterbalanced across participants, the order of blocks within each task (alcohol or control) was fixed, with alcohol stimuli always presented first.

Furthermore, as discussed earlier, specific features of the upgraded Stroop-task (i.e. appetitive control stimuli, pre-existing colour associations, and complete re-colouring of stimuli with colour filters) might have played a role in the absence of AB for beer pictures. Future studies could explore different variations of sets of pictorial stimuli, as for example stimuli that are matched in terms of pre-existing colour associations, neutral control stimuli (for example items of stationary; Di Lemma, Dickson, Jedras, Roefs, & Field 2015), or stimuli with coloured outlines instead of superimposed colour-filters (see for example cocaine-related Stroop task by Hester et al. 2006), to identify optimal formats of stimuli for the pictorial alcohol Stroop-task.

Regarding personalized stimuli, the current study only recruited beer drinkers due to limitations in developing multiple versions of the app. Although this methodology still offers meaningful comparisons, future studies should consider tailoring personalized stimuli to each participant (e.g. Christiansen & Bloor 2014), thus offering a broader range of personalized stimuli.

Finally, the two suggested solutions to improve psychometric properties (personalization of stimuli and pictorial presentation) were combined within a single alcohol-Stroop task (the upgraded Stroop-task). Although there were no reliability or validity differences between the basic and upgraded Stroop, future studies could disentangle pictorial and personalized features to examine the effects of each on the psychometric properties of the alcohol-Stroop task.

3.4 Study two

3.4.1 Introduction

Study two builds on a number of findings from study one. First, study one revealed that the alcohol Stroop-task has acceptable reliability when administered in the EMA modality. However, in that modality, administration on a smartphone was combined with administration in the real world, making difficult to determine whether the acceptable reliability can be attributed to the ubiquitous nature of smartphones or the increased ecological validity of the home environment. On one hand, individuals become increasingly habituated to smartphone use (for a related discussion see Miller, 2012; Oulasvirta, Rattenbury, Ma, & Raita, 2012; Raento, Oulasvirta, & Eagle, 2009). This might make responses provided on a smartphone-based cognitive task more consistent. On the other hand, factors related to the underlying motivational state of AB (e.g. presence of alcohol cues, substance availability; see Field et al. 2016) are more prominent in participants' homes, and therefore alcohol-related cues might attract attention more systematically in those environments.

Second, the previous study showed that the upgraded Stroop-task was not necessary to achieve acceptable reliability in the EMA administration modality. A basic Stroop-task with non-personalized alcohol words also showed acceptable internal reliability. Also, the attention-grabbing properties of soft-drinks may have reduced the magnitude of difference in reaction times between beer-related stimuli and soft-drinks stimuli, masking any AB for beer stimuli

Building on the findings described above, Study two examined the internal reliability and predictive validity of a smartphone-based alcohol Stroop-task when administered in a neutral laboratory (i.e. neutral room containing desks and computers), a semi-naturalistic laboratory (i.e. the University of Liverpool's Bar-Lab) and in participants' natural environments (i.e. real-world). Only one type of alcohol-Stroop task was utilized in Study two; a Stroop-task containing general alcohol-related words (e.g. pub, liquer). Acceptable reliability in all environments would indicate that the ubiquitous nature of smartphones leads to more consistent responding, regardless of ecological validity of the setting. Alternatively, acceptable reliability only in the naturalistic environments (real-world and/or Bar-Lab) would indicate that ecological validity is more important for more consistent responding. It was expected that the smartphone-based alcohol-Stroop task would show acceptable reliability in the real world, replicating study one. However, since this is a novel topic without much empirical evidence so far, specific hypotheses as to whether reliability will also be acceptable in the other environments were not formulated.

3.4.2 Methods

3.4.2.1 Participants

A sub-sample of eighty-two participants was randomly drawn from two bigger EMA studies in this thesis (study one in chapter four and the study in chapter five), forty-one participants from each study. An additional forty-one participants were recruited from the University of Liverpool and local community via advertisements and word of mouth. Overall, one hundred and twenty-three participants (97 female, average age 26.83 (±8.99) years) were included in this study. Participants were fluent English speakers and regular social drinkers (consuming at least one alcoholic drink in a typical week). Participants were excluded if they had any current or previous alcohol abuse or dependence diagnosis, were pregnant or breastfeeding, or colour-blind. Participant characteristics are shown in Table 3.3. All participants provided informed consent prior to participation and the study was approved by the University of Liverpool Ethics committee.

3.4.2.2 Experimental conditions

Data for the neutral-lab condition were drawn from a larger EMA study (chapter five), where participants came for an initial lab visit in the university, before the start of the EMA phase. In that visit, participants completed an alcohol-Stroop task on a smartphone (Alcatel OneTouch Pixi 3 and Samsung Galaxy Ace 3) the AUDIT and the TLFB. Block order in the Stroop (alcohol block or control) was counterbalanced across participants.

Data for the real-world condition were drawn from a larger EMA study (see study one in chapter four), where participants were lent a smartphone (Samsung Galaxy Ace GT-S5830i) for one week to complete three alcohol-Stroop assessments per day, in their natural environments. Participants also completed an AUDIT and a TLFB during an initial visit in the lab, before the start of the EMA phase. The current study included data from the first assessment (10:30-12:30) on the first EMA day. The reason for selecting this particular assessment was because it represented the first time that participants used the smartphone-based alcohol-Stroop task; therefore, task novelty remains constant across conditions. Block order in the EMA study were the data were drawn from was counterbalanced across assessments, therefore in the first assessments on the first day the alcohol-block was always presented first.

Data for the semi-natural condition were drawn by extra participants recruited for the purpose of this study. Participants were invited in the Bar-Lab, which is a laboratory specially adapted to look like a pub (the room has pub-style tables and chairs, a bar, a large display of bottles of alcoholic drinks, a fridge with alcoholic drinks, and posters of alcoholic drinks). In the Bar-Lab participants completed an alcohol Stroop-task on a smartphone (Samsung Galaxy Ace 3), an AUDIT and a TLFB. Block order in the Stroop was counterbalanced across participants.

3.4.2.3 Materials

Participants completed the AUDIT and the TFLB questionnaires, as well as the alcohol-Stroop task. For a description of these materials see study one.

3.4.2.4 Alcohol Stroop-task stimuli

The alcohol Stroop-task contained words generally related with alcohol. See general-words stimuli in study one.

3.4.2.5 Data reduction and analysis

AB scores and internal reliability were calculated in the same way as study one.

Predictive validity of the Stroop-task was examined as described in study one, although no multicollinearity occurred in the current study, so interaction terms between AB score and condition were entered in the model.

Presence of AB for alcohol-related cues was examined with a 2x3 mixed ANOVA on reaction times, with word type (alcohol or control) as the within-subjects factor and condition (neutral lab, semi-natural, real world) as the between-subjects factor. Differences in AB magnitude were examined with a one-way ANOVA on AB scores, with condition as the between-subjects factor.

3.4.3 Results

3.4.3.1 Participants' characteristics

Conditions were compared in terms of gender ratio, age, TLFB and AUDIT scores. Regarding gender ratio, there was no significant difference among conditions *($\chi^2(2) = 5.95$, $p = .051$, Cramer's $V = .22$).*

Due to age and TLFB scores being skewed even after transformation, a non-parametric test was used (Kruskal-Wallis). Conditions differed in terms of age *($H(2) = 12.79$, $p = .002$).* Pairwise comparisons[10] (with p values corrected for multiple comparisons with a Bonferroni correction), revealed that participants in the semi-natural condition were older than in the neutral-lab condition *(Std. test statistic = -3.52, $p = .001$).* All other comparisons were not significant *($ps > .05$).* Conditions did not differ in terms of TLFB scores *($H(2) = 1.16$, $p = .560$).*

[10] Taking the difference between the average ranks of the different conditions and comparing it to a critical difference.

Due to AUDIT scores being skewed, a square root transformation was conducted before entering data in a one-way between-subjects ANOVA. There was no difference among conditions in terms of AUDIT scores *($F (2, 119) = 1.52$, $p = .223$, $\eta_p^2 = .025$).*

3.4.3.2 Internal reliability

Acceptable levels of reliability were achieved in the real-world condition *(α=.70),* but not in the neutral-lab condition *(α=.55)* or the semi-natural condition *(α=.57).*

3.4.3.3 Predictive validity: Regression between AB, condition, and alcohol involvement (Table 3.4)

After controlling for age and gender, neither AB nor the interaction between AB and condition predicted alcohol involvement *(ps > .05).*

3.4.3.4 Presence of AB

A 2x3 mixed ANOVA revealed a significant main effect of type of word (alcohol or control) on reaction times *($F (1, 120) = 52.26$, $p < .001$, $\eta_p^2 = .30$)* with participants responding slower to alcohol-related words than control. This reveals a significant AB for alcohol-related cues.

There was also a significant main effect of condition (neutral-lab, semi-natural or real world) on reaction times *($F (2, 120) = 3.36$, $p = .038$, $\eta_p^2 = .05$).* Post-hoc analysis, with p values corrected for multiple comparisons with Bonferroni correction, showed that participants had slower reaction times in the neutral-lab compared to the real world (p = .032). All other comparisons were not significant *(ps > .05).*

There was no significant interaction between type of word and condition *($F (2, 120) = .19$, $p = .829$, $\eta_p^2 = .003$).* This demonstrates that participants showed AB for alcohol across all environments.

3.4.3.5 Magnitude of AB

A one-way ANOVA showed no effect of condition on magnitude of AB *($F (2, 120) = .19$, $p = .829$, $\eta_p^2 = .003$)*[11]. This finding indicates that AB was equally strong across all environments.

[11] Repetition of the same statistics for the interaction term (condition and type of word) and the one-way ANOVA is not a typo. Both analyses return exactly the same set of numbers.

Table 3.4. Participants characteristics - Study two

	Condition			
	Neutral-lab (n = 41)	Semi-natural (Bar-Lab) (n = 41)	Real world (n = 41)	Sample (N = 123)
Gender (F:M)	37:4	28:13	32:9	97:26
Age	24.63 (±8.73)	30.07 (±9.34)	25.78 (±8.13)	26.83 (±8.99)
AUDIT [a]	8.32 (±3.29)	9.93 (±4.16)	9.13 (±4.95)	9.12 (±4.20)
Weekly consumption (units[b])	17.96 (±18.95)	16.61 (±11.80)	18.43 (±11.93)	17.68 (±14.53)
Reaction time (RT) for alcohol worlds (ms [c])	1157.42 (±161.07)	1109.73 (±135.04)	1073.03 (±175.92)	1113.39 (±160.74)
RT for control words (ms [c])	1103.16 (±149.74)	1063.21 (±143.54)	1015.98 (±162.42)	1060.79 (±155.04)

Values are means ±SD, except for gender for which they represent frequencies. a.AUDIT = Alcohol use disorders identification test scores range between 0 (minimum) and 40 (maximum). Scores of 8 or above indicate at least hazardous use of alcohol. b. 1 UK unit = 8g of alcohol. c. ms = milliseconds

Table 3.5. Regression showing the association between AB and alcohol involvement, adjusting for age and gender - Study two.

Steps	Model fit		Predictors[b]	Individual coefficients
	F	R^2		β^c
Step 1	5.60*	.09	Gender	.24*
			Age	-.19*
Step 2	3.01*	.09	Gender	.24*
			Age	-.19*
			AB[a]	.08
			Condition	.04
Step 3	2.39*	.09	Gender	.24*
			Age	-.19*
			AB[a]	.09
			Condition	.04
			AB[a] x condition	-.01

a. AB = Attentional Bias, measured in milliseconds. b. Age, AB and outcome variable (alcohol involvement) were standardized before inserted in the model. * $p<.05$
AB scores were not associated with alcohol involvement, regardless of condition.

Figure 3.3. Mean reaction time in the alcohol Stroop-task, split per condition and type of stimuli - Study two

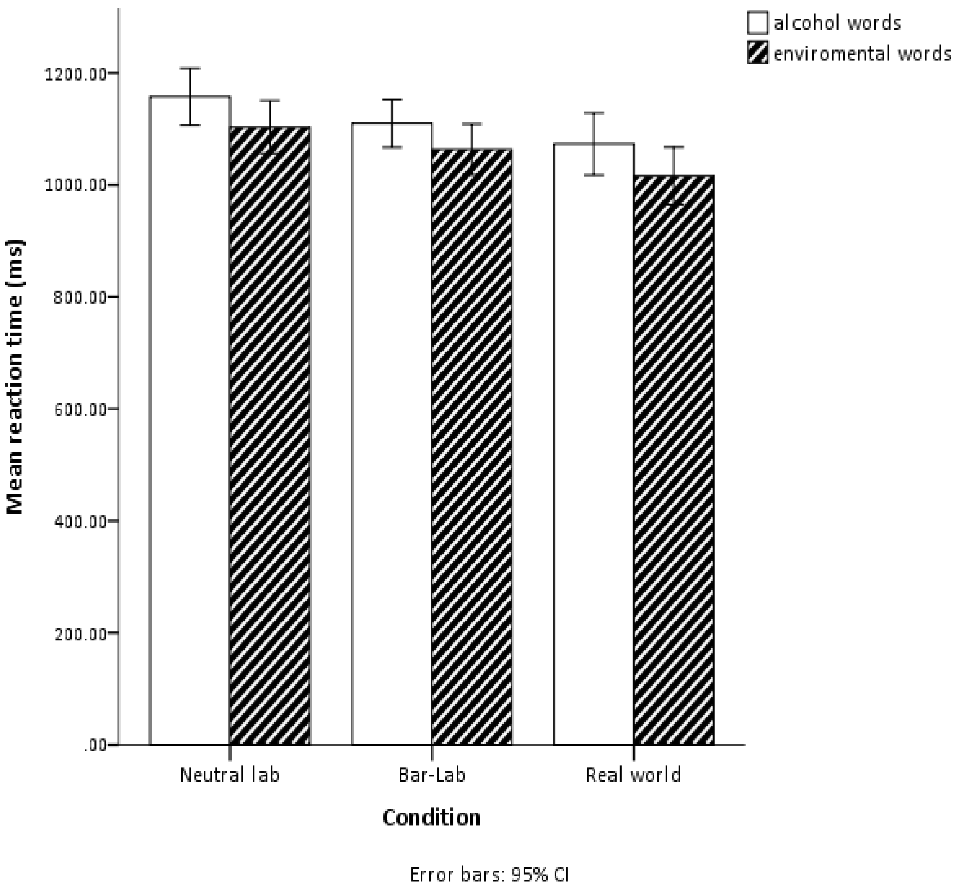
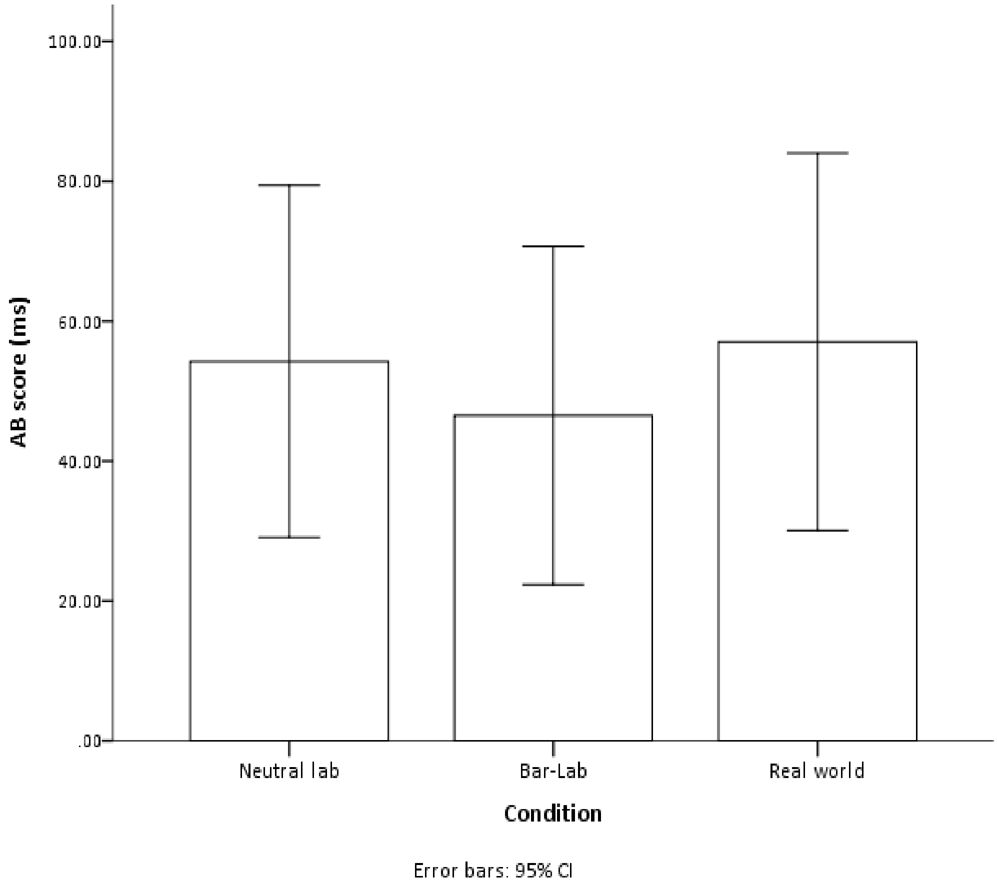


Figure 3.4. AB scores, split per condition - Study two



3.4.4 Discussion

The alcohol Stroop-task was administered on a smartphone in a neutral laboratory in University of Liverpool (neutral-lab condition), in the Bar-Lab of University of Liverpool (semi-natural condition) and in participants' natural environments (real-world condition). In each environment, the study examined the internal reliability and predictive validity of the task, as well as presence and magnitude of AB. Findings showed acceptable reliability of the smartphone-based alcohol Stroop-task in the real world, replicating results from study one. In all other environments, reliability was not acceptable, showing that the ubiquitous nature of smartphones is not a sufficient factor on its own to produce consistent responding. Also, none of the AB measures was associated with alcohol use behaviour, revealing generally poor predictive validity. Finally, AB of comparable strength was found across environments.

Interestingly, the smartphone-based alcohol Stroop-task had acceptable reliability in the real world but not in the Bar-Lab, which is a room specially adapted to create the impression of a real pub. One explanation might be that when participants were in the Bar-Lab, they still felt like their behaviour was closely observed, which might have made them less comfortable, compared to a real environment. They may also have realized that although alcohol is present it will not be available for consumption. A previous study in University of Liverpool (Christiansen, Townsend, Knibb, & Field 2017) compared the Bar-Lab to a neutral lab and they also failed to detect any effects of the different contexts. In both rooms, they provided participants with an alcohol placebo (a non-alcoholic drink masked as alcoholic) or a non-alcoholic control drink. Although consumption of the alcohol-placebo drink was positively associated with all outcome variables, this was not moderated by environmental context. Taken together, results might indicate that the naturalistic appearance of the Bar-Lab might not outweigh participants' knowledge that they are in a controlled environment.

This study had some limitations. First, data for the neutral-lab and real-world condition were drawn from separate, larger studies. However, the same recruitment process was used in both studies and participants were drawn from the same population (mainly stuff and students of University of Liverpool).

Second, presentation order of blocks (alcohol and control) in the Stroop-task was not counterbalanced in the real-world condition. Data for that condition were drawn from the first assessment of a larger EMA study (study one chapter four) where block order was counterbalanced across assessments rather than participants. As a result, the first assessment had the same presentation order (alcohol first) for all participants. However, it is unlikely that this would have an impact on the consistency of responses. Counterbalancing could have been restored by drawing data from multiple EMA assessments, although that would come at the expense of task novelty. Mixing measures when participants used the task for the first time with measures when participants have already had some practice might have made responding appear as less consistent. Indeed, a post –hoc analysis showed an internal reliability of α =.45 when novel and non-novel assessments were combined (block order was counterbalanced in this case).

Third, the specific environment were participants completed the alcohol Stroop-task in the real-world condition was not recorded. It is notable that the task was reliable even though it might have been completed in a great variability of locations (e.g. at home, at work, etc.). However, it would be important for future studies to gather information on the different natural environments (potentially using the phone's geo-location - GPS) where participants completed the alcohol Stroop-task. This would allow for a more fine-grained examination of the psychometric properties of the task within specific variants of "natural environments".

3.5 General Discussion

Studies one and two contrasted the psychometric properties of the alcohol Stroop-task when administered in EMA modality (on smartphones in the real world) with more traditional modalities (laptops/smartphones in laboratory environments). Both studies found acceptable internal reliability in EMA modality of administration. However, reliability was not acceptable when the task was administered in laboratory environments, regardless of means of administration (smartphone or laptop). Taken together, findings support EMA administration as a better alternative for measuring AB.

Comparable results have been found by Waters et al. (2012) who lent a personal digital assistance (PDA) to heroin and cocaine dependent patients so that they could complete a drug Stroop-task as they went about their day. The drug Stroop-task showed acceptable reliability when administered on the PDA and outside of a laboratory, albeit only for the heroin-related stimuli. Although the sample in Waters' study differed from the current studies (drug dependent patients instead of non-dependent social alcohol drinkers), their study provides another example of reliable AB measures in EMA modality of administration.

The alcohol Stroop-task was reliable when administered on a smartphone in the real world (study one and two) but not when administered on the smartphone in two laboratory environments (the neutral lab and the semi-naturalistic lab - study two). This suggests that the context in which the measures are taken (i.e. the real world) might play a more important role in increasing reliability, rather than the mean of administration (i.e. the smartphone).

One possible explanation for this finding might be that measures in natural environments are more sensitive to the underlying motivational aspects of AB than laboratory measures. Alcohol-related cues might automatically attract attention in a more systematic way in a natural environment, because internal and contextual factors that influence AB (as well as fluctuations in these factors) may be more prominent in natural environments. Some examples of such factors can be the presence of alcohol in the environment (Cox et al. 2003; Ramirez et al. 2015), expectancy that alcohol will be available for consumption (Field, Hogarth, et al. 2011; for a review see Jędras et al. 2013), craving (see meta-analysis by Field et al. 2009), and current mood (Field & Powell 2007; Field & Quigley 2009; Grant et al. 2007). Another explanation could be that participants might feel more comfortable and relaxed in their natural environments, where there is no close observation of their behaviour.

Despite the acceptable reliability in EMA modality, both studies revealed poor predictive validity of the alcohol Stroop-task. Measures of AB were not associated with alcohol-use behaviour, adding to the equivocal literature. Although some studies have found cross-sectional associations between AB and alcohol-use behaviour (Field, Christiansen, et al. 2007; Garland, Franken, & Howard 2012, for a review see Field & Cox 2008), many studies have failed to replicate this (Christiansen, Mansfield, et al. 2015; Wiers et al. 2017; for a review see Christiansen, Schoenmakers, et al. 2015).

However, the lack of association between AB and alcohol involvement does not necessarily mean that the alcohol Stroop-task is an inappropriate measure of AB. Operationalization of predictive validity as a cross-sectional association between AB and alcohol-use behaviour (specifically retrospective weekly consumption/AUDIT scores) might be flawed. This operationalization would assume that AB is a trait-like phenomenon, indicative of the general motivation to drink alcohol. However, the state hypothesis of AB (Field et al. 2016) would suggest that AB might be best conceptualized as a state-phenomenon which fluctuates under the influence of the current motivational context, and which would show a short temporal relationship with alcohol use behaviour, rather than a general association across different times and places.

Taken together, present findings provide evidence for the acceptable reliability of real-world measures of AB and show that EMA methodologies may offer a way of more reliably assessing AB. Although the utilization of smartphones is a practically sensible option, the ubiquitous nature of those devices does not seem to contribute to the acceptable reliability observed here. Present findings are not encouraging regarding the predictive validity of real-world AB measures, although future studies could investigate whether the observed lack of predictive validity is indicative of poor psychometric properties of the alcohol Stroop-task or a product of the erroneous conceptualization of AB as a trait phenomenon. This question will be explored in the EMA studies in the following chapters.

Chapter four: EMA studies one and two

4. Close encounters: Temporal associations among attentional bias, craving, and alcohol consumption, in the real world.

4.1 Abstract

Previous findings regarding the prospective associations between attentional bias (AB) and alcohol consumption have been inconsistent. One explanation for this might be that AB reflects a state, rather than a trait, so it will only have a short-term temporal relationship with alcohol consumption (the state-hypothesis for AB). The two studies in this chapter utilized ecological momentary assessment (EMA) to examine the associations among AB, craving, intention to drink alcohol, and alcohol consumption. Regular social drinkers carried a smartphone for one week with an app (AlSAp) installed. At three times per day (morning, early afternoon, late afternoon), participants completed a Stroop-task with alcohol-related words (study one) or a personalized pictorial Stroop-task (study two) and reported craving for alcohol. At the beginning of each day participants reported their intention to drink alcohol, and at the end of the day (or first thing the next morning) they reported their quantity of alcohol consumption (in units). Participants' alcohol consumption was restricted to after completing the last smartphone assessment (late afternoon). Both studies revealed that a) global measures (measures taken across the whole day) of craving and AB were not associated with each other, b) craving, but not AB, was higher on days when participants had plans to consume alcohol, c) on days when alcohol was consumed, craving, but not AB, was gradually increasing and peaked in the late afternoon, and d) craving predicted alcohol use when measured in the late afternoon but not earlier. Also, study one revealed that participants showed significant AB in all assessments of the day, and that AB predicted alcohol use when measured in the late afternoon but not earlier. However, study two did not replicate this. These studies provide evidence for the importance of temporal proximity between alcohol consumption and AB/craving assessments, in order to understand their prospective relationships. However, they provide only partial support to the state hypothesis of AB. They also demonstrate the importance of incorporating EMA and smartphone applications in the study of AB and craving in social drinkers.

4.2 Introduction

Cues in the environment that have been repeatedly paired with substance use acquire motivational salience (Robinson & Berridge 1993). These cues can then automatically grab and hold the attention of the substance user, a phenomenon known as attentional bias (AB; Franken 2003). It is argued that AB plays an important role in the development and maintenance of alcohol misuse (for example, Field & Cox 2008).

Although some studies in non-dependent drinkers have found an association between AB and alcohol use (e.g. Christiansen and Bloor, 2014; Fadardi and Cox, 2009; Field, Christiansen, Cole and Goudie, 2007), there are numerous examples where these associations have not been replicated (e.g. Christiansen, Mansfield, Duckworth, Field and Jones, 2015;Field, Mogg, Zetteler and Bradley, 2004;Groefsema, Engels, Kuntsche, Smit and Luijten, 2016). Evidences are also equivocal regarding the clinical utility of AB in predicting relapse after a period of abstinence (see for example Cox, Hogan, Kristian and Race, 2002, but also Field, Mogg, Mann, Bennett and Bradley, 2013, for a review see Christiansen, Schoenmakers, et al. 2015).

An explanation that has been suggested for the inconsistent results described above is the state hypothesis of AB (AB is a motivational state that would show only transient associations with substance use; see Field et al. 2014; Field et al. 2016). This hypothesis would suggest that AB should fluctuate under the influence of the current motivation to use the substance. For example, AB has been found to be associated with feelings of craving, both in the laboratory (Field, Mogg, Mann, Bennett and Bradley, 2013; Manchery, Yarmush, Luehring-Jones and Erblich, 2017; for a meta-analysis see Field, Munafò and Franken, 2009) and in the real world (Waters et al. 2012; Waters, Szeto, Wetter, Cinciripini, Robinson and Li, 2014). Moreover, AB has been found to elevate when participants expected the substance to become available for consumption (Field, Hogarth, Bleasdale, Wright, Fernie and Christiansen, 2011; Jones, Hogarth, Christiansen, Rose, Martinovic and Field, 2012; for a review see Jędras, Jones and Field, 2013).

However, it is reasonable to expect that the current motivation to consume the substance would change rapidly. Indeed, previous naturalistic studies have shown that craving increased when measured in contexts where motivation to drink alcohol was expected to be high. For example, a study found that smokers in a non-smoking flight showed gradually increasing craving, which peaked in the assessment most proximal to landing (Dar, Rosen-Korakin, Shapira, Gottlieb and Frenk, 2010). Notably, the study also demonstrated that this was a function of proximity to the opportunity to smoke rather than duration of abstinence. In another study (Dar, Stronguin, Marouani, Krupsky and Frenk, 2005) Orthodox Jews were found to have higher levels of craving on days when smoking was generally available, compared to days when they abstained for religious reasons

As AB is expected to constantly fluctuate under the influence of the motivational context, the state hypothesis also suggests that AB would show only short-term associations with substance use. Consequently, AB measured in a laboratory is likely to show little concordance with substance-use behaviours occurring in different contexts in the real world, several days or even weeks later (as for example when trying to predict relapse after discharge from an inpatient treatment program). Instead, associations between AB and substance use should become apparent if AB is measured shortly prior to substance use and within the same real-world contexts that substance use takes place (for example using ecological momentary assessment – EMA; Shiffman, Stone and Hufford, 2008)

Indeed, an ecological momentary assessment (EMA) study with drug-dependent inpatients (Marhe, Waters, van de Wetering and Franken, 2013) showed that AB measured during an episode of temptation to use drugs (but not in random moments) predicted relapse in the sort- to medium-term, but not in the long-term. The study also found that, surprisingly, craving did not predict relapse. However, other EMA studies have demonstrated that craving was associated with substance use in the short-term but not in the long-term (Fatseas et al. 2015) and that craving was a better predictor of relapse after a period of abstinence when measured proximally to substance use (for a meta-analysis see Serre et al. 2015).

However, no studies have used EMA to examine AB for alcohol-related cues and its association with alcohol consumption. The aim of the two studies presented in this chapter was to examine variations in AB and craving over the course of a drinking day, as well as the temporal associations between AB craving and initiation of alcohol consumption. The studies also examined the effect of intention to consume alcohol on a given day on the global measures (measures takes across the whole day) of AB and craving, as well as the association between global AB and craving. Study one measured AB with a classic alcohol-words Stroop-task, while a personalized alcohol pictures version was used in study two.

For both studies, it was predicted that a) global measures of AB and craving will be significantly associated with each other, b) AB and craving will be greater on days when participants had plans to drink, c) AB and craving will fluctuate over the course of a drinking day and will peak in the session that was most proximal to alcohol use, and d) AB and craving will be a better predictor of alcohol use when measures are taken proximally to initiation of drinking.

4.3 Study one

4.3.1 Methods

4.3.1.1 Participants (table 4.1)

Participants were recruited from the University of Liverpool and local community, via intranet advertisement and word of mouth. To be eligible, participants should be fluent English speakers and regular social drinkers (self-reporting alcohol consumption at least once per week). They were excluded if they had any current or previous alcohol abuse or dependence diagnosis, were pregnant or breastfeeding, or suffered from any form of colour-blindness (all based on verbal and written confirmation by the participants). Sixty-three participants were initially recruited, but one participant dropped out of the study before the beginning of the EMA phase, leaving 62 participants to be included in the analysis (46 females, average age 25.89 (±8.25) years). All participants provided informed consent prior to participation and the study was approved by the University of Liverpool Ethics committee.

4.3.1.2 Computer and smart-phones

A training tutorial on the app was completed at baseline, in the laboratory, on a standard laptop computer (Toshiba Satellite Pro A200). Smart-phones (Samsung Galaxy Ace GT-S5830i) were loaned to participants for the daily assessments over the EMA week.

4.3.1.3 Alcohol Stroop-task stimuli

The alcohol Stroop-task stimuli were words generally related to alcohol and to environmental features. For more details see chapter three, section 3.3.2.3. Stimuli were presented to participants in three different styles; a block of alcohol-related stimuli followed by a block of control stimuli, a block of control stimuli followed by a block of alcohol-related stimuli, or two blocks of mixed alcohol-related and control stimuli. Each presentation style appeared once every day. Order of presentation styles (whether they appeared in the morning, early afternoon or late afternoon) was counterbalanced across the week.

4.3.1.4 Baseline measures

The *Time Line Follow Back (TLFB; Sobell & Sobell 1990)* and the *Alcohol Use Disorder Identification Test (AUDIT; Saunders et al. 1993)* were used to assess typical alcohol consumption and hazardous drinking habits, respectively. For a detailed description see chapter two, sections 2.1.1 and 2.1.3.

4.3.1.5 EMA measures (table 4.2)

The Alcohol-Stroop App (AlSAp) was used to measure intention to drink, AB and craving. For a detailed description of AlSAp see chapter two, section 2.2. A paper diary was used to assess quantity of alcohol consumption (in units) for each day over the EMA period. For a detailed description see chapter two, section 2.1.2.

4.3.1.6 Procedure

Interested participants visited a laboratory in University of Liverpool where they were screened for eligibility and provided informed consent. Training for using the app was provided in the form of a laptop task, identical to that on the smartphones. At the end, participants were given a smartphone with the AlSAp installed.

Participants were instructed to start using the app the following day and for the next seven days, three times per day. Assessments were required to be completed within pre-specified time windows; a morning period from 10:30 to 12:30, an early-afternoon period from 13:30 to 15:30 and a late-afternoon period from 16:30 to 18:30. Testing was not extended beyond 18:30 as there were more chances for participants to initiate a drinking sessions after that time (no alcohol should have been consumed by participants before completing the last AlSAp assessment, see below). No prompts to remind participants to access the app were provided but participants were given a paper table outlining the assessments schedule for each day and were instructed to consult the table every time they felt unsure.

The AlSAp was programmed to show the intention to drink question only in the morning sessions. AB and craving (craving was measured as a continuous variable with a 1-11 Likert scale, where 1 denoted no craving and 11 very strong craving) were assessed in every AlSAp session (in the morning, early afternoon, and late afternoon). At the end of each day, or first thing the next morning, participants reported on the diary the number of alcohol units they have consumed that day. Participants were instructed not to drink alcohol before completing the late-afternoon AlSAp session, although compliance to this could not have been confirmed in any other way than self-report. Alcohol use was restricted to evenings to create a continuum of temporal proximity of AlSAp assessments to alcohol use, from distant proximity (morning sessions) to moderate proximity (early-afternoon sessions) and close proximity (late-afternoon sessions). It also ensured sobriety when completing the AlSAp assessments.

In the event of any technical failure (e.g. the app crashing half-way a session), participants were instructed to repeat the session immediately until it was completed. In the event of persisting problems, participants were asked to return the faulty phone, receive a replacement and continue the study for as many days left until completing seven days of testing. Seven phones (11.3%) were returned due to persisting problems. The internal memories of these phones were wiped out and they were restored to industrial settings before used again in the study.

Upon completion of the EMA phase, participants returned the phone, received a thorough debriefing, and were given compensation in the form of either course credits or shopping vouchers (worth £20). Participants received full compensation when initiating at least 16 out of 21 sessions (75% study engagement), reduced compensation for at least 11 initiated sessions (50% study engagement) and no compensation for fewer than 11 initiated sessions (see sections 4.4.1.1 and 4.4.2.1 for compliance results).

4.3.1.7 Data reduction and analysis

AB scores were calculated by subtracting reaction times (RTs) for control stimuli from RTs for alcohol stimuli. For a detailed description of AB scores calculation see chapter 2, section 2.3.1. Positive AB scores (slower RTs for alcohol compared to control stimuli) are indicative of AB towards alcohol-related stimuli.

Alcohol consumption was analysed as a dichotomous variable, coded as no or any alcohol consumption (initiation of drinking). The rationale for this choice was that elevations in AB prior to a drinking session will increase the motivation to choose drinking over not drinking. However, once drinking is initiated, acute alcohol effects will affect both AB and control over drinking (for a review see Field et al. 2010). Also, although the present studies included non-clinical samples, this analytical approach was chosen to resemble clinical studies, where the outcome of interest is whether participants would choose to consume alcohol or not, rather than how much alcohol they will drink.

Data had a hierarchical structure, with AlSAp sessions (first level) clustered within days (second level) and within participants (third level). Analysis was performed using multilevel models, as this allowed for dependent observations caused by clustering, and for unequal numbers of data points (Hayes 2006; Quené & van den Bergh 2004 - analysis was performed with MLwiN 2.30; Rasbash, Charlton, Browne, Healy, & Cameron 2009). The sample size in this study was above the recommended size (N > 50; Maas & Hox 2005) to ensure unbiased estimates of error in multilevel modelling. Regarding model parameters, regression intercepts were set at random, while slopes were kept fixed. For a detailed explanation of multilevel analysis and hierarchical data structures see chapter two, section 2.3.2. A prerequisite of multilevel analysis is for the outcome variable to show significant variation in the highest level of the structure (Peugh 2010). In cases where this assumption was violated, a regular single-level analysis was performed.

Table 4.1. Participants characteristics - Study one

Characteristic	Mean (SD) or Freqency (%)
Age	25.89 (±8.25)
Females	46 (74.19%)
Alcohol Use Disorder (AUDIT)[a] score	9.41 (±5.19)
Quantity of drinking in the week before the study (in units)	18.07 (±11.47)
Quantity of drinking in the EMA week (in units)	21.82 (±15.58)
Number of alcohol-consuming days in the EMA week	3.18 (±1.32)

a. Cut off score for hazardous drinking = 8.

Table 4.2. Daily measures over the EMA week - Study one

Measure	Mean (SD) or Frequency (%)
Days when alcohol was consumed	197 (45.4%)[a]
Quantity of drinking per day (in units)	3.09 (±5.26)
Days when intention to drink was declared	140 (32.3%)[b]
Assessments	
Morning	368 (35.1%)[c]
Early afternoon	347 (33.1%)[c]
Late afternoon	332 (31.7%)[c]
Attentional Bias scores (ms)	
Morning	31.30 (±70.56)
Early afternoon	39.11 (±77.71)
Late afternoon	18.37 (77.88)
Craving score[d]	
Morning	1.78 (±1.72)
Early afternoon	2.20 (±1.96)
Late afternoon	2.92 (±2.57)

a. Percentage is out of 429 days for which data on alchol consumption were available.

b. Percentage is out of 367 days for which data on intention to drink were available.

c. Percentages are out of 1047 assessments for which participants provided data.

d. Scored on a scale from 1 (no craving) to 11 (very strong craving).

4.3.2 Results

4.3.2.1 EMA adherence and reactivity

Regarding AB, participants initiated 1,091 alcohol Stroop-tasks (83.8% of possible 1,302) and completed 1,039 of them (79.8% of possible 1,302 and 95.2% of 1,091 initiated). The 52 incomplete alcohol Stroop-tasks were excluded from the analysis. Regarding craving, participants returned 1,025 (78.7% of possible 1,302) answered questions[12]. They also answered 367 (84.6% of possible 434) intention to drink questions and provided 429 (98.8% of possible 434) alcohol diary entries. Participants provided data in 368 morning (84.8% of possible 434), 347 early-afternoon (80% of possible 434), and 332 (76.5% of possible 434) late-afternoon assessments.

To examine for reactivity effects (participants reducing their alcohol consumption over the EMA week due to intense self-monitoring), quantity of alcohol consumption the week before the study was compared to quantity consumed over the EMA week. A bootstrapped dependent t-test (used due to violation of parametric assumptions) showed no differences (drinking before EMA: 18.07 (±11.47) units, drinking over EMA: 21.82 (±15.58) units; $t(59) = -2.02$, $p = .05$), indicating no reactivity effect.

4.3.2.2 Craving

4.3.2.2.1 Selection of optimal data structure

When craving was examined within all days (drinking and non-drinking), a 3-levels structure (sessions within days within participants) fitted the data better than any other data structure (a 'sessions within days' 2-levels structure ($\chi^2 = 175.45$, $df = 1$, $p < .001$), a 'sessions within participants' 2-levels structure ($\chi^2 = 28.77$, $df = 1$, $p < .001$), or a single-level structure ($\chi^2 = 401.85$, $df = 2$, $p < .001$) of the data). Mean craving per session was $\beta_0 = 2.29$ (*Standard Error, SE = 0.18*). Thirty-eight per cent of variance in craving was attributed to between participants differences (3rd level of analysis), while 13.7% was attributed to between days differences within the same participant (2nd level of analysis), and 48.2% to between sessions differences within the same day (1st level of analysis).

When craving was examined within alcohol-drinking days only, a 3-levels structure (sessions within days within participants) fitted the data better than a 'sessions within days' 2-levels

[12] Although a craving question appeared at the end of every alcohol Stroop-task, different completion rates might be due to the app being aborted prematurely, either by the participant or by the app crashing.

structure *(χ^2=79.24, df=1, p<.001) but not* a ‘sessions within participants’ 2-levels structure *(χ^2=4.97, df=1, p= .052)*. A ‘sessions within participants’ 2-levels structure, fitted the data better than a single-level structure (*χ^2=157.56, df=1, p<.001*). Therefore, the ‘sessions within participants’ 2-levels structure was retained as best fitting the data. Mean craving per session was *β_0 = 2.66 (SE = 0.22).* Forty-one point one per cent of variance in craving was attributed to between participants differences (2nd level), while 58.9% was attributed to between sessions differences within the same participant (1st level).

4.3.2.2.2 Association of craving with intention to drink and AB.

The research question was whether craving would be higher on days when participants have pre-determined plans to consume alcohol and when AB was also high. A 3-levels (sessions within days within participants) regression model was fitted to predict craving from AB and intention to drink as first and second level explanatory variables, respectively. The model predicted 5% of variance in craving scores at the session level, 23.05% at the day level, and 9.89% at the participant level. Intention to drink (*β = 0.80, (SE) = 0.11, p < .001, n=942 sessions*), but not AB (*β* < 0.001, (*SE*) = 0.001, *p* = 1.00, *n=942 sessions*), was positively associated with craving. Craving was higher in sessions that belonged to a day with a declared intention to drink, compared to days with no intention to drink. However, sessions that measured high in AB did not necessarily measure high in craving too. This suggests that craving was influenced by participants’ expectancy to consume alcohol that day, but not by their AB.

4.3.2.2.3 Fluctuations in craving over the course of a drinking day

The research question was whether craving will gradually increase over the course of a drinking day as a function of its proximity to alcohol use. To test that, only days when alcohol was consumed were included in the analysis. A 2-levels regression model was fitted to predict craving from period of testing (morning, early afternoon, or late afternoon) as first level explanatory variables. The model was fitted three times. The first time, late-afternoon sessions were excluded from the analysis, so that the dummy variable in the regression would represent the comparison between morning and early-afternoon sessions. Morning sessions were excluded the second time to compare early afternoon with late afternoon, and early-afternoon sessions were excluded the third time to compare morning with late afternoon. P values were corrected for multiple comparisons.

The first model (morning vs early afternoon) explained 25.7% of variance in craving in the participant level and 21.5% in the session level. Craving was higher in the early afternoon than in the morning (*$\beta = 0.49$, $SE = 0.18$, $p = .021$, $n = 333$ sessions*).

The second model (early afternoon vs late afternoon) did not explain any further variance in craving in the participant level, compared to an empty model with no explanatory variables. However, it explained 20.9% of craving in the session level. Craving was higher in the late afternoon than in the early afternoon (*$\beta = 1$, $SE = 0.19$, $p < .001$, $n = 314$ sessions*).

The third model (morning vs late afternoon) did not explain any further variance in craving in the participant level, compared to an empty model without explanatory variables. However, it explained 1% of variance in craving in the session level. Craving in the late afternoon was higher than in the morning (*$\beta = 1.51$, $SE = 0.20$, $p < .001$, $n = 329$ sessions*).

Results indicate the craving was gradually increasing over the course of a drinking day and peaked when measured most proximally to alcohol use (Figure 4.1; changes in craving over the course of non-drinking days are also presented for comparison).

4.3.2.3 Attentional bias

4.3.2.3.1 Selection of optimal data structure

When AB was examined within all days (drinking and non-drinking) a 3-levels structure (sessions within days within participants) did not fit the data better than a 'sessions within days' 2-levels structure, or a 'sessions within participants' 2-levels structure (*all $ps > .05$*). Also, none of the 2-levels structures fitted the data better than a single-level structure (*all $ps > .05$*). Therefore, a single-level structure was retained as best fitting the data and a regular single-level analysis was performed.

When AB was examined within drinking days only, the pattern was the same as above. Therefore, in this case too, a single-level model was retained as best fitting the data and regular single-level analysis was performed.

4.3.2.3.2 Presence of AB over the course of the day

The research question was whether there was significant AB for alcohol-related cues across the testing periods (morning, early afternoon, and late afternoon). Due to violation of parametric assumptions, a bootstrapped one-sample t-test was used to examine whether average AB score was significantly different from 0. Average AB score in all testing periods was positive and significantly different from 0 (*morning: $t(365) = 8.49$, $p = .001$, $d = .44$, $n =$*

366 sessions; early afternoon: t (344) = 9.35, p =.001, d = .50, n = 345 sessions and late afternoon: t (326) =4.26, p =.001, d = .24, n = 327 sessions). Results indicate that a significant AB for alcohol-related cues was present at all assessments over the course of the day.

4.3.2.3.3 Association of AB with intention to drink

The research question was whether AB would be higher when participants have pre-determined plans to consume alcohol. Due to violation of parametric assumptions, the Mann-Whitney U test was used to compare days when intention to drink was declared with days when no intention was declared. No difference in magnitude of AB was found between groups of days (*Z = -.26, p = .80, n = 959 sessions*). Results indicate that AB was not affected by participants' expectancy to consume alcohol on a given day.

4.3.2.3.4 Fluctuations in AB over the course of a drinking day

The research question was whether AB will gradually increase over the course of a drinking day, as a function of its proximity to alcohol use. To test that, only days when alcohol was consumed were included in the analysis. Due to violation of parametric assumptions, a Kruskal-Wallis test was used to examine differences in magnitude of AB among testing periods (morning, early afternoon, and late afternoon). No differences were found (χ^2 *(2) = 5.96, p = .051, n = 493 sessions*). Results indicate the AB remained stable over the course of a drinking day (Figure 4.2; changes in AB over the course of the day for non-drinking days are also presented for comparison).

4.3.2.4 Initiation of alcohol consumption

4.3.2.4.1 Selection of optimal data structure

A 2-levels structure (days within participants) did not fit the data better than a single-level structure *(χ^2 = .01, df = 1, p = .94).* Therefore, a single-level structure was retained as best fitting the data and regular single-level analysis was performed with the data.

4.3.2.4.2 Prediction of initiation of alcohol consumption by AB and craving (table 4.3)

The research question was whether measures of AB and craving would be a better predictor of alcohol use when measured in greater proximity to initiation of drinking. A Ridge logistic regression (ridge parameters derived through the method described by Cule & De Iorio 2013) was used to test a model that predicted initiation of alcohol consumption with the three AB and craving measures (in the morning, early afternoon, and late afternoon) as predictors. Variables were measured repeatedly within a short period of time, and within the same individuals. For that reason, a Ridge regression was adopted, as opposed to a standard binary logistic regression, due to concerns over multicollinearity (see Månsson & Shukur 2011; Shen & Gao 2008). For example, correlations between craving measures were particularly high (*.54 to .71*).

Results showed a significant ridge logistic regression model (*multiple* $R^2 = 0.11$, $p < .001$, $n = 244$ *days*). Although AB in the morning (*scaled* $B = 0.74$, $SE=.39$, $p =.058$) and early afternoon ($B = 0.53$, $SE =.39$, $p =.17$) did not predict initiation of drinking, AB in the late afternoon showed a positive association with initiation of drinking ($B=0.92$, $SE=0.39$, $p =.018$). A similar pattern was also found for craving. Although morning ($B =0.20$, $SE=0.42$, $p =.63$) and early-afternoon ($B = 0.35$, $SE=.43$, $p =.42$) craving did not predict initiation of drinking, late-afternoon craving showed a positive association ($B= 1.40$, $SE=0.43$, $p=.001$) with initiation of drinking. Associations were also maintained after controlling for intention to drink. Results indicate that, over the course of a day, AB and craving showed a short-term, but not a mid- to long-term, association with alcohol use.

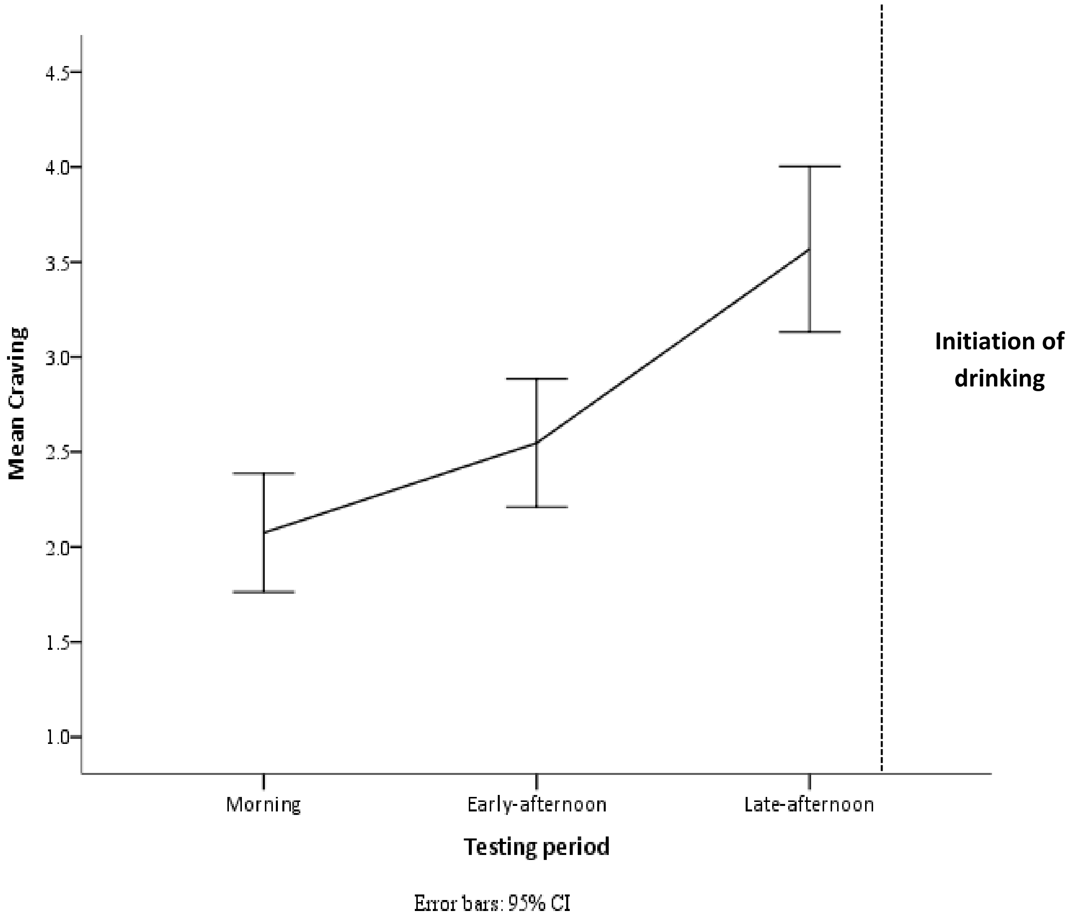
Table 4.3. Ridge logistic regression predicting initiation of drinking by AB and craving over the course of the day - Study one

	scaled B(SE)	multiple R^2
AB morning	.74 (.39)	
AB early afternoon	.53 (.39)	
AB late afternoon	.92 (.42) *	.11 **
Craving morning	.20 (.42)	
Craving early afternoon	.35 (.43)	
Craving late afternoon	1.40 (.43) *	

* p < .05; **p < .001

Figure 4.1. Craving (1-11 scale) over the course of drinking days (panel one) and non-drinking days (panel two) - Study one

Panel 1



Panel 2

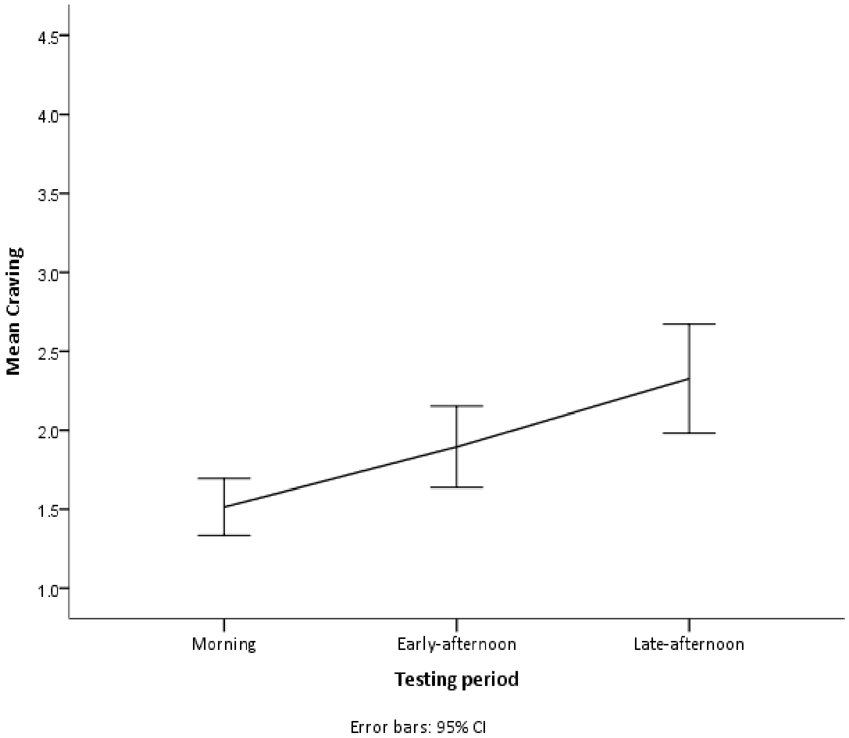
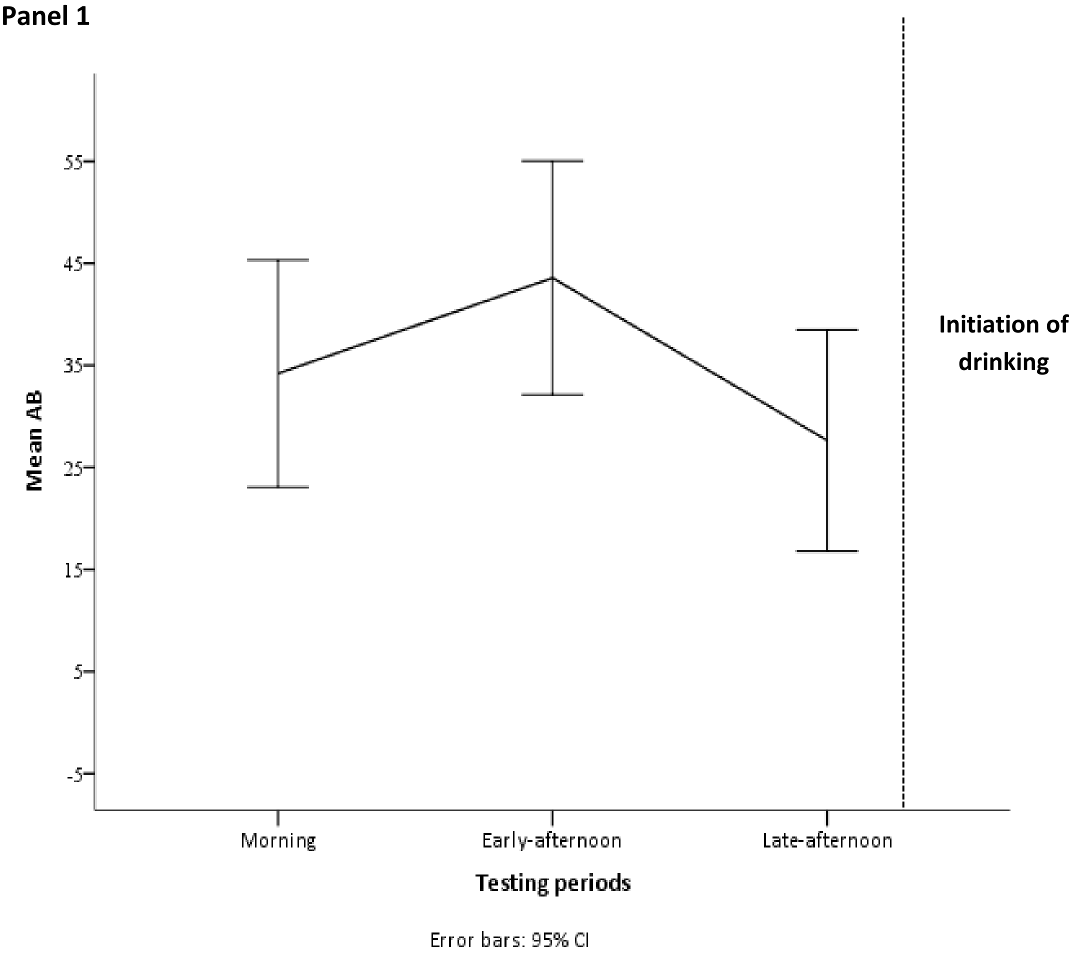
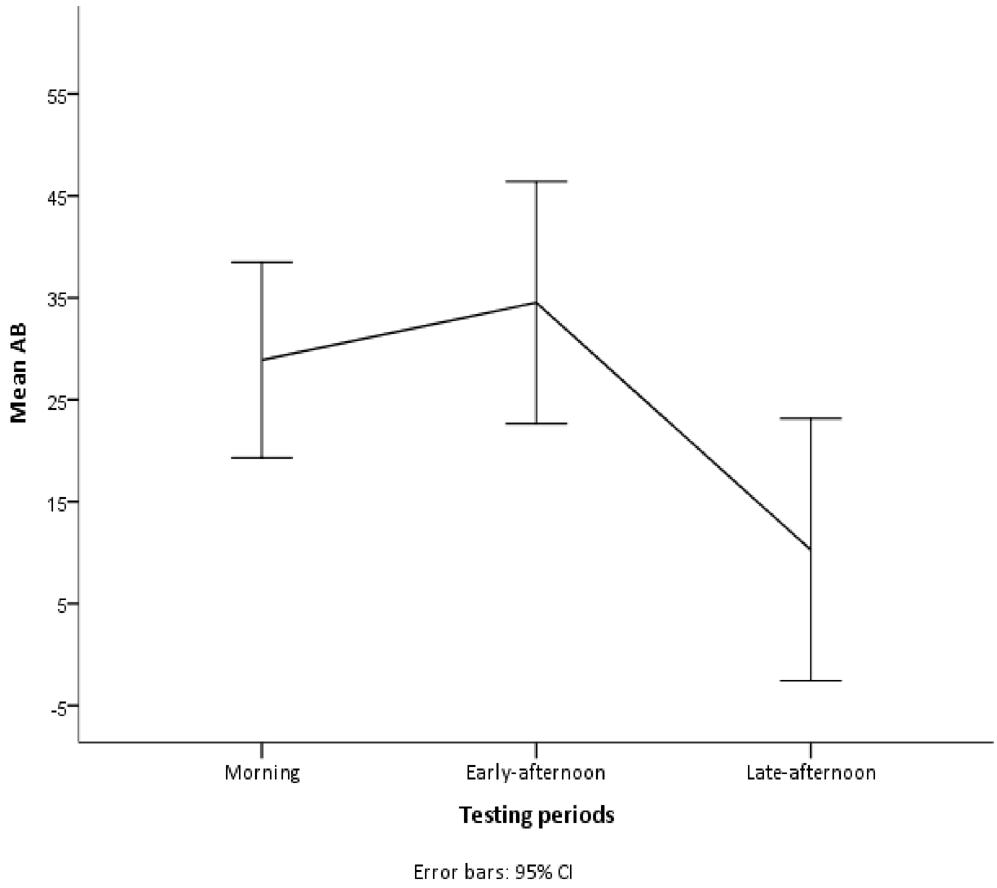

4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
Mean Craving
Morning
Early-afternoon
Late-afternoon
Testing period
Error bars: 95% CI

Figure 4.2. AB (ms) over the course of drinking days (panel one) and non-drinking days (panel two) - Study one

Panel 1



Panel 2



Mean AB
55
45
35
25
15
5
-5
Morning
Early-afternoon
Late-afternoon
Testing periods
Error bars: 95% CI

4.3.3. Interim summary of findings

Study one used EMA to examine AB (measured with a classic alcohol-words Stroop-task), craving, intention to drink and alcohol consumption over seven days. Findings showed that global measures of AB were not associated with global measures of craving. They also showed that craving, but not AB, was higher on sessions that belonged to days with declared plans to consume alcohol, compared to days with no plans. Moreover, over the course of a drinking day, craving, but not AB, was gradually increasing, peaking in the late afternoon (the assessment that was most proximal to initiation of drinking). Finally, measures of AB and craving predicted whether participants will choose to consume alcohol or not, only when taken proximally to initiation of drinking (late-afternoon assessment).

4.4 Study two

In study two, a different version of the alcohol Stroop-task was used which included personalised pictures of alcohol and pictures of soft-drinks. In all other aspects, the same methodology was applied.

4.4.1 Methods

4.4.1.1 Participants (table 4.4)

Sixty-one participants (40 females, average age of 20.86 (±4.30) years) were recruited. Recruitment procedures, eligibility criteria and ethical approval were the same as for study one. An additional inclusion criterion was that participants should self-report beer as their favourite alcohol drink.

4.4.1.2 Computer and smart-phones

Device information is the same as in study one, apart from the fact that smartphones loaned to participants were Alcatel Pixi 3 and Samsung Galaxy Ace 3 GT-S7275R.

4.4.1.3 Alcohol Stroop-task stimuli

The stimuli used in the alcohol Stroop-task were pictures of beer and soft-drinks. For more details see chapter three, section 3.3.2.3. Stimuli were presented to participants in two styles; a block of alcohol-related stimuli followed by a block of control stimuli, or vice-versa. Order of presentation styles (whether they appeared in the morning, early afternoon or late afternoon) was counterbalanced across the week.

4.4.1.4 Baseline measures

See study one.

4.4.1.5 EMA measures *(table 4.5)*

EMA measures were the same as in study one, apart from the fact that the diary to assess quantity of alcohol consumption (in units) was provided in a digital format installed on the smartphone. For a detailed description see chapter two, section 2.1.2.

4.4.1.6 Procedure

Procedures were the same as in study one, apart from the fact that after providing informed consent, participants also reported their four most preferred alcoholic drinks. Beer had to be the first drink they reported to be considered eligible (in order for beer pictures to be considered personalised stimuli). Eligible participants then completed a battery of questionnaires regarding their alcohol-use behaviour (TLFB, AUDIT). Another difference with study one was that the table outlining the assessment schedule was provided to participants in a digital format installed on the smartphone.

4.4.1.7 Data reduction and analysis

See study one.

Table 4.4. Participants characteristics - Study two

Characteristic	Mean (SD) or Freqency (%)
Age	20.86 (±4.30)
Females	40 (65.57%)
Alcohol Use Disorder (AUDIT)[a] score	10.62 (±4.20)
Quantity of drinking in the week before the study (in units)	27.34 (±15.38)
Quantity of drinking in the EMA week (in units)	31.45 (±18.73)
Number of alcohol-consuming days in the EMA week	3.62 (±1.39)

a. Cut of score for hazardous drinking = 8.

Table 4.5. Daily measures over the EMA week - Study two

Measure	Mean (SD) or Frequency (%)
Days when alcohol was consumed	219 (51.3%)[a]
Quantity of drinking per day (in units)	4.47 (±6.04)
Days when intention to drink was declared	151 (35.4%)[b]
Assessments	
Morning	375 (34%)[c]
Early afternoon	371 (33.6%)[c]
Late afternoon	357 (32.4%)[c]
Attentional Bias scores (ms)	
Morning	-3.13 (±83.94)
Early afternoon	-6.73 (±75.69)
Late afternoon	-10.36 (±81.10)
Craving score[d]	
Morning	1.87 (±1.68)
Early afternoon	2.35 (±1.10)
Late afternoon	3.15 (±2.52)

a. Percentage is out of 427 days for which data on alchol consumption were available.

b. Percentage is out of 370 days for which data on intention to drink were available.

c. Percentages are out of 1103 assessments for which participants provided data.

d. Scored on a scale from 1 (no craving) to 11 (very strong craving).

4.4.2 Results

4.4.2.1 EMA adherence and reactivity

Regarding AB, participants initiated 1,105 alcohol Stroop-tasks (86.3% of possible 1,281) and completed 1,090 (85.09% of possible 1,281, and 98.6% of 1,105 initiated). The 15 incomplete tasks were excluded from the analysis. Regarding craving, participants returned 1,090 answered questions (85.09% of possible 1,281). They also answered 370 intention to drink questions (86.7% of possible 427) and provided 427 alcohol diary entries (100% of possible). Participants provided data in 375 morning sessions (87.8% of possible 427), 371 early-afternoon sessions (86.9% of possible 427), and 357 late-afternoon sessions (83.6% of possible 427).

To examine for reactivity effects (participants reducing their alcohol consumption over the EMA week due to intense self-monitoring), quantity of alcohol consumption the week before the study was compared to quantity consumed over the EMA week. A bootstrapped dependent t-test (used due to violation of parametric assumptions) showed that participants consumed more alcohol over the EMA week (drinking before EMA: 27.34 (±15.38) units, drinking over EMA: 31.45 (±18.73) units; $t(60) = -2.09, p = .04$). This would suggest an alteration in participant's drinking patterns, although to the opposite direction than expected.

4.4.2.2 Craving

4.4.2.2.1 Selection of optimal data structure

When craving was examined within all days (drinking and non-drinking) a 3-levels structure (sessions within days within participants) fitted the data better than all other data structures (a 'sessions within days' 2-levels structure ($\chi^2 = 85.42, df = 1, p < .001$), a 'sessions within participants' 2-levels structure ($\chi^2 = 32.02, df = 1, p < .001$), and a single-level structure ($\chi^2 = 206.31, df = 2, p < .001$) of the data). Mean craving per session was $\beta_0 = 2.47$ (*Standard Error, SE = 0.15*). Twenty-one point eight per cent (28.1%) of variance in craving scores was attributed to between participants differences (3rd level of analysis), while 15.8% was attributed to between days differences within the same participant (2nd level of analysis), and 62.4% to between sessions differences within the same day (1st level of analysis).

When craving was examined within drinking days only, a 3-levels structure (sessions within days within participants) fitted the data better than a 'sessions within days' 2-levels structure (χ^2*=75.26, df=1, p<.001*), but not a 'sessions within participants' 2-levels structure (χ^2*=0, df=1, p=1*). The 'sessions within participants' 2-level structure fitted the data better than a single-level structure (χ^2*=75.26, df=1, p<.001*). Therefore, a 2-levels 'sessions within participants' structure was retained as best fitting the data. Mean craving per session was β_0*=2.89* (*SE=0.14*). Thirty-five point one (35.1%) of variance in craving was attributed to between participants differences (2nd level), while 64.9% was attributed to between sessions differences within the same participant (1st level).

4.4.2.2.2 Association of craving with AB and intention to drink.

The research question was whether craving would be higher on days when participants have pre-determined plans to consume alcohol, and when AB is also high. A 3-levels (sessions within days within participants) regression model was fitted to predict craving from AB and intention to drink as first- and second-level explanatory variables, respectively. The model predicted 3.38% of variance in craving scores at the session level, 30.15% at the day level, and 7.49% at the participant level. Intention to drink ($\beta = 1.12$*, (SE) = 0.15, p < .001, n = 1011 sessions*), but not AB ($\beta < 0.001$*, (SE) = 0.001, p = 1.00, n = 1011 sessions*), was positively associated with craving. Craving was higher in sessions that belonged to a day with a declared intention to drink, compared to days with no intention to drink. However, sessions that measured high in AB did not necessarily measure high in craving. This shows that craving was affected by participants' expectancy to consume alcohol that day, but not by their AB.

4.4.2.2.3 Fluctuations in craving over the course of a drinking day

The research question was whether craving will increase over the course of a drinking day, as a function of its proximity to initiation of drinking. To test this, only days when alcohol was consumed were included in the analysis. A 2-levels regression model was fitted to predict craving from period of testing (morning, early afternoon, or late afternoon). The model was fitted three times. The first time, late-afternoon sessions were excluded from the analysis, so that the dummy variable in the regression model would represent the comparison between morning and early-afternoon sessions. Morning sessions were excluded the second time to compare early afternoon with late afternoon, and early-afternoon sessions were excluded the third time to compare morning with late afternoon. P values were corrected for multiple comparisons.

The first model (morning vs early afternoon) explained 56.9% of craving in the participant level and 14.8% in the session level. Craving was higher in the early afternoon than in the morning *(β = 0.69, SE = 0.19, p = .001, n = 380 sessions).*

The second model (early afternoon vs late afternoon) explained 13.35% of variance in craving in the participant level. However, in the session level, it did not predict any additional variance in craving, compared to an empty model with no explanatory variables. Craving was higher in the late afternoon than in the early afternoon (*β = 1.24, SE = 0.23, p < .001, n = 380 sessions*).

The third model (morning vs late afternoon) explained 24.3% of variance in craving in the participant level. However, it did not explain any further variance in craving in the session level, compared to an empty model with no explanatory variables. Craving was higher in the late afternoon than in the morning (*β = 1.97, SE = 0.21, p < .001, n = 368 sessions*).

Overall, results indicate that craving was gradually increasing over the course of a drinking day and peaked when the assessment was most proximal to alcohol use (Figure 4.3; changes in craving over the course of the day in non-drinking days are also presented for comparison).

4.4.2.3 Attentional bias

4.4.2.3.1 Selection of optimal data structure

When AB was examined within all days (drinking and non-drinking), a 3-levels structure did not fit the data better compared to a ‘sessions within days’ 2-levels structure or a ‘sessions within participants’ 2-levels structure *(all ps > .05)*. Also, none of the 2-levels structures fitted the data better than a single-level structure (*all ps > .05*). Therefore, a single-level structure was retained as best fitting the data and regular single-level analysis was performed.

When AB was examined within drinking days only, pattern of results remained the same. Therefore, in this case too, a single-level structure was retained as best fitting the data and a regular single-level analysis was performed.

4.4.2.3.2 Presence of AB over the course of the day

The research question was whether there was AB for alcohol-related cues across all testing periods (morning, early afternoon, and late afternoon). Due to violation of parametric assumptions, a bootstrapped one-sample t-test was used to examine whether average AB score was significantly different from 0. Average AB score was not different from 0 in the morning *t (370) = -0.72, p=.48, d = -.04, n = 371 sessions*) and early-afternoon sessions (*t (365) = -1.70, p=.09, d = -.09, n = 366 sessions*). Average AB was significantly different from 0 in the late-afternoon sessions (*t (352) = -2.40, p=.021, d = -.13, n = 353 sessions*) although the attentional bias was away from, rather than towards, alcohol. Taken together, these results indicate that there was no AB for alcohol in any assessment over the course of the day. In the late-afternoon assessment there was significant AB for soft-drinks cues.

4.4.2.3.3 Association of AB with intention to drink

The research question was whether AB would be higher on days when participants have pre-determined plans to consume alcohol. Due to violation of parametric assumptions, the Mann-Whitney U test was used to compare days when intention to drink was declared with days when no intention was declared. No difference in magnitude of AB was found between groups of days (*Z = -1.27, p = .20, n = 1017 sessions*). Results indicate that AB was not higher in sessions which belonged to intention to drink-days, compared to no-intention days.

4.4.2.3.4 Fluctuations in AB over the course of a drinking day

The research question was whether AB will increase over the course of a drinking day, as a function of its proximity to initiation of drinking. To test that, only days when alcohol was consumed were included in the analysis. Due to violation of parametric assumptions, a Kruskal-Wallis test was used to examine the differences in magnitude of AB across testing periods (morning, early afternoon, and late afternoon). Analysis revealed no differences (χ^2 *(2) = 0.82, p = .66, n = 1090 sessions*). Results indicate the AB remained stable over the course of a drinking day (Figure 4.4; changes in AB over the course of the day for non-drinking days are also presented for comparison).

4.4.2.4 Initiation of alcohol consumption

4.4.2.4.1 Selection of optimal data structure

A 2-levels structure (sessions within days) did not fit the data better than a single-level structure (χ^2 *= 0.16, df = 1, p =1*). Therefore, a single-level structure was retained as best fitting the data and regular single-level analysis was performed.

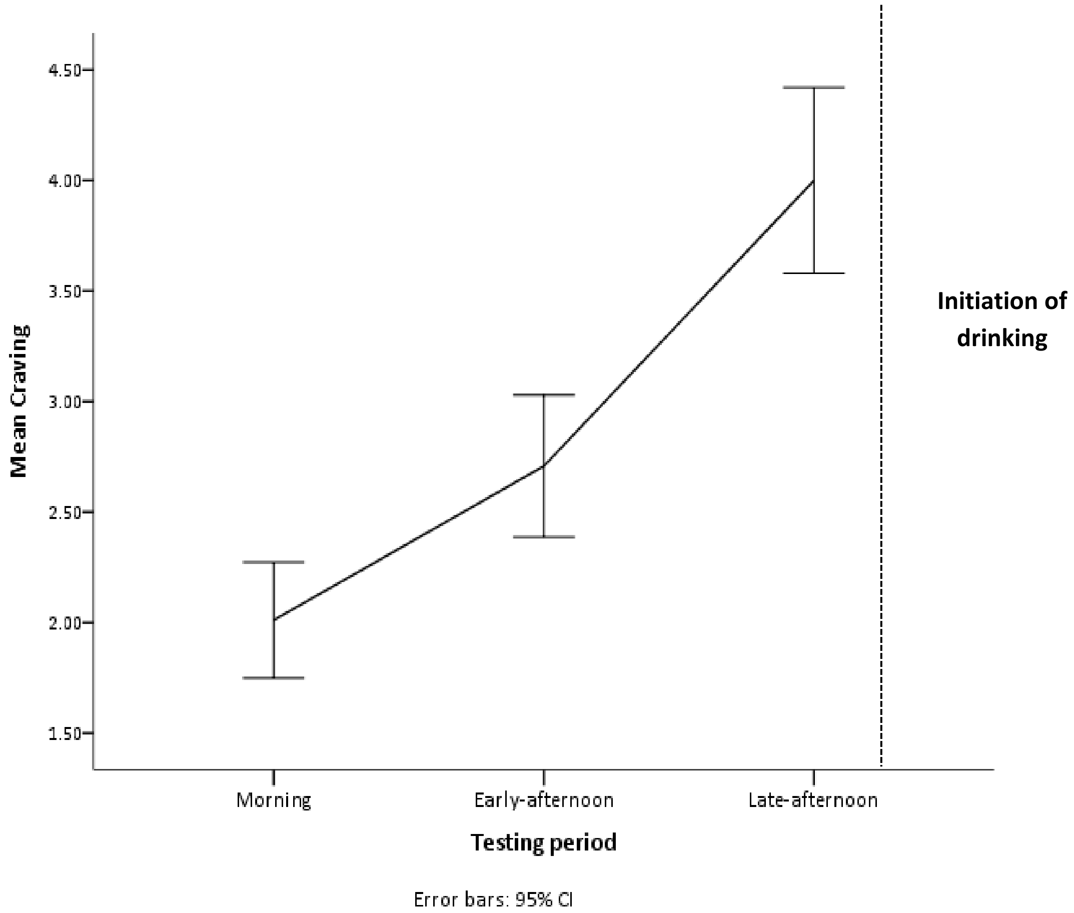
4.4.2.4.2 Prediction of initiation of alcohol by AB and craving

The research question was whether measures of AB and craving in different periods of testing (morning, early afternoon, and late afternoon) would be associated with initiation of drinking in the evening. A ridge logistic regression (ridge parameters derived through the method described by Cule & De Iorio 2013) was used to test a model that predicted initiation of drinking with the three AB and craving measures as predictors. Variables were measured repeatedly within a short period of time and within the same individuals. For that reason, a Ridge regression was adopted, as opposed to a standard binary logistic regression, due to concerns over multicollinearity (see Månsson & Shukur 2011; Shen & Gao 2008). For example, correlations between craving measures were particularly high *(.46 to .62)*.

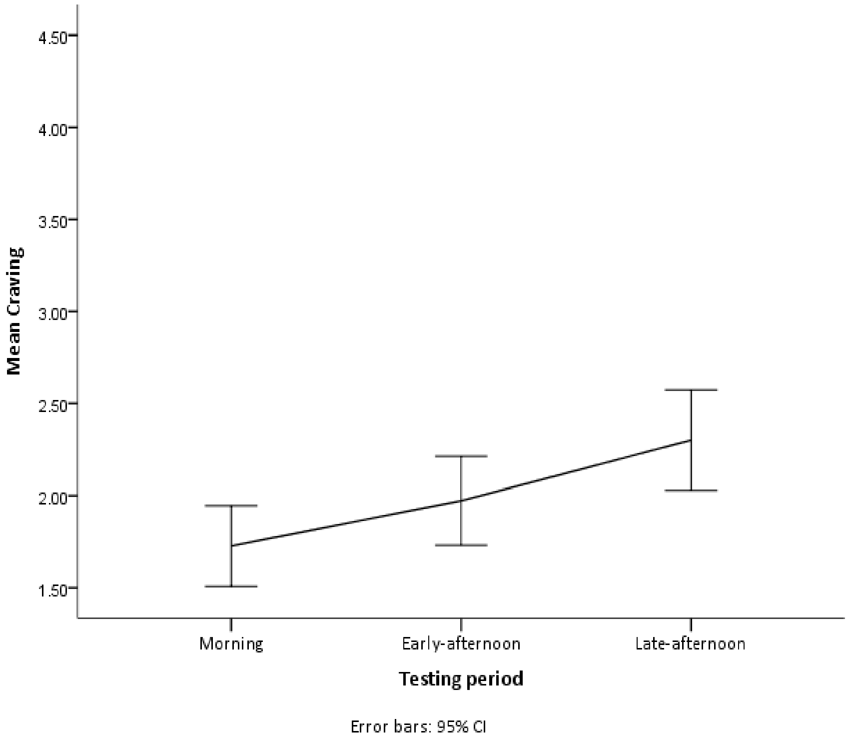
Results showed a significant ridge logistic regression model (*multiple* $R^2 = 0.11$, $p < .001$, $n = 284$ *days*). AB did not predict initiation of drinking, regardless of when it was measured within the day (morning: *scaled* $B = 0.29$, $SE=0.35$, $p =.40$; early afternoon: $B = - 0.07$, $SE=0.35$, $p =.84$; late afternoon: $B=0.04$, $SE=0.35$, $p=.90$). Regarding craving, morning ($B = -0.43$, $SE= 0.35$, $p =.22$) and early afternoon craving ($B = 0.42$, $SE= 0.35$, $p =.24$) did not predict initiation of drinking. However, late afternoon craving was positively associated with initiation of drinking ($B= 1.83$, $SE=0.35$, $p < .001$). Results indicate that, over the course of a day, measures of craving, but not AB, predicted alcohol use in the short-term, but not mid- to long-term. When intention to drink was added as a predictor in the model, an unexpected result was that morning craving was negatively associated with initiation of drinking ($B= -0.96$, $SE=0.46$, $p = .004$). For all other variables, pattern of results remained the same.

Figure 4.3. Craving (1-11 scale) over the course of drinking days (panel one) and non-drinking days (panel two) - Study two

Panel 1



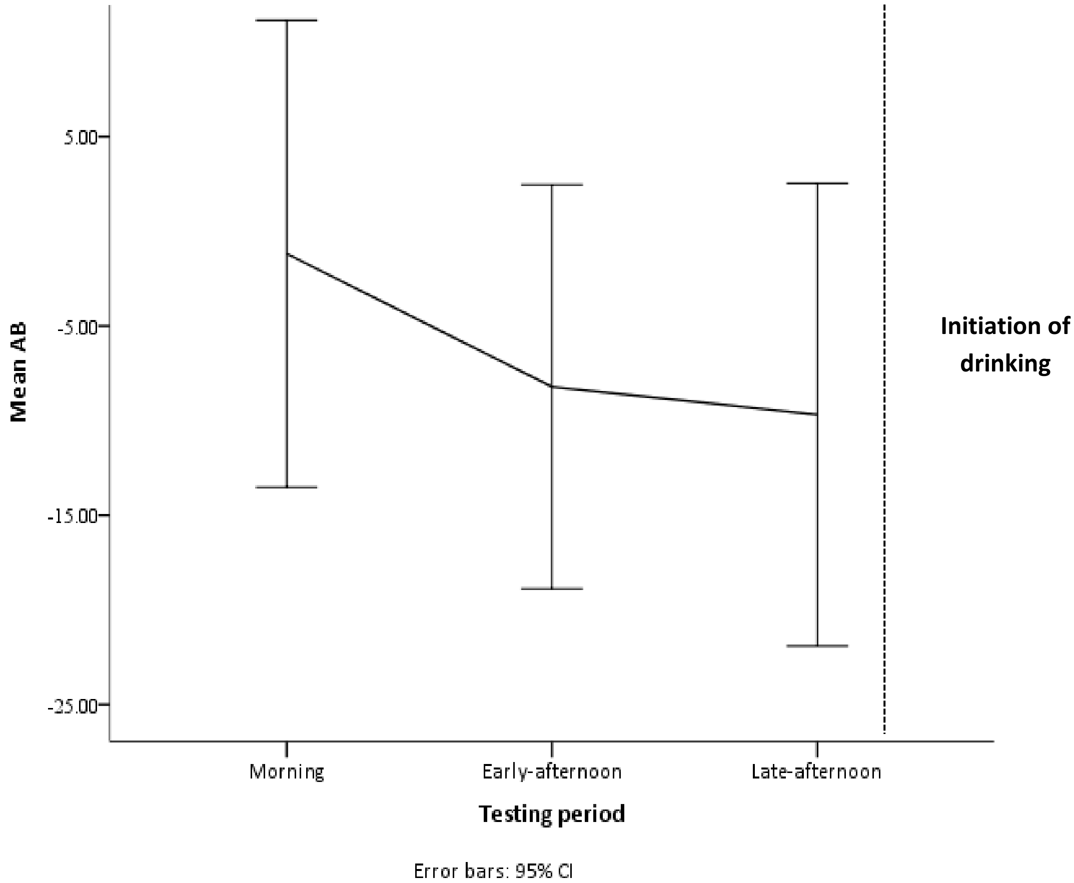
Panel 2



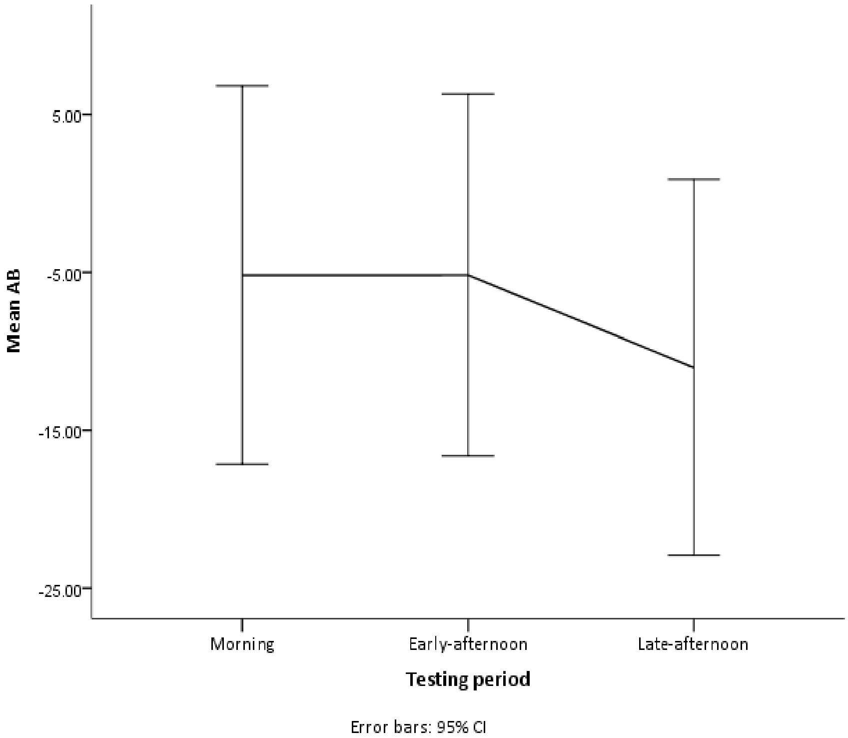
4.50
4.00
3.50
3.00
2.50
2.00
1.50
Mean Craving
Morning
Early-afternoon
Late-afternoon
Testing period
Error bars: 95% CI

Figure 4.4. AB (ms) over the course of drinking days (panel one) and non-drinking days (panel two) - Study two

Panel 1



Panel 2



5.00
-5.00
-15.00
-25.00
Mean AB
Morning
Early-afternoon
Late-afternoon
Testing period
Error bars: 95% CI

4.4.3 Interim summary of findings

Study two used EMA to examine AB (measured with a personalized pictures Stroop-task), craving, intention to drink, and alcohol consumption, over seven days. It should be noted that although participants did not reduce their alcohol consumption over the EMA week (as would be expected if there was a reactivity effect to self-monitoring drinking behaviour), they registered higher volumes of alcohol consumption in the EMA diaries compared to the TLFB questionnaire prior to the EMA period. However, this might indicate that quantity of alcohol consumption was registered more accurately over the EMA period compared to the TLFB (daily reports as opposed to retrospective reports for the last 14 days), rather than that participants altered their typical drinking patterns over the EMA period.

Findings showed that global measures of AB were not associated with global measures of craving. They also showed that craving, but not AB, was higher on sessions that belonged to days with declared plans to consume alcohol, compared to days with no plans. Moreover, over the course of a drinking day, craving, but not AB, was gradually increasing, peaking in the late afternoon (the assessment that was most proximal to initiation of drinking). Finally, craving predicted whether participants will choose to consume alcohol or not, only when it was measured proximally to initiation of drinking (late-afternoon assessment). These findings replicated study one. However, in contrast to study one, study two did not find an association between any measure of AB (morning, early afternoon, or late afternoon) and initiation of drinking.

4.5 Discussion

Both studies in this chapter utilized EMA to a) explore the relationship between global measures (measures taken across the whole day) of AB and craving, b) explore the effect of intention to drink on AB and craving, c) examine how AB and craving fluctuated over the course of a drinking day, and c) test the extent to which AB and craving would be better predictors of initiation of drinking when measured proximally to alcohol use. In study one AB was examined with a classic alcohol-worlds alcohol-Stroop task, while study two utilized a personalized pictures of alcohol version.

4.5.1 AB

Regarding AB, the main purpose of the two studies was to provide evidence that would support the state hypothesis of AB. This hypothesis would predict that AB fluctuates under the influence of the current motivation to drink alcohol and that AB would show only short-term associations with alcohol use. Although findings in study one provided partial support to the hypothesis, finding in study two did not support the hypothesis.

Both studies in this chapter showed no association between AB and craving. This is in contrast to what was hypothesized, as well as to previous findings in the lab (for a meta-analysis see Field et al. 2009) and the real world (Waters, Szeto, Wetter, Cinciripini, Robinson, & Li 2014).

The meta-analysis by Field et al. (2009) on the association between AB and craving might provide some context for the current findings. The meta-analysis revealed an existing, but weak relationship between AB and craving. It also showed that this relationship tends to become stronger for illicit substances (e.g. cannabis, cocaine, heroin), when there is strong craving, and when AB is inferred by direct measurements of attention (e.g. eye-movement tracking instead of reaction-time tasks). The two studies in this chapter examined AB and craving for alcohol, measured AB indirectly (manual reaction-times) and participants generally showed low levels of craving. In future studies, it would be worth using mobile eye-tracking to measure AB in the real world, as well as conducting EMA studies in clinical populations, where craving would be expected to be generally higher.

Also, the real-world study which found an association between AB and craving (Waters et al. 2014) was conducted in smokers attempting to quit, as opposed to social drinkers not motivated to reduce their drinking in the present study. Although it has been found that treatment-seeking status does not moderate the association between AB and craving (Field et al. 2009), the presence of the association in Waters et al. (2014) but not here might be explained by differences in the intensity of craving in the two samples. As showed by the meta-analysis discussed earlier, the association between AB and craving is stronger when craving is high. Stronger craving by smokers attempting to quit, compared to drinkers not motivated to reduce their drinking, would be consistent with predictions of the cognitive process model of addiction (Tiffany & Conklin 2000; see section 1.4.3 in chapter one), which suggests that craving occurs as a response to any interruptions to the substance using behaviour. It would be reasonable to expect that the more intensive and more frequent the attempts to interrupt the substance use behaviour are (as when attempting to quit smoking), the stronger the craving generated by these attempts might be. Although participants in the present study also showed some momentary increased in craving over the course of the day, their average craving remained at medium levels. Taken together, the association between AB and craving might become more apparent in moments of intense craving and the participants in this study might not have reached that level of intensity, despite the momentary increases over the course of the day.

The sample in the studies in this chapter was social drinkers not seeking to reduce their alcohol consumption. Participants trying to abstain from the substance might generally experience stronger cravings compared to participants who can consume the substance freely and, therefore, their craving might be more strongly associated with AB. Indeed, the meta-analysis discussed above (Field, Munafò, & Franken 2009; although results were based in laboratory studies) showed that the association between AB and craving was stronger when craving was higher.

The lack of association between intention to drink alcohol and AB, found in both studies, contradicts previous evidence which show that AB is elevated when participants expect to receive the substance (for a review see Jędras et al. 2013). However, Werthmann, Roefs, Nederkoorn, and Jansen (2013) manipulated the level to which participants perceived chocolate as available and found no effect on AB for chocolate-related cues.

One explanation for the present findings might be that groups of days (intention and no-intention to drink) were compared in terms of global AB, averaged across the three time-periods of the day (morning, early afternoon, and late afternoon). However, participants have been instructed to abstain from alcohol until the late-afternoon assessment, so those who had plans to consume alcohol might have started to anticipate alcohol only after the late-afternoon assessment. However, post-hoc analyses in both studies, where comparison of groups of days (intention and no-intention) was restricted to late-afternoon AB assessments, revealed no differences (Kruskal-Wallis test, $ps > .05$).

Opposite to expectations, AB remained stable over the course of a drinking day, in both studies. The assumed mechanism behind elevation of AB when alcohol use approached, would be parallel increases in craving, perceived availability of the substance, and incentive value of alcohol (all of which have been found to enhance AB; craving: for a meta-analysis see Field et al. 2009; expectation: for a review see Jędras et al. 2013; incentive value: Rose et al. 2013; Rose et al. 2018). However, the present studies were not able to test this mechanism fully, as they measured increases in craving (see section 4.5.2.2) but not in perceived availability and incentive value. Future studies could use EMA to examine whether perceived availability and incentive value of alcohol are elevated as alcohol use approaches, and whether this is associated with parallel increases in AB.

Regarding the predictive ability of AB, the two studies revealed different results. Results from study one supported the state hypothesis of AB, showing that AB predicted initiation of drinking when measured proximally to alcohol use (late afternoon assessments), but not when measured earlier over the course of the day (morning and early afternoon). However, study two showed no association between any measure of AB and initiation of drinking.

Null findings in the second study might be attributed to the stimuli used in the alcohol Stroop-task (pictures of beer and soft-drinks). Time taken to react to the control stimuli (pictures of soft drinks) might have been inflated by the automatic attraction of attention by appetitive stimuli (see post-hoc analysis in study two, chapter three) and by the strong pre-existing colour associations of some of the stimuli (e.g. a Seven-Up bottle always brings "green" in mind; see Cox et al. 2006). In the alcohol-Stroop, AB is operationalized as the level of interference caused by alcohol-related stimuli. If the task has become less sensitive to that interference, it is not surprising that the study failed to detect any associations between AB and alcohol use behaviour. However, regarding using soft-drinks as control stimuli, one could argue that losses in magnitude of difference in interference are balanced by purer measures of AB towards alcohol, as the task controls for the appetitive nature of stimuli.

4.5.2 Craving

As expected, craving was higher on days when intention to drink was declared. This adds to the mounting evidence suggesting that anticipation of the substance enhances cue-elicited urges to consume the substance (for reviews see Jędras et al. 2013 and Wertz & Sayette 2001). This finding also agrees with a naturalistic study (Dar, Stronguin, Marouani, Krupsky, & Frenk 2005) where researchers asked Orthodox Jewish heavy smokers to complete craving measures in their homes, at various points during the day, in three occasions; a normal working-day, a working-day when participants have been instructed to abstain and the Sabbath, when participants habitually abstain from smoking for religious reasons. Craving in the Sabbath, when participants had no plans to smoke, was significantly lower, compared to normal working days when opportunity to smoke was present (although there was no difference between normal and abstaining working-days).

Participants in the present studies were instructed not to consume alcohol before the late afternoon and their craving was found to gradually increase over the course of the day, peaking in the late-afternoon assessment, which was the most proximal assessment to initiation of drinking. This supports the hypothesis and is consistent with findings from a naturalistic study with flight attendants (Dar, Rosen-Korakin, Shapira, Gottlieb, & Frenk 2010). Flight attendants are not allowed to smoke during a flight. When their craving was assessed over the course of a flight, it was found to gradually increase, peaking prior to landing, which was the most proximal assessment to the opportunity to smoke. Most importantly, comparisons across flights of different durations showed that increases in craving were a function of proximity to the opportunity to smoke (landing), rather than duration of being abstinent.

In line with what was hypothesized, craving predicted initiation of drinking when measured proximally to alcohol use (late-afternoon assessments), but not when measured earlier over the course of the day (morning and early-afternoon assessments). Past research (Fatseas et al. 2015) has identified comparable patterns, where measures of craving predicted substance use in the short- to mid-term (measured on average four hours later) but not in the long-term (measured on average eight hours later). Also, although participants in the present studies where not seeking to abstain from alcohol, results resemble those of studies examining relapse after a period of abstinence. Serre et al. (2015) reviewed EMA studies that examined the association between craving and relapse for a range of substances (e.g alcohol, tobacco, cocaine, etc) and found that craving was a better predictor of substance use when measured proximally to substance use.

4.5.3 Synthesis of findings

Taken together, the overall results regarding AB and craving reveal a picture more complex to what was initially hypothesized. The theoretical rationale of the present studies was that AB would be a better predictor of initiation of drinking when measured in close temporal proximity to alcohol use, because at that point AB would be at its highest, driven up by parallel increases in craving. Indeed, AB predicted initiation of drinking when measured in the late afternoon, when craving was at its highest. Yet, there was no elevation of AB in the late afternoon, compared to earlier assessments. This pattern is in partial agreement with another EMA study examining AB, craving and substance use (Marhe et al. 2013), although that study examined relapse of drug dependent patients. Marhe et al. (2013) found that AB predicted relapse only in the short-term and only when measured on moments of temptation to use drugs. However, they also found that AB on temptation moments was higher compared to moments of no-temptation.

This might suggest that AB is a better predictor of alcohol use when other motivational variables (e.g. high craving) operate at the same time, without AB necessarily rising to higher levels. That would mean that AB measures taken over the course of the same day, but in different motivational contexts, might differ in their predictive ability of alcohol use, even if they are of comparable magnitudes. The implications of this would be that programs aiming to predict imminent alcohol consumption should not necessarily search for momentary peaks of AB but for momentary co-occurrence of AB and high craving. Future studies could ask participants to complete AB assessments every time they feel an urge to drink alcohol and examine the extent to which those measures would predict alcohol use, in comparison to AB measured in no-urge moments (this question is address in chapter five).

However, a post-hoc analysis in the present studies did not support this assumption. AB and craving were averaged across the three testing periods (morning, early-afternoon, late-afternoon) to calculate a global AB and craving measure for the day. A ridge logistic regression examined whether initiation of drinking would be predicted by global AB and craving, and their interaction. In study one, initiation of drinking was significantly predicted by global AB and craving (positive prospective associations; $p = .04$ and $p < .001$, respectively) but not their interaction ($p = .524$). In study two, initiation of drinking was significantly predicted by global craving (positive prospective association; $p < .001$) but not from global AB or its interaction with craving ($p = .758$ and $p = .980$, respectively).

4.5.4 Limitations

The present studies have some weaknesses. First, in these studies it was not possible to objectively assess whether participants complied to the restriction not to consume alcohol before the late-afternoon assessment. Efforts to increase compliance included to stress out in the initial visit that this restriction was a necessary and obligatory element of the study, and to allow completion of the late-afternoon assessments from as early as 16:30. However, future studies could explore the option of equipping participants with breathalysers or other wearable devices (see Mahmud et al. 2018 for examples of available options) that would be able to electronically register and timestamp measures of biomarkers of drinking, in order to ensure abstinence from alcohol prior to every assessment (although that could be financially and practically demanding).

Second, anticipation of alcohol was inferred from participants self-reports regarding their alcohol-drinking plans for the day. However, this plan was declared in the morning assessment and might have changed over the course of the day. Days that were categorized as alcohol-anticipation days, might have changed to no alcohol-anticipation after the morning assessment and vice versa. Future studies could monitor participants' plans throughout the day to obtain a more detailed picture of alcohol anticipation.

Finally, different time-periods of the day (morning, early afternoon, and late afternoon) were used as a proxy for different contexts which would have different effects on the motivation to drink alcohol. For example, one might encounter less alcohol-related cues and perceive alcohol as less available when they go to work in the morning, compared to the end of the day in the late afternoon. However, the present studies did not directly test any variations in such contextual factors. Future studies could examine whether contextual factors that are known to affect AB and craving vary over the course of the day.

4.5.5 Conclusion

In conclusion, this chapter utilized EMA to explore the state-motivational nature of AB in the real world, and its association with alcohol use. This chapter provides no evidence that AB is a state-like variable. However, results indicate that AB might be able to predict alcohol use only when measured proximally to alcohol use, although lack of replication in the second study suggests caution. The chapter also demonstrates that measures of craving fluctuate momentarily and they are a better predictor of alcohol use when measured proximally to alcohol use. Both studies provide support for the importance of incorporating EMA and smart-phone applications in the study of alcohol attentional bias and craving.

Chapter five: EMA study three

5. Close encounters 2: Variations in AB and craving, and their association with alcohol consumption, in real-world contexts of high motivation to drink alcohol.

5.1 Abstract

Chapter four, study one, demonstrated that measures of AB taken in the late afternoon, which was the assessment most proximal to alcohol use, predicted initiation of drinking, although AB did not increase in magnitude prior to initiation of drinking. The current study expands on this by examining AB, craving, and their association with alcohol use, in two additional motivational contexts. AB and craving was measured when participants felt tempted to consume alcohol (temptation assessments - TA) and immediately prior to when participant served themselves a drink (total availability assessments - TAvA). Sixty social drinkers were lent a smartphone for one week and completed measures of craving (single-item question) and AB (alcohol-Stoop task), and they also registered quantity of alcohol consumption in a digital diary. Measures of AB and craving were taken at fixed time-windows (control assessments - CA), and in TA and TAvA. Results showed that craving, but not AB, was higher in both TA and TAvA, compared to CA. In CA, none of the variables predicted alcohol quantity or initiation of drinking. In TAvA, craving, but not AB, predicted quantity of alcohol consumption. In TA, none of the variables predicted quantity or initiation of drinking. In conclusion, the motivational context in which measures were taken, had an impact on craving and its association with alcohol use. However, motivational context had no effect on either AB or its association with alcohol use.

5.2 Introduction

As it has been discussed previously in the thesis (chapter one, section 1.7), AB shows inconsistent prospective associations with alcohol use, both in non-dependent drinkers (e.g. Christiansen, Mansfield, et al. 2015; Field et al. 2004; Groefsema et al. 2016) and clinical samples (for a review see Christiansen, Schoenmakers, et al. 2015). An explanation that has been provided for the inconsistent findings is the state-hypothesis of AB (Field et al. 2014; Field et al. 2016), which suggests that magnitude of AB, as well as its association with alcohol use, is dependent on the motivational context in which measures are taken.

However, chapter four have provided only partial support to the state-hypothesis. Study one demonstrated that when AB was measured in high temporal proximity to alcohol use, when craving was also high, it predicted initiation of drinking, without increasing in magnitude. This might suggest that it is the strength of the association between AB and alcohol use, rather that the magnitude of AB itself, that fluctuates under the influence of the motivational context.

The present study examined AB and craving, and their association with alcohol use, within two specific situations when motivation to drink was expected to be high, rather than just at different times of day, like in the previous studies in the thesis. The first situation was when participants experienced a strong urge to drink alcohol, representing a moment of temptation to drink (temptation assessments - TA).

Previous naturalistic studies (Marhe et al. 2013; Waters et al. 2012) have shown that AB and craving in drug-dependent patients were higher in moments of temptation, compared to random times over the course of the day. They also found that AB was associated with relapse when measured in moments on temptation but not when measured in random times. Surprisingly, craving did not predict relapse. In another study (Jones, Tiplady, et al. 2018), craving in heavy drinkers motivated to abstain has also been found to increase in moments of temptation in the real world. However, although craving predicted alcohol consumption when measured in random moments, craving measured in moments of temptation explained no additional variance in alcohol consumption.

The second motivational situation was when participants were about to serve themselves an alcoholic drink and consume it. This represented a moment with high proximity to alcohol consumption, when alcohol was physically present and imminently available for consumption (total availability assessments - TAvA).

Despite the lack of evidence regarding the effect of these motivational features on AB in the real world, several laboratory studies have demonstrated that AB was increased when substance-related cues were physically present (in vivo exposure to alcohol; Cox et al. 2003; Ramirez et al. 2015) and when participants expected that the substance would be available for consumption (for a review see Jędras et al. 2013).

Regarding craving, previous naturalistic studies have demonstrated that craving is a better predictor of substance use when measured proximally to substance use and when the substance is present. For example, Fatseas et al. (2015) measured craving in participants dependent to a variety of substances (alcohol, tobacco, opioids, and cannabis) and found that craving was associated with substance use in the short- to mid-term (up to four hours later), but not in the long-term (eight hours later). A meta-analysis of EMA studies (Serre et al. 2015) showed that craving was a better predictor of relapse after a period of abstinence when measured most proximally to substance use. Also, Dulin and Gonzalez (2017) examined the effect of various craving-triggers (such as time of day, other people drinking, and being around alcohol) on the association between craving and alcohol consumption and found that one of the strongest associations between craving and alcohol use was detected when craving was triggered by the presence of alcohol.

Moreover, studies conducted in the real world have shown that craving is higher when the substance is considered available for consumption. For example, Dar et al. (2010) measured smoking craving in flight attendants over the course of a non-smoking flight and found that craving peaked close to landing, when cigarettes would become available again. Similarly, Dar et al. (2005) measured craving in orthodox Jewish in normal weekdays, weekdays when participants were asked to abstain, and on a day when participants abstained for religious reasons (the Sabbath), and found that participants had the lowest levels of craving on the Sabbath, when cigarettes were considered as generally not available.

In the current study AB and craving where measured over the course of seven days in fixed time-windows (control assessments; CA), as well in specific events of interest (TA and TAvA). It was hypothesized that global measures (across the whole day) of AB will be associated with global measures of craving, and that AB and craving will be higher in TA and TAvA, compared to CA. Finally, it was also hypothesised that AB and craving will predict alcohol use when measured in TA and TAvA, but not in CA.

5.3 Methods

5.3.1 Participants (table 5.1)

Participants were recruited from the University of Liverpool and local community, via intranet advertisement and word of mouth. To be eligible, participants should have been fluent English speakers and regular social drinkers (self-reporting alcohol consumption at least once per week). They were excluded if they had any current or previous alcohol abuse or dependence diagnosis, were pregnant or breastfeeding, or suffered from any form of colour-blindness (all assessed with verbal and written confirmation provided by participants). All participants provided informed consent prior to participation and the study was approved by the University of Liverpool Ethics committee.

Seventy participants were initially recruited but ten were excluded, leaving 60 participants that were included in the analysis (53 females, average age 25.81 (±9.15) years old). Reasons for exclusion of the ten participants were as follows: Data from one participant were lost due to technical failure of the smartphone. One participant withdrew from the study before completing any assessments. Data from two participants were discarded because their alcohol consumption diary entries were not available. Data from three participants were discarded as they did not comprehend the instructions properly (e.g. mislabelling sessions or completing the wrong task). Data from three participants were discarded as they provided insufficient amount of data (less than 50% completion rate, see section 5.3.5 for definition of completion).

5.3.2 Materials

5.3.2.1 Smart-phones

An Alcatel Pixi 3 or Samsung Galaxy Ace 3 GT-S7275R was loaned to participants for the daily assessments over the EMA week.

5.3.2.2 Alcohol Stroop-task stimuli

Stimuli in the alcohol Stroop-task were words generally related to alcohol or to environmental features. For more details see chapter 3, section 3.3.2.3. Stimuli were presented to participants in a blocked format; a block of alcohol-related stimuli followed by a block of control stimuli, and vice-versa. Order of blocks was counterbalanced across sessions.

5.3.2.3 Training to recognize moments of temptation

As participants were required to recognize moments of temptation, a brief training session on how to recognise alcohol craving was provided, before the EMA phase commenced. Participant watched a brief automated presentation, based on Cognitive Behavioural Therapy (CBT) manuals for treating addiction (Carroll 1998; Kadden, Caroll, Donovan, Cooney, Monti, Abrams et al. 1995), which explained what craving feels like (physical/somatic and psychological/cognitive expressions) and what could trigger craving (e.g. exposure to people, places, and situations that are associated with drinking). The presentation also provided practical examples (e.g. “You had a very tiring and stressful day and you finally got back home. As you take your shoes off and lay on the couch, you start thinking that a drink would be great right now”).

5.3.3 Measures

5.3.3.1 Baseline measures

The *Time Line Follow Back (TLFB; Sobell & Sobell 1990)* and the *Alcohol Use Disorders Identification Test (AUDIT; Saunders et al. 1993)* were used to assess typical alcohol consumption and hazardous drinking habits, respectively. For a detailed description see chapter two, sections 2.1.1 and 2.1.3.

5.3.3.2 EMA measures (table 5.2)

A variation of the Alcohol-Stroop App (AlSAp) was used to measure AB and craving. For a detailed description of AlSAp see chapter two, section 2.2. The version used in the present study was slightly different to the one described in chapter two, section 2.2 in that at the beginning of each AlSAp session participants were asked why they have initiated this session, choosing one of the three options; 1. This is a fixed-assessment (denoting a control assessment; CA), 2. I experienced craving to drink alcohol (denoting a temptation assessment; TA), 3. I am about to drink alcohol (denoting a total availability assessment; TAvA). Also, the intention to drink question, described in chapter two, section 2.2., was not included here.

As participants were not trying to reduce their alcohol consumption, it was likely that moments of temptation might lead to alcohol consumption in some cases. For that reason, at the end of TA, participants were asked if they are going to have an alcohol drink right now (yes or no). Finally, a digital diary was used to assess quantity of alcohol consumption (in units) each day over the EMA period (see chapter two, section 2.1.2).

5.3.4 Procedure

Interested participants visited a laboratory in University of Liverpool where they were screened for eligibility and provided informed consent. Eligible participants completed a battery of questionnaires regarding their alcohol use behaviour (TLFB, AUDIT) and the brief training on craving recognition. Training for using the app was provided by asking participants to access the app on the smartphone, complete a full assessment, and ask any questions. Finally, participants were given a smart-phone with the AlSAp installed and had the study protocol (table 5.3) explained to them

Participants were instructed to start using the app the following day and for the next seven days, up to five times per day. Two control assessments (CA) were required to be completed within pre-specified time windows; one assessment from 10:30 to 12:30 and another one from 13:30 to 15:30. Participants could complete a CA any time within these time-windows. The number of CA was set at two per day, so that in cases where the first one was missed, participants would have another opportunity to provide control measures. Also, CA were not taken later than 15:30, as there were more chances for participants to choose to drink alcohol after that time (no CA should be completed after consuming alcohol, see description of TAvA below).

Participants were also asked to access the app when they felt tempted to consume alcohol. To make number of assessments more manageable, and to increase adherence to the protocol, participants were told to complete a maximum of two TA per day.

Finally, participants were asked to access the app just when they were about to serve themselves the first alcoholic drink of the day to consume it. After having their first drink, participants were asked to stop using the app. Participants were explicitly instructed that no assessments should be initiated after they had their first drink, even if they have another drink later, experience their first or second temptation for the day, or have an uncompleted CA on pending. This was to avoid their performance being affected by possible intoxication (Field et al. 2010). At the end of each day or first thing the next morning, participants reported on the digital diary the number of alcohol units they have consumed.

No prompts to remind participants to access the app were provided. Participants were provided with a digital manual, saved on the smartphone, which outlined the assessments schedule and they were instructed to consult it every time they felt unsure. In the event of any technical failure (e.g. the app crashing half-way a session), participants were instructed to repeat the session immediately until completed. In the event of persisting problems, they were asked to return the faulty phone, receive a replacement and continue the study until seven days of testing was completed. One phone (1.6%) was returned due to battery charging problems. A new charger was given to the participant who then continued with the study as normal.

Upon completion of the EMA phase, participants returned the phone, received a thorough debriefing, and were given compensation in the form of either course credits or shopping vouchers (worth £20). Participants received full compensation for at least 75% completion, reduced compensation for at least 50% completion and no compensation for less than 50% (see section 5.4.1 for compliance results). As it was impossible to assess whether participants might have experienced temptation but neglected to access the app, completion rates were based on CA (should have two) and TAvA (should have one for every day that alcohol consumption was evident on the diary).

5.3.5 Data reduction and analysis

AB scores were calculated by subtracting reaction times (RT) for control stimuli from RT for alcohol stimuli. For a detailed description of AB scores calculation see chapter two, section 2.3.1. Positive AB scores are indicative of AB towards alcohol-related stimuli.

In previous EMA studies in the thesis, alcohol consumption was analysed as a dichotomous variable (initiation of drinking), coded as no ('0') or any alcohol consumption ('1': rationale of this choice is described in chapter four, section 4.3.1.7). However, as TAvA happened only on drinking days, AB and craving measures taken in TAvA were regressed on to quantity of alcohol consumption (number of units consumed). AB and craving measures taken in TA were regressed on to quantity of alcohol consumption and initiation of drinking.

Data had the same hierarchical structure as in chapter four and the same analytic strategy was used (see chapter four, section 4.3.1.7). For a detailed explanation of multilevel analysis and hierarchical data structures see chapter two, section 2.3.2.

Table 5.1. Participants characteristics

Characteristic	Mean (SD) or Freqency (%)
Age	25.81 (±9.15)
Gender	53 females (88.3%)
Alcohol Use Disorder (AUDIT)[a] score	8.55 (±3.39)
Quantity of drinking in the week before the study (in units)	18.29 (±15.12)
Quantity of drinking in the EMA week (in units)	18.41 (±12.85)
Number of alcohol consuming days in the EMA week	2.88 (±1.32)

a. Cut of score for hazardous drinking = 8.

Table 5.2. Daily measures over the EMA week

Measure	Mean (SD) or Frequency (%)
Days when alcohol was consumed	173 (41.2%)[a]
Quantity of alcohol consumed per day (in units)	2.62 (±4.35)
Assessments	
Control assessments (CA)	703 (73.9%)[b]
Temptation assessments (TA)	101 (10.6%)[b]
Total availability assessments (TAvA)	147 (15.5%)[b]
Attentional Bias scores (ms)	
CA	55.76 (±77.24)
TA	62.08 (±74.11)
TAvA	50.99 (±76.69)
Craving score[c]	
CA	2.17 (±1.92)
TA	6.20 (±2.44)
TAvA	7.61 (±2.61)

a. Percentage is out of 419 days for which data on alchol consumption were provided.
b. Percentage is out of 951 assessments completed in total.
c. Scored on a scale from 1 (no craving) to 11 (very strong craving).

Table 5.3. Type and number of AB and craving assessments over the EMA period

Type of assessment	Name of assessment	Trigger of completion	Number required per day	Restrictions
Time-related	Control assessments (CA)	Time	Two	Should happen within the provided time-windows (10:30 -12:30 and 13:30-15:30)
Event-related	Temptation assessments (TA)	Temptation/Urge/Craving to drink alcohol	Maximum two	No restrictions
	Total availability assessments (TAvA)	Serving oneself the first alcoholic drink of the day	Maximum one	No further assessments should take place after this

5.4 Results

5.4.1 EMA adherence and reactivity

Participants initiated 971 alcohol Stroop-tasks and completed 950 (97.8%). The 21 incomplete tasks were excluded from the analysis. Participants returned 942 completed craving questions[13]. They also provided 419 (99.8% of possible 420) alcohol consumption diary entries. Participants returned 703 CA (83.7% of possible 840), 101 TA (possible number of TA cannot be determined) and 147 TAvA (85% of possible 173).

To examine for reactivity effects (participants reducing their alcohol consumption over the EMA week due to intense self-monitoring), quantity of alcohol consumption the week before the study was compared to quantity consumed over the EMA week. A bootstrapped dependent t-test (used due to violation of parametric assumptions) showed no differences (drinking before EMA: 18.29 (±15.12) units, drinking over EMA: 18.41 (±12.85) units; $t(57) = -.01$, $p = .999$), indicating no reactivity effect.

5.4.2 Craving

5.4.2.1 Selection of optimal data structure

A 3-level structure (sessions within days within participants) fitted the data better than the 'sessions within days' 2-level structure ($\chi^2=62.24$, $df=1$, $p < .001$) but not the 'sessions within participants' 2-level structure ($\chi^2=0$, $df=1$, $p=1$). The 'sessions within participants' 2-level structure fitted the data better than a single level structure ($\chi^2 = 68.94$, $df = 1$, $p < .001$). Therefore, the 'sessions within participants' 2-level structure was retained as the optimal data structure.

Mean craving per session (scores could range from 1-no urge to 11-very strong urge) was $\beta_0=3.39$ *(Standard Error, SE = 0.18).* Fourteen point seven percent (14.7%) of variance in craving scores was attributed to between participants differences (2nd level of analysis), while 85.3% was attributed to between sessions differences within the same participant (1st level of analysis).

5.4.2.2 Association between AB and craving

The first research question was whether global measures of AB and craving (measures across the whole day) would be associated with each other. A 2-level regression model was

[13] Although a craving question appeared at the end of every alcohol Stroop-task, different completion rates might be due to the app being aborted prematurely, either by the participant or by the app crashing.

fitted to predict craving, with AB as a first-level explanatory variable. At the participant level, the model did not explain any further variance in craving, compared to a 2-level model with no explanatory variables. At the session level, it explained 0.05% of variance in craving. Global measures of AB were not associated with global measures of craving *(β = -0.001, SE = 0.001, p = .318, n=941 sessions).*

5.4.2.3 Fluctuations in craving when measured in temptation assessments (TA), total availability assessments (TAvA), and control assessments (CA).

The second research question was whether craving will be higher in TA and TAvA compared to CA. A 2-level regression model was fitted to predict craving, with type of assessment (CA, TA and TAvA) as a first level explanatory variable. Type of assessment was inserted as dummy variables, with CA being the reference level. At the participant level, the model did not explain any further variance in craving, compared to a 2-level model with no explanatory variables. At the session level, it explained 60.7% of variance in craving. Compared to CA, craving was higher in TA *(β = 4.13, SE = 0.20, p < .001, n = 942 sessions)* and TAvA *(β=5.50, SE=0.16, p < .001, n=942 sessions)*[14]. Results indicate that, as expected, craving elevates in moments of temptation to drink. It also elevates when alcohol is physically present and about to be consumed (figure 5.1).

5.4.3 Attentional bias

5.4.3.1 Selection of optimal data structure

A 3-level structure (sessions within days within participants) fitted the data better than the 'sessions within days' 2-level structure *(χ^2=6.94, df=1, p= .016) but not the* 'sessions within participants' 2-level structure *(χ^2=0, df=1, p=1).* The 'sessions within participants' 2-level structure fitted the data better than a single-level structure *(χ^2=6.94, df=1, p=.016).* Therefore, the 'sessions within participants' 2-level structure was retained as best fitting the data.

Mean AB per session was *β_0=55.86 ms (Standard Error, SE = 3.11).* Three point six per cent (3.6%) of variance in AB scores was attributed to between participants differences (2nd level of analysis), while 96.4% was attributed to between sessions differences, within the same participant (1st level of analysis).

---

[14] For the sake of comparison, analysis was run again using TA as the reference level. Craving was higher on TAvA compared to TA (β = 1.38, SE = 0.24, p < .001, n = 942 sessions).

5.4.3.2 Presence of AB in temptation assessments (TA), total availability assessments (TAvA), and control assessments (CA).

The first research question was whether there was AB for alcohol-related cues across types of assessment (CA, TA, TAvA). To test for that, average AB score in each type of assessment was compared to 0 with a one-sample t-test. As there is no multilevel equivalent of one-sample t-tests, analysis was performed at a single-level (the unit of analysis was sessions). Significant AB for alcohol-related cues was present across all types of assessment *(CA: $t(701)=19.13$, $p <.001$, $d = .72$, n=702 sessions; TA: $t(100) =8.42$, $p<.001$, $d = .84$, n=101 sessions, TAvA: $t(146)=8.02$, $p<.001$, $d = .66$, n=147 sessions).*

5.4.3.3 Fluctuations in AB when measured in temptation assessments (TA), total availability assessments (TAvA), and control assessments (CA).

The second research question was whether AB would be higher in TA and TAvA, compared to CA. A 2-level regression model was fitted to predict AB, with type of assessment (CA, TA or TAvA) as first level predictors. Type of assessment was inserted as dummy variables with CA used as the reference level. At the participant level, the model did not explain any further variance in AB, compared to a 2-levels model with no explanatory variables. At the session level, it explained 0.34% of variance in AB. None of the dummy variables was significantly associated with AB *(all ps > .05, n=950 sessions)*[15]. This suggests that there were no differences in magnitude of AB between CA and TA or between CA and TAvA (figure 5.2).

5.4.4 Alcohol consumption

The research question was whether AB and craving would be better predictors of alcohol consumption when measured in TA and TAvA, compared to when measured in CA. Participants provided measures of AB and craving for both TA and TAvA on only 43 out of 420 days (10.2%). To increase the available data points, TA and TAvA were examined in different models.

5.4.4.1 Prediction of alcohol consumption by AB and craving measured in total availability assessments (TAvA) and control assessments (CA)

As TAvA happened only on days when alcohol was consumed, days when alcohol was not consumed were excluded from the analysis and the outcome variable was quantity of alcohol consumption measured in units (alcohol as continues variable). A 2-level model

[15] For the sake of comparison, the analysis was run again using TA as the reference level. There was no difference in AB between TA and TAvA (p > .05).

(days within participants) fitted the data better than a single-level model *(χ^2=24.64, df=1, p<.001).*

The research question was whether measures of AB and craving taken in TAvA would be better predictors of alcohol use, compared to measures taken in CA. To test that, a 2-level regression model predicted quantity of alcohol consumption, with CA and TAvA measures of AB and craving as first level predictors. At the participant level, the model did not explain any further variance in quantity of alcohol consumption, compared to a 2-level model with no explanatory variables. At the day level, it explained 2.48% of variance in quantity of alcohol consumption. TAvA craving measures were positively associated with quantity of alcohol consumption *(β = 0.36, SE=0.18, p=.04, n=134 days).* All other predictors were non-significant *(ps > .05).* This suggests that measures of craving, taken when alcohol is present and about to be consumed, can predict how much alcohol will be consumed, even after controlling for levels of craving at control moments. However, measures of AB, taken when alcohol was present and about to be consumed, did not predict how much alcohol will be consumed by participants.

5.4.4.2 Prediction of alcohol consumption by AB and craving measured in temptation assessments (TA) and control assessments (CA).

To retain consistency with previous EMA studies in the thesis (see chapter four), and the TAvA analysis above, TA measures of AB and craving were regressed on to quantity of alcohol consumption (units of alcohol consumed)[16], as well as with initiation of drinking (drinking, no drinking). In both cases, a "days within participants" 2-level structure did not fit the data better than a single-level structure *(quantity of alcohol: χ^2=0, df=1, p=1; initiation of drinking: χ^2=0, df=1, p=1).* Therefore, analysis was conducted on a single level.

The research question was whether measures of AB and craving, taken in TA, would be better predictors of alcohol use (either quantity or initiation), compared to CA. A ridge regression (ridge parameters derived through the method described by Cule & De Iorio 2013) was used to test a model that predicted alcohol use with AB and craving measures in CA and TA as predictors. As variables were measured repeatedly, within a short period of time and within the same individuals, a ridge regression was adopted, as opposed to a standard linear or logistic regression, due to concerns over multicollinearity (see Månsson

[16] In contrast to the TAvA analysis, non-drinking days were retained in the analysis, as TA measures took place on both drinking and non-drinking days.

& Shukur 2011; Shen & Gao 2008). For example, CA and TA measures of craving were significantly correlated *($r_s = .36$, $p = .001$).*

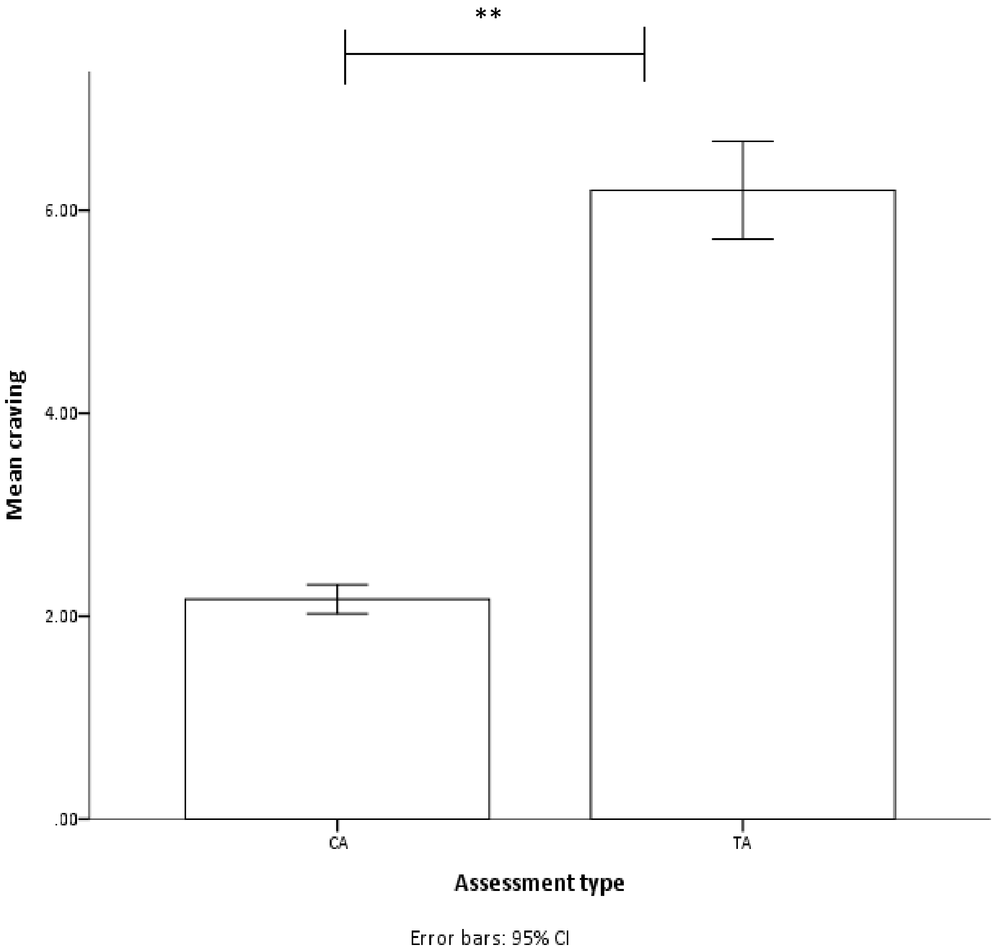
Regarding quantity of alcohol, results showed a non-significant ridge regression model *(multiple $R^2 = 0.02$, $p = .770$, n=78 days)* with all predictors being non-significant *(ps > .05).* Regarding initiation of drinking, results also showed a non-significant ridge logistic regression model *(multiple R^2=0.02, p=.845; n =78 days)* with all predictors being non-significant *(ps > .05).* This indicates that AB and craving did not predict quantity or initiation of alcohol consumption, regardless of whether they were measured in moments of temptation to drink or not.

5.4.4.3 Secondary analysis – When moments of temptation led to drinking

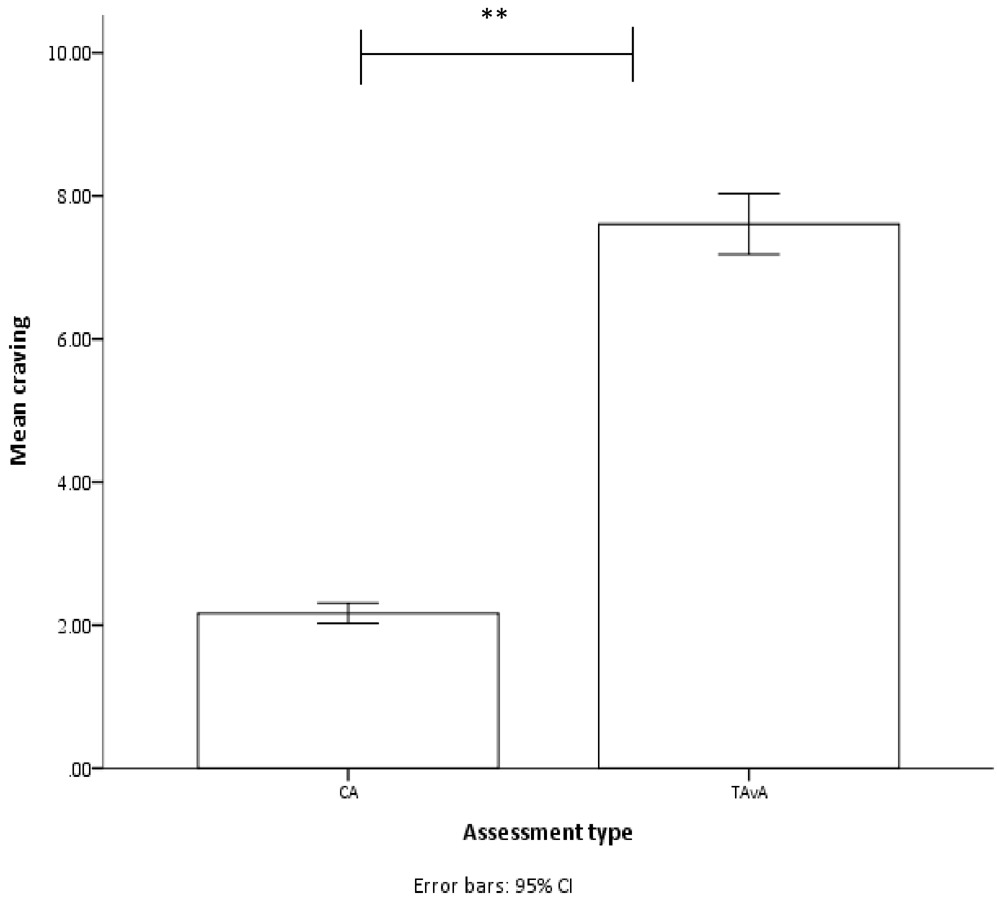
Participants in the present study were not motivated to abstain from alcohol or reduce their consumption. For that reason, it was likely that temptation to drink might have led participants to eventually serve themselves and consume an alcoholic drink. Indeed, 17.8% of TA were immediately followed by alcohol consumption and so these TA could also be considered as TAvA. Analyses on alcohol use described in section 5.4.4.1 were repeated, after recoding those TA as TAvA. Pattern of results, after recoding, remained the same.

Figure 5.1. Mean craving (1-11 scale) per type of assessment

Panel 1



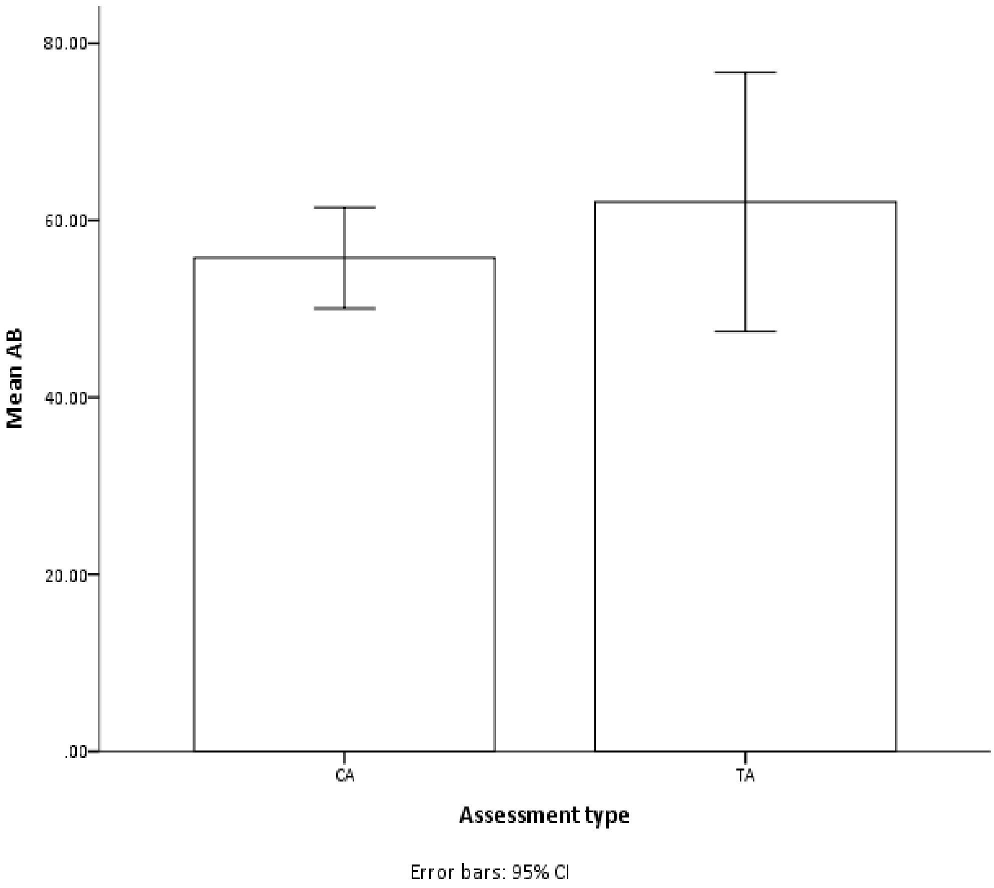
Panel 2



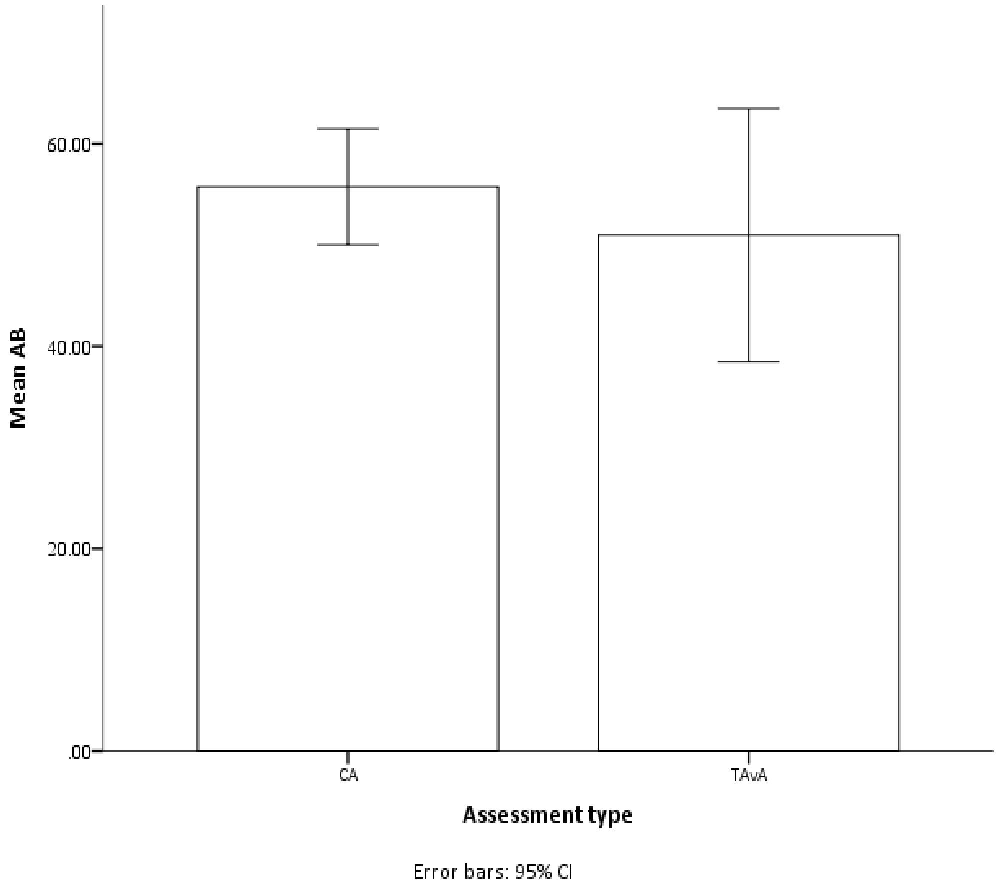
Panel 1: Difference between CA and TA measures of craving, Panel 2: Difference between CA and TAvA measures of craving. CA = Control assessments, TA=Temptation assessments, TAvA=Total availability assessments. ** p < .001

Figure 5.2. Mean AB (ms) per type of assessment

Panel 1



Panel 2



Panel 1: Difference between CA and TA measures of AB, Panel 2: Difference between CA and TAvA measures of AB. CA=Craving assessments, TA=Temptation assessments, TAvA=Total consumption assessments. All differences were non-significant (ps > .05).

5.5 Discussion

The current study used a mixed EMA design (combination of time- and event-related assessments) to examine AB and craving at fixed time-windows (control assessments – CA), as well as when participants felt tempted to drink (temptation assessments – TA), or when alcohol was physically present and about to be consumed (total availability assessments – TAvA). It was predicted that global measures (measures across the whole day) of AB and craving would be correlated, and that AB and craving would be higher in TA and TAvA, compared to CA. It was also hypothesized that AB and craving measures taken in TA and TAvA, but not in CA, would be associated with alcohol use. Findings did not support any of the hypotheses regarding AB, while they partially supported predictions for craving.

5.5.1 AB

The main purpose of the present study was to examine whether the state hypothesis of AB can be supported in its original format (AB magnitude and its association with alcohol use fluctuates according to the motivational context) or whether it might have to be refined according to findings from study one in chapter four (the association of AB with alcohol use, but not the magnitude of AB, fluctuates according to the motivational context). However, findings in the present study did not support the state hypothesis or its refinement.

Contrary to the hypothesis, global measures of AB and craving were not associated with each other. However, these results replicate findings from the previous EMA studies in this thesis (see chapter four), inasmuch as there was no association between global measures of AB and craving. As also discussed in chapter four (see section 4.5.1.1) these findings contradict past research in the laboratory (for a review see Field et al. 2009) and in the real world (Waters et al. 2014). However, these discrepancies may be explained by the following factors (for a more detailed discussion see chapter four, section 4.5.1.1); first the association of AB with craving has been found (Field et al. 2009) to be stronger for illicit substance (as opposed to tobacco and alcohol) and when AB is measured directly (eye-tracking, rather than reaction-time indices; Field et al. 2009), and second, participants attempting to abstain from using the substance (as for example the smoker attempting to quit in the EMA study by Waters et al. 2014) might experience stronger cravings than social drinker not seeking to reduce their drinking (sample of the present study).

Contrary to predictions, AB was not higher at TA and TAvA, compared to CA. This is consistent with chapter four, inasmuch as that AB did not vary in magnitude, irrespective of the context in which it was measured. However, examination of the multilevel structure of AB in the present study revealed that 96.4% of variation in AB was attributed to between sessions differences, within the same participant. This can be interpreted as AB changing from moment to moment but in a random way, which cannot be explained by current temptation or total availability of alcohol.

Comparable magnitudes of AB in TA and CA is in contrast to findings from a previous EMA study (Waters et al. 2012), which found that AB in cocaine and heroin dependent patients was higher at moments of temptation, compared to random measures. Nonetheless, it is not entirely surprising that AB, during an episode of temptation, was greater in drug dependent patients trying to abstain (Waters at al.), compared to non-dependent drinkers with no motivation to reduce their consumption (present study).

One reason might be that addiction, as opposed to patterns of use in non-addicted individuals, is characterized by lack of inhibition and compulsive substance use (for a review see Lubman, Yücel, & Pantelis 2004). It would then be reasonable to expect that an episode of temptation in addicted individuals might induce a higher state of preoccupation with the substance, leading to higher AB.

Another reason could be that dependent patients in rehabilitation, as opposed to non-dependent social drinkers, have a goal of abstinence which can be threatened by an episode of temptation (i.e. increased risk for relapse). As a result, during an episode of temptation, abstinence-oriented patients might perceive substance-related stimuli as threatening. Perceived threat of the stimuli in an emotional Stroop-task can cause a general slowdown of reactions (Algom, Chajut, & Lev 2004), which in the case of an addiction Stroop-task would have been interpreted as higher AB for the substance. Moreover, during an episode of temptation, abstinence-oriented patients might try to avoid any substance-related stimuli and supress any substance-related thoughts. Past studies have found that participants strongly motivated to abstain and avoid alcohol, but not those lacking such motivation, have showed AB for alcohol-related cues (Moss et al. 2013), and that alcohol dependent participants showed AB for alcohol after an alcohol-thought suppression exercise, but not when they could express these thoughts freely (Klein 2007).

The hypothesis that AB would predict alcohol consumption when measured at TA and TAvA, but not CA, was rejected. Regardless of whether the predicted outcome was quantity of alcohol consumption (number of units consumed) or initiation of drinking (drinking/not drinking), AB was not associated with alcohol use when measured in TA, TAvA, or CA.

The first EMA study (chapter four) found that AB predicted alcohol use when measured most proximally to alcohol use. This pattern was not found here, when AB was measured immediately prior to alcohol use (in TAvA). However, the first EMA study examined whether AB would predict initiation of drinking (drinking/no drinking), while this study examined whether AB would predict quantity of alcohol consumption. It is possible that AB can affect the choice between drinking or not drinking but once drinking is initiated, quantity of consumption is affected by other factors (e.g. acute effects of alcohol on inhibitory control; see Field et al. 2010)

Present findings are also in contrast to EMA findings in heroin and cocaine dependent patients (Marhe et al. 2013). In this population, AB measured in moments of temptation to use drugs, but not at random times, predicted chances of relapse in the sort-term (but not in the mid- to long-term). However, for the reasons discussed earlier (greater disinhibition and compulsivity), it is not unreasonable that AB, during an episode of temptation, might have been a stronger driver of substance use in a dependent sample, compared to a non-dependent sample.

Taken together, the present results indicate that AB did not change in magnitude, nor did it predict alcohol use, regardless of whether it was measured in control moments, moments of temptation to drink, or moments of total availability of alcohol. This shows that neither AB, nor its association with alcohol use, fluctuated under the influence of the motivational context. These findings reject the state hypothesis of AB, as well as its refined format suggested in study one (chapter four).

5.5.2 Craving

Results supported the hypothesis that craving will be higher in situations when motivation to drink was expected to be high. However, although craving measured in moments of total availability of alcohol was a better predictor of alcohol consumption, this was not found for craving measured in moments of temptation.

As hypothesised, craving was higher when measured in TA and TAvA, compared to CA. Higher craving in TA compared to CA is consistent with the findings of Jones, Tiplady, et al.

(2018), who examined non-dependent heavy drinkers trying to cut down their alcohol consumption and found that craving measured in moments of temptation to drink was higher, compared to measures taken at random times (although there was a very small number of temptation episodes in that study overall).

Craving being elevated in TAvA compared to CA, is consistent with results in chapter four, where it was found that craving was higher in the late afternoon compared to earlier assessments over the course of the day. In both cases, craving was elevated in the assessment that was most proximal to alcohol use (TAvA here and late-afternoon assessments in chapter four). Moreover, higher craving in TAvA also showed that craving increased when alcohol was available for consumption, which is consistent with previous studies conducted in the real world (Dar et al. 2010; Dar et al. 2005).

As hypothesised, craving predicted alcohol consumption when measured in TAvA, but, surprisingly, not in TA. More specifically, craving in TAvA predicted quantity of alcohol consumption but not initiation of drinking. Craving in TA did not predict initiation of drinking or quantity of alcohol.

Prediction of quantity of drinking by craving in TAvA, but not in CA, showed that craving was a better predictor of alcohol consumption when measured most proximally to alcohol use. This is consistent with findings in chapter four, where craving predicted initiation of drinking when measured in the late afternoon but not in earlier assessments. The finding is also consistent with previous naturalistic studies demonstrating the importance of temporal proximity to substance use in the ability of craving to predict substance use (Fatseas et al. 2015; Serre et al. 2015)

Another implication of the present finding would be that craving is a better predictor of alcohol consumption when measured in the presence of alcohol. This is across the same lines with a previous naturalistic study (Dulin & Gonzalez 2017), which found one of the strongest associations between craving and alcohol use when craving was triggered by the presence of alcohol.

Surprisingly, craving measured in TA did not predict alcohol use (neither quantity nor initiation). Despite this being in contrast to the hypothesis, it is in line with results from previous EMA studies. Marhe et al. (2013) measured craving of drug dependent individuals in moments of temptation, as well as random times, and examined its association with chances of relapse. Their results showed that elevated craving was not associated with increased chances of relapse, regardless of the motivational context in which craving was measured. Likewise, Jones, Tiplady, et al. (2018) measured craving of heavy non-dependent drinkers in moments of temptation, as well as random ones. Due to the low number of temptation episodes reported by participants, this study offered no direct comparison between types of assessments. However, although craving measured at random times was associated with alcohol use, when measures taken in moments of temptation were added to the predictive model, they did not account for any additional variance in alcohol use.

The lack of association between TA measures of craving and alcohol consumption found here, might be explained by ceiling effects. For example, when participants were trained to understand and recognize moments of temptation, these moments were explained to them as moments of strong craving. Therefore, it is reasonable to expect that craving scores were always high every time participants reported a temptation episode. Predicting an outcome variable from an explanatory variable would require a certain level of variability in both variables, so that a predictive model can identify the extent of change in the outcome variable attributed to changes in the explanatory variable. When moments of temptation are isolated, differences in strength of craving might become negligible, with most participants scoring towards the higher end (scores become negatively skewed).

However, a *post-hoc* descriptive analysis on the current data did not support the previous suggestion. Within temptation assessments, measures of craving ranged from one (scale minimum) to eleven (scale maximum), while in 52.5% of the assessments participants reported craving scores below or equal to the scale median (≤ 6). Also, distribution of craving scores within temptation assessments was not notably negatively skewed *(skewness statistic was -.16 and its SE was .24, which shows that the skewness statistic was smaller than two times its SE).*

Taken together, the fact that craving predicts quantity of alcohol consumption when measured in TAvA, but not TA, might be linked back to the distinction made previously between dependent and non-dependent drinkers (compulsive use with lack of control in the first case, as opposed to more voluntary decisions to drink in the second). High craving in a moment of temptation might not necessarily drive a non-dependent drinker to seek for alcohol, if they have other reasons not to do so (e.g. being at work). However, in a moment that participants have already decided to drink and are ready to serve themselves (TAvA), higher craving can signal that a larger quantity of alcohol will be consumed.

5.5.3 Limitations and strengths

One of the central comparisons in the present study was between measures taken in control moments and those taken in moments of temptation to drink alcohol. In clinical literature, relapse is an important treatment outcome and, therefore, moments of temptation are of particular interest, as they might increase chances of relapse (see for example Waters et al. 2012 and Marhe et al. 2013). However, moments of temptation in the present non-clinical sample would differ qualitatively from temptations experienced by dependent patients. They would probably be weaker, pose less risk for the individual's goals (if indeed they have any goals to limit drinking), and elicit less efforts to resist. Also, non-clinical samples may often engage in alcohol drinking not because they fail to resist a strong urge but for social and re-creational reasons.

However, this study has also some notable strengths. In chapter four, participants could have initiated drinking any time after the late-afternoon assessments, which introduced great variability in the level of proximity of these assessments to alcohol use. Here, participants were assessed immediately prior to serving themselves an alcoholic drink (TAvA), fairly standardizing and minimizing the time interval between assessment and alcohol use.

5.5.4 Conclusion

The motivational context in which measures were taken had an impact on craving but not on AB. The present findings do not support the hypothesis that AB, or its association with alcohol use, will fluctuate under the influence of the current motivation to drink alcohol. On the other hand, in situations where motivation to drink was expected to be high, craving was elevated and, in some cases, it also predicted alcohol consumption.

Chapter six: Focus-group interview

6. My finger is not a rectangle: A focus-group exploring attitudes of users of the smartphone-based alcohol Stroop-task

6.1 Abstract

Smartphone apps designed to measure or modify substance-use behaviour are widely used in research and are readily available for download. Given their popularity, it is important to examine the attitudes of prospective users. The aim of the present study was to examine the attitudes of respondents using the app specially developed for this thesis to measure AB in the real world (a smartphone-based alcohol Stroop-task). Six regular social drinkers used the app and discussed their views in a focus group. Using an inductive thematic analysis, three themes were identified: users as research participants, users as clients, and users as researchers. The themes represented the main perspectives from which respondents interacted with the app. Each theme was further split into subthemes, representing the important aspects of the app within each theme. Users as participants focused on stand-alone features of the app, and its day-to-day usability. Users as clients focused on the health-behaviour content of the app, its interface, and its potentials for commercialization. Users as researchers focused on the stimuli used in the alcohol Stroop-task, and on other general methodological issues. Overall, apart from the aspects strictly related to the app's scientific validity, developers should carefully consider aspects making the app more appealing and more user-friendly.

6.2 Introduction

Smartphones (or personal digital assistants – PDAs) have often been used in substance-use research to deliver interventions or assess substance-related behaviours (for a review see Capon, Hall, Fry, & Carter 2016) and, as described extensively in this thesis, they have also been used to assess attentional bias (AB) for substance-related cues (see for example Marhe et al. 2013; Waters et al. 2012; Waters et al. 2014). Their portable and ubiquitous nature makes them ideal for ecological momentary assessment (EMA; for the importance of this method in substance-use research see Shiffman 2009), while their ability to sense and/or record the environment (e.g. cameras, microphones, GPS, etc.) puts them in a unique position to identify antecedents of substance use in real time, and help with relapse prevention (see for example Gustafson, Shaw, Isham, Baker, Boyle, & Levy 2011; McClernon & Choudhury 2013). Also, the fact that larger and more diverse samples can download an app in their smartphones, compared to partaking in any given laboratory study, allows for the examination of more subtle effects (increased statistical power) that would be otherwise undetectable (Brown, Zeidman, Smittenaar, Adams, McNab, Rutledge et al. 2014).

Apart from their use in research, there is also a plethora of commercialized apps that can be downloaded in personal smartphones. Available options include apps for drinking reduction (e.g. Crane, Garnett, Brown, West, & Michie 2015), smoking cessation (e.g. Abroms, Lee Westmaas, Bontemps-Jones, Ramani, & Mellerson 2013), and even alcohol AB modification (e.g. Cox, Intrilligator, et al. 2015). However, the great variety of mental health and health behaviour-change apps in the market has often been accompanied by criticism regarding their lack of efficacy evidence (Donker, Petrie, Proudfoot, Clarke, Birch, & Christensen 2013), use of theory in app development (Helf & Hlavacs 2016), and incorporation of valid behaviour-change techniques (Crane et al. 2015).

Given the popularity of smartphones in substance-use research and treatment, there has been an interest for understanding the attitudes of users of these apps. Users of health behaviour-change apps reported that they valued features, such as instant access to information, behaviour monitoring, ability to integrate use of the app with their normal use of the smartphone, credibility of the app source, and being able to control the actions performed by the app (Dennison, Morrison, Conway, & Yardley 2013). However, app users were also concerned about privacy and information security, and the reliability of the well-advertised ability of apps to sense the environment.

The great number of health-behaviour apps available for download (including substance-use behaviours, such as smoking and drinking alcohol) would also raise the question regarding the factors that influence initial selection and continuous engagement with an app. For example, smokers and alcohol drinkers who were interested in using an app to help them reduce their use, reported that they would select which app to download based on inviting outlooks, positive reviews, a recognizable source, and a title relevant to the app's purpose (Perski, Blandford, Ubhi, West, & Michie 2017). After the initial selection, respondents mentioned that continuous engagement was more likely if the app boosted motivation to use and autonomy, and created a sense of personal relevance and trust. However, continuous engagement was negatively affected by any type of public disclosures (e.g. social media posts) embedded in the app.

Furthermore, a study (Crane, Garnett, Brown, West, & Michie 2017) has examined the attitudes of respondents motivated to reduce their drinking, regarding an app which offered various behaviour-change modules (e.g. normative feedback and action planning), including an AB modification (ABM) program[17]. Respondents often expressed confusion and uncertainty on how to use the app. They also found that it was important for the app to be easy to use, have a clear benefit for them, and be rewarding, so they feel motivated to use it again. Regarding ABM, participants appeared sceptical about its ability to help them. They saw no direct links between the task (reacting to probes replacing pictures) and their drinking behaviour, and they found the proposed mechanism of change (retraining of AB) quite simplistic. Taken together, the literature suggests that app users prefer apps that are easy to use, demonstrate clear benefits, and are generally inviting (e.g. nice outlook and positive reviews). In the contrary, they repel apps that can be perceived as too complex or too intrusive.

Previous studies in the thesis used AlSAp (the AB measuring app developed exclusively for this thesis, see section 2.2 in chapter two) to quantitatively explore variations in AB and craving, and their association with alcohol use in the real world. To supplement these findings, this qualitative study examines the attitudes of prospective users of AlSAp. The present study also fills a gap identified in the literature, as there was a lack of studies examining the views of users of an app that would assess (but not re-train) their AB in real time.

[17]The program was a Visual Probe Training, where participants reacted to probes (e.g. an arrow) that replace alcohol or control cues. For more details see chapter one, section 1.6.1.2

The present study interviewed regular social drinkers who used AlSAp for the first time and analysed their responses using Thematic Analysis (Braun & Clarke 2006). The main aim of the study was to examine usability of the app (features that work well and those that do not), potentials of the app to be commercialized, and general suggestions for improvement.

6.3 Methods

6.3.1 Respondents

Six respondents (four females, 67%) were recruited from the student and staff population of the University of Liverpool using intranet advertisements and word of mouth. Their average age was 29.70 years (±11.57) and, on average, they consumed 20.63 (±16.21) units of alcohol in a typical week.

Criteria for inclusion in the focus group required all respondents to be over 18 years old, speak fluent English, know how to use a smartphone, and consume at least one alcohol drink in a typical week (regular social drinker). Also, prospective respondents should not have any history of alcohol use disorder or suffer from colour-blindness (as the Stroop-task requires colour identification). Age and typical alcohol consumption were self-reported prior to the focus group.

6.3.2 The Alcohol-Stroop app (AlSAp)

For a detailed description of the app, see chapter two, section 2.2 and figure 2.2

6.3.3 Interview

The focus group was facilitated by the lead interviewer. An assistant interviewer, with prior experience of conducting focus groups, was also present to take notes and provide assistance. The interview was split in three parts; presentation of the app, guided use of the app, and discussion of respondents' opinions.

During the first part, a short presentation was given where the phenomenon of AB and its hypothesized association with alcohol use was briefly explained. Following this, an explanation of the app and a short demonstration was given. In the second part, smartphones with the app installed and step-by-step instructions were given out. Respondents were left to complete one full assessment with the app, with the lead interviewer being available to answer any questions.

In the third part, a semi-structured interview guide was used to stimulate discussion. The topic guide (table 6.1) was adapted from the interview questions used in Crane et al. (2017). Respondents were initially asked their opinions about the app in general (e.g. how easy of difficult they find it), before asked to consider how they would feel about using the app repeatedly over a period of time (as in the context of an EMA study for example). They were then asked to express opinions on the scientific credibility of the app and its theoretical background. At the end of the discussion, they were given the opportunity to make general comments and suggest general improvements.

6.3.4 Procedure

Those who expressed interest for the study were provided with more details and a date for the focus group. Those confirming they are happy to participate were provided with a participant number and a link to complete a Time Line Follow Back questionnaire online (TLFB; Sobell & Sobell 1990). Informed consent and basic demographic information (age and gender) were taken at the beginning of the TLFB. Respondents were informed that completing the online questionnaire was a pre-requisite for the focus group.

On the date of the focus group, respondents were informed about the ground rules (e.g. discussion will be recorded and conducted in a relaxed manner with no right or wrong answers), they provided another written informed consent, and they signed a confidentiality agreement (regarding not sharing any content of the discussion with people outside of the focus group). The focus group lasted for one hour and at the end respondents were debriefed and compensated for their time.

6.3.5 Analysis

The recording of the focus group was transcribed verbatim and analysed using Thematic Analysis (Braun & Clarke 2006). Thematic Analysis was chosen as it has been used effectively in previous qualitative explorations of health-behaviour apps (Crane et al. 2017; Dennison et al. 2013; Perski et al. 2017), and because it does not require researchers to interlink their identified themes to create a coherent final theory (as opposed to other qualitative methods like Grounded Theory; Charmaz 1996). The later point was important, as the aim of the present study was to simply summarize respondents' opinions about AlSAp, rather than crate a theory of how users interact with digital tools for behaviour assessment.

Regarding the specific features of the analysis (see Braun & Clarke 2006), the epistemology was essentialist/realist (considering the meaning that is expressed through language as a straightforward indication of experience), the coding was inductive (coding explored all possible themes, without trying to fit to predetermined themes), and themes were identified at the semantic level (focusing on the explicit meaning, without trying to infer underlying ideas, assumptions, etc.).

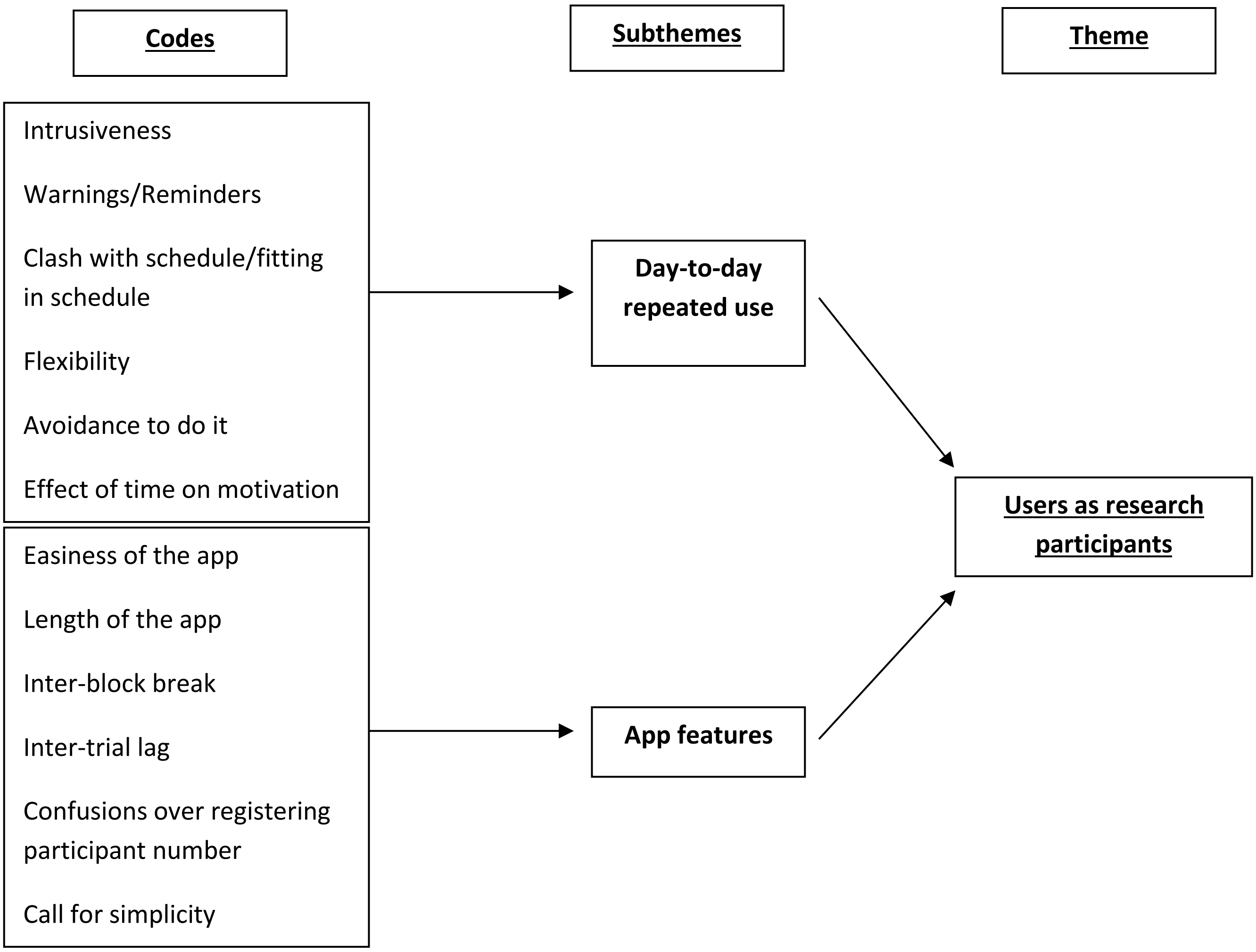
First, extracts from respondents' disclosures were coded, with each code attempting to summarize the central meaning of the extract (e.g. '*I'd probably prefer to have the flexibility to do it (the app) at three rough times during the day'* coded as 'flexibility'). Next, individual codes were organized under general themes. Themes were further split in subthemes, with each subtheme describing a distinct aspect of the theme (see figure 6.1 for an example). Following that, a process of going back and forth between codes and themes was followed, to ensure that all codes fit well within their allocated themes, and that when the themes were taken together, they adequately described the interview. At the end of this process, themes were refined (if needed) and named.

Table 6.1. Topic guide for the focus group

Interview topics
I. Was there anything you particularly liked or disliked? How did you find the length and difficulty level?
II. If you were asked to use the app again, when on your own, without the researcher being present to provide help, was there anything you would find particularly difficult or easy to do?
III. How many times per day would you be willing to use the app and for how long? Would you find repeated use of the app intrusive?
IV. How would you prefer to be reminded to use the app during the day? For example, would you prefer to rely on your memory or receive some prompts?
V. If our aim is to examine how much your attention is automatically attracted by alcohol-related cues, what type of cues would you like to see there? What would you think would be the best comparison level (control cues)?
VI. How believable would it be to you that the app could predict if/how much you will drink later?
VII. Do you have any suggestions for how the app could be improved or other general comments?

Topics were adapted from Crane et al. (2017)

Figure 6.1. Construction of the theme "Users as research participants"



6.4 Findings of the Thematic Analysis

Three themes were identified (table 6.2), reflecting the main perspectives respondents had on the app. Subthemes were also identified, describing the different aspects of the app that were found to be important within each theme. The first theme was "users as research participants" with two subthemes; "repeated day-to-day use" and "app features". The second theme was "users as clients" with three subthemes; "health-behaviour content", "app interface" and "marketing/commercialization". The third theme was "users as researchers" with two subthemes; "stimuli" and "general methodology".

6.4.1 Theme one: Users as research participants

This theme described respondents' views, as if they were to use the app as participants of a research project. They commented on specific features of the app (subtheme one) that they liked or did not like, and on the usability of the app if it was to be used repeatedly every day (subtheme two).

6.4.1.1 Subtheme one: App features

In this subtheme, stand-alone features of the app, not related to its daily use, were found to be important. Positive features were that the app was easy to use and that there was a break half-way through.

> *"I've done other ones in the past and they've just gone on and on and on and had no break in between, so that's good" - participant 2.*
>
> *"It's not actually as complicated as it looks" - par. 3*
>
> *"I found it ok. I didn't think it was too tricky" – "I was happy to have the break" - par. 4*
>
> *"it's nice to have a little break in the middle I think, just to stop you"* - par. 5

However, negative comments were made about the inter-trial lag. The time between tapping a response button and the app moving to the next trial (inter-trial lag) was found to be too long and it disrupted the flow of the app.

> *"three seconds I think is a long time. I don't think I close to ever getting timed out on it" - par 1.*

"when you press the button, it didn't click on straight away from the next one" - par 2

"I think the breaks between the trials could be shorter. It felt like I was waiting quite a bit until the next trial was loaded" - par 6

Also, respondents found the process of inserting their subject number confusing. In the experiment preparation part of the app (see section 2.2.2, chapter two), before the Stroop-task begun, respondents had to manually insert their subject numbers. They had to use a pop-up keyboard to erase the number 0 which appeared there by default and type their own number.

"it would make sense if you typed in your subject number, cos that's like relevant to you, rather than just being a random number" – "I don't think that'd be too tricky, if people knew that was their number" - par 1.

"I found it fine, accept the participant number [...] I might get mixed up [...] You could see that that could go wrong" - par 2

"I think, cos it already said zero in there as well, and you had to get rid of" - par 4

Overall, respondents asked for more simplicity on the steps required to get to the Stroop-task (the experiment preparation part). The task was programmed like a computerized cognitive experiment that could also be completed on smartphones. A drawback of this was that some steps which would normally be completed by the experimenter, were left to be completed by the participants themselves.

"Maybe if the keyboard only came up if you clicked on the box [...], because you needed to enter something" - par. 1

"So like simplify it, so there's not an entire keyboard, just like an enter button" – "if someone doesn't know it's coming it might throw them, so maybe just [...] get rid of the beginning bit" - par. 3

"with those bits that we didn't need to change, did we need to see them? [...] they could be hidden" - par. 5

"enter it like just a number, without having to delete anything" – "I think it would be maybe easier if it was just like 'do you want to continue; yes or no' without having

to maybe type too many things" – "People might be put off by the fact that they might not remember what to do next" - par. 6.

6.4.1.2 Subtheme two: Day-to-day repeated use

Respondents also commented on the usability of the app if they had to use it repeatedly as they go about their lives (EMA modality). There was no consensus regarding whether the app was too long and whether completing it three times per day would be too intrusive. One respondent found the app to be long, while others believed that completing it three times per day was just as acceptable (but no more than that). One respondent even changed their mind during the focus group.

> *"I think it's short enough. It's not like you're sat there for an hour having to fill something out" – "it's not like a huge inconvenience on people [..]. It's not gonna take ages" – "I wouldn't do any more - So three times a day is already a lot. And if you're doing that for a week like how many is that? 21 times in a week of the same game. People will just say 'I'm not doing it'" - par. 1*

> *"I'd say it was maybe a bit long" - par. 3*

> *"I think three would be okay but more than that..." - par. 4*

> *"I think three is the maximum" - par. 6*

However, they all agreed that if they had to access the app repeatedly over a long period of time, potential clashes with their personal schedules would be a concern.

> *"you could be in an important meeting or dunno" - par. 2*

> *"cos if you couldn't do it, if you were in a meeting or something" - par. 4*

> *"say if I'm at work, I might have things on that particular day" - par. 5*

For those reasons, they stressed the importance of being able to fit the repeated assessments in their personal schedules.

> *"you could do it when you get up and when you go to bed and then it's part of your routine" - par. 1*

> *"you can sort of work it into your schedule then can't you?" - par. 3*

"if it was like, sort of like nine till 12, [...] and it just said 'do this for the morning' and then [...] do it for the afternoon and do it for the evening" - par. 4

"I think you have to be able to build it into whatever you might be doing on that particular day" - par. 5

"I for example get up really early so I could do it at 6am where some other people might not get up till 8am" - par. 6

This led to a consensus that flexibility would be important, if the app was to be used in an EMA research project. Otherwise, participants might avoid completing the task, and their motivation will reduce over the time.

"if you set people times, it's just not gonna work" – "They'd just avoid doing it, cos it's gonna be the same thing over and over" - par. 1

"you're dreading the actual notification coming up and you're going 'oh God'" – "I think cos the first day's okay but then it goes and you start dreading them (the assessments)" - par. 2

"(it would be better) if you have hour windows, like between nine and 10, between 12 and one" - par. 3

"I think that would be better. To snooze it and do it later" - par. 4

"I'd probably prefer to have the flexibility to do it at three rough times during the day" - "you might start with enthusiasm and every intention of doing it, but by Wednesday or Thursday you find you might have missed one, and then I suppose, if you've missed one, do you carry on?" - par. 5

Indeed, EMA studies in the thesis (see chapter four and five) did offer some level of flexibility. Alcohol Stroop-task assessments were organized in three two-hour windows, in the morning, the early afternoon, and the late afternoon. Anchoring assessments on the beginning, the middle, and the end of the day probably made it easier for participants to fit them in their daily routines. For example, they could follow the order of their meals; breakfast, lunch and dinner. Participants also had two hours in each assessment period to choose a moment that would be most convenient to them.

Finally, it was agreed among all respondents that some warnings/reminders would make it much easier to remember using the app and reduce missed assessments.

"I personally don't think I would do it that length, three times in a day, unless I got told to do it" - par. 1

"It would be good to have some kind of warning that, say, an hours' time that the test will be coming" - par. 2

"(It would be good to use a) pop up message (as a reminder)" - par. 3

"if I wasn't prompted, I don't think I would do it three times" – "Maybe a text (as a reminder)" - par. 4

"Without a reminder I don't think I would remember to do it three times a day" – "Like a notification that you have to swipe and maybe open up the app straight away" - par. 5

However, it the EMA studies in the thesis, no reminding prompts were provided and yet the rate of initiated alcohol Stroop-task assessments was quite high (e.g. 86.3% in study two in chapter four).

6.4.2 Theme two: Users as clients

This theme described respondents' views, as if the app was commercially available to be downloaded by users looking to reduce their alcohol consumption. Participants commented on what the app could do to help people (health-behaviour content - subtheme one), how it should look (interface - subtheme two), and how it should be adapted to become more commercial (commercialization/marketing - subtheme three).

6.4.2.1 Subtheme one: Health-behaviour content

Respondents discussed the alcohol-related content of the app, what the app is advertised to do, and how it can help people. Most respondents were not convinced by the explanation that the extent to which attention is automatically attracted by alcohol-related words could predict alcohol use in the sort-term. However, one respondent found it believable

"I don't know if you could use it to see how likely it is to relapse" - par. 1

"it might give a small indication but I couldn't say 'oh that person's definitely going to drink a pint in the next half an hour - "it's human behaviour at the end of the day" - par. 2

“you’d think actually ‘Yeah, if I’m taking longer to answer, it’s because it (alcohol) is meaning something to me” - par. 3

“when you try something for the first time you don’t always think it’s going to have an impact” - par. 5

“it doesn’t actually convince me personally” - par. 6

This is consistent with the views expressed by respondents in Crane et al. (2017) study who completed an AB modification module in an alcohol-reduction app. For example, one of their participants mentioned “What was the point of that? Seriously. Really, what was the point of that? Am I missing something? No, I’m not impressed, I don’t know what it was, I don’t know why I’ve just done that.” (Crane et al. 2017, p.10).

A factor that made the theoretical mechanism of the app less believable was that the actual measures (reaction times) did not seem very relevant to the behaviour in question (alcohol use).

“If it gave a percentage or something like that: ‘The chances of you having a drink today is kind of 90%’, that would be good” - par. 2

“it only provides you with the accuracy and your reaction time [...] personally, like, I wouldn’t necessarily think ‘oh my reaction time is now slower than it was in the previous trials’ and think ‘oh maybe I want to drink more or less’” - par. 6

One reason for why some respondents were not convinced by the theoretical explanation of the app might be that there are multiple interpretative stages between the actual measures (reaction times) and the behaviour of interest (alcohol use). Reaction times are used to infer interference caused by alcohol words, compared to control. This is used to quantify AB for alcohol, which is then translated into the current motivational state to drink alcohol. Finally, this motivation is considered indicative of the chances that one would seek and consume alcohol.

Moreover, the theoretical explanation may have been perceived by some respondents as too mechanistic. In contrast to dependent drinkers, whose drinking is more automatic and compulsive (Lubman et al. 2004), non-dependent drinkers may have more complex motives (e.g. mood, peer-pressure, social occasion, etc.). Not surprisingly, respondents were sceptical about taking a behaviour they consider to be of their own volition and reducing it to reaction-time differences.

In order to make the app more convincing and more useful, respondents suggested to provide more evidence for its effectiveness, add a clearer statement of its purpose, and provide feedback to the user. Respondents also said that merely providing a warning that one is at risk of drinking (behaviour assessment) is not enough, unless it is followed by suggestions on how to avoid this (behaviour-change intervention). The same topic was also reported by Crane et al. (2017), where respondents made it clear that they wanted the app to be beneficial and they wanted to know how it can help them.

> *"People [...] not necessarily got an alcohol problem. So, if it's just predominantly aimed at how likely you are to relapse, there's going to be a limited market" – "(you can provide performance feedback) like a leader board" - par. 1*

> *"I'd want to see like how it's helped people before to then go and buy it, or whatever, or download it" – "I would want it to be, like, not just collecting the data but also, maybe, doing something to help you – "it could say like 'do these things or go to this website or do whatever" -"like on FitBit, you have a graph on the app with your steps each day and stuff" - par. 4*

> *"I guess you'd need to see some results almost for you to give it a trial - You need to see the benefits of using it I" – "Make it like a tool" – "it could either give you the warning signs that you're maybe going to have a drink later or give you some tips, to do things that might stop that happening" – "If there was some way you could see your progress" - par. 5*

> *"to record that (alcohol use) as well and say 'oh the reaction time was slower, and you consumed more alcohol on that day'" – "I would want to see some kind of evidence like how effective it is, how many trials have been done. With what populations and, you know, that kind of statistics basically" - par. 6*

Although offering an alcohol-use reduction intervention was beyond the scope of the current app, these comments clearly indicate that if the app was to become commercially available for individuals concerned about their alcohol use it should be extended to include some behaviour-change techniques.

6.4.2.2 Subtheme two: Interface

Respondents also evaluated how inviting the interface of the app would be to someone downloading it. There was a consensus that the aesthetics of the app required some jazzing up.

> *"it needs to be like some pictures or a white screen, to make it seem inviting whereas the black (referring to the screen background colour) is very like 'urg, God, I have to sit and do this'" - par. 1*

> *"the actual game bit (referring to the Stoop-task) [...] just jazz it up a bit. Make it look like not- so just like- just letters and words." – "It looks a bit dated" - par. 3*

> *"could have some graphics to bring it to life a bit more" - par. 5*

More specifically, respondents found the app too technical (especially when the initial menu screens came up before the Stroop-task in the experiment preparation part) and that the format was too small.

> *"Bigger format. I felt like I was, like, pressing right with the edge of my finger" - par. 1*

> *"if you had a bigger phone it would be better" - par. 2*

> *"it (referring to the experiment preparation part) looks like computer code typing - it looks like's it's gone wrong almost" - par. 3*

> *"It's a little bit technical at the start" – "a bigger box in the middle, cos also the font's quite small" - par 4*

> *"Bigger buttons, because I was conscious of getting on that little rectangle" -par 5*

Notable, many respondents found that the app had a non-natural feel, because of the rectangle response-buttons, and the fact that they had to hold the phone horizontally with both hands. Smartphones use has become so ubiquitous that it can sometimes foster the formation of habits (e.g. constant screen checking; Oulasvirta, Rattenbury, Ma, & Raita 2012). In the same way, respondents seemed to be so used to holding their phones up-right with one hand, and tapping with their thumbs (e.g. when texting or browsing the internet), that they were irritated by the different way of use required by the app.

“A circle would be good (referring to the shape of the buttons)” – “if you have it like that (upright position) it’s literally like all accessible. You can use either hand” - par. 1

“(in real life) you don’t have to spin your phone around, do you?” - par. 3

“I didn’t like that it was rectangular buttons. My finger isn’t a rectangle” - par. 4

“I found it slightly uncomfortable doing it that way (pointing sideward). I always use my phone that way (upright)” - par. 5

“people text like that (by holding the phone upright with one hand)” - par. 6.

6.4.2.3 Subtheme three: Commercialization/Marketing

Respondents made several comments regarding the ways the app could be become more commercial. A popular proposal was to gamify the app.

“making it out as a game” – “Like if it increased in difficulty” – “That would make it more challenging and people would be like ‘oh I’ve got to do it’” - par.1

“(setting performance targets) It’s like a little competition with yourself” - par.3

“(it would be nice) if you have to keep doing it every day to make sure that you keep getting a certain number of points” - par. 4

“you could go faster (at each level of difficulty)” - par. 6

Brown et al. (2014) examined the feasibility of adapting some classic cognitive tasks (e.g. stop-signal tasks) into gamified apps, and they have found that the apps were quite successful, in terms of being downloaded by thousands of users (although that project was intensively and widely advertised). Most importantly, they found that when completed in an uncontrolled environment, the gamified tasks retained the same effect sizes shown in the laboratory.

However, Brown et al. (2014) did not include any AB related tasks (e.g. the Stroop task or the VPT). Boendermaker et al. (2016) compared a classic computerized VPT training for reducing AB with a gamified version (for more details on VPT training see section 1.6.1.2 in chapter one). They have found that AB was reduced only in the classic version, while motivation to do the training reduced over time in both versions (but more so in the gamified version). The authors note that if gamification does not meet the users’ standards (which can be quite high, given the current advancement in computer graphics and technology), it may reduce, rather than enhance, motivation to engage.

Gamification of the current app, although not necessarily prohibiting, might present difficulties. Rewarding participants for better accuracy (e.g. choosing the correct colour in all trials) might be too easy to maintain their interest, while rewarding them for faster responses might increase their error-rate (leaving less valid trials available to calculate their AB scores). The app could become more challenging by introducing more difficult versions, as for example if words were presented in white fonts with only a thin coloured outline (see Fadardi & Cox 2009 for an example of a task loosely based on the alcohol-Stroop task with increased levels of difficulty). However, it should be carefully explored whether any such adaptations might affect the reaction-time differences between alcohol and control stimuli.

Respondents in the focus group also asked to make features of the app more relatable, while some of them came up with possible names for the app.

> *"it's just got an O (the logo of open sesame, which is the software the app was programmed with) as the little icon [...] you need to have something that relates to it (alcohol use) so you, kind of, know what the app is about" - "Open sesame (the name of the software the app was programmed with) says nothing about the app to me. I'd probably think it was like some magic" – "(the name could be something like) Colour Pick" - par. 1*

> *"(the name could be) Guess the Colour (or) Match the Colour" - par. 3*

However, it was also suggested that the name and appearance of the app should be discreet, and not give away that the person using the app is concerned with their alcohol use.

> *"It can just be a game. And then no one will know [...]. Cos to anyone looking, it would probably be like just a normal app" - par. 1*

> *"It's quite a personal thing to say 'I wanna reduce the amount I'm drinking and I'm a bit worried about it" – "I think alcohol is very-, like it's a touchy subject" - par. 4*

The topics of appropriate naming and social desirability were also discussed in Perski et al. (2017) and Dennison et al. (2013). When describing how they would choose an alcohol-reduction or smoking-cessation app, respondents in Perski et al. (2017) said they would be looking for a title with key-terms as "reduce alcohol" or "quit smoking". In Dennison et al.

(2013), users of health-behaviour apps admitted they could feel embarrassed if seen using such an app, as one's efforts to achieve a healthy life-style were considered a private matter.

6.4.3 Theme three: Users as researchers

This theme described respondents' views, as if they were a scientist/research who would be interested in using the app for their research. Respondents commented on the alcohol and control stimuli used in the alcohol Stroop-task (subtheme one) and on other, more general, methodological issues (subtheme two).

6.4.3.1 Subtheme one: Stimuli

For the alcohol stimuli, respondents noted that they did not find them salient enough to measure participants' automatic attention. Nevertheless, in most of the EMA studies in the thesis (see chapter four and five), participants showed significant AB for the alcohol words. Respondents in the focus group also found that the alcohol words were too broad/not specific, and unrelated to participants' personal drinking preferences.

> *"To be honest, the words don't really have that much of an impact" – "Personally they meant nothing to me" – "Too broad - Like it said 'spirits'. So, split it up like vodka, gin, stuff like that" – "Big gin drinkers; if they saw gin they'd be like 'bang, done' because they'd be attracted to that word" - par. 1*

> *"Can't remember. The only word I can remember, that was alcoholic, was 'pub'" - par. 2*

> *"They were too broad. Like having more specific types of alcohol maybe" – "I never say 'I really fancy some spirits tonight'" – "Things that people like. It makes them think 'ahh I had one of them the other day'" - par. 3*

> *"I wasn't consciously aware of making any associations" – "Having the specific drinks might be something to look at doing, rather than the general categories" - par. 5.*

> *"I for example drink wine, gin and prosecco, so I wouldn't be attracted to beer" - par. 6.*

Respondents suggested some alternative alcohol stimuli instead of words generally related to alcohol. They mentioned alcohol brand-names, pictures of alcoholic drinks, and a combination of words and pictures.

"(as an alternative use) maybe brands" - par. 1

> *"I thought pictures as well. You could still have the word just underneath it, couldn't you?" - part. 4.*

> *"such as Heineken instead of beer" – "you could also use pictures" - par. 6*

Past quantitative studies have also shown that personalised and pictorial stimuli can be a promising alternative to words generally related to alcohol. AB has been found to be stronger for personalized, rather than general alcohol stimuli (Christiansen, Mansfield, et al. 2015), while measures of AB for personalized stimuli had better internal reliability (Christiansen, Mansfield, et al. 2015) and predictive validity (Christiansen & Bloor 2014), compared to measures of AB for general alcohol stimuli. Also, measures of AB for pictorial substance-related stimuli have been found to have better internal reliability, compared to measures of AB for word stimuli (Ataya et al. 2012).

This thesis also explored potential superiority of personalised and pictorial stimuli in a smartphone-based alcohol Stroop-task (see chapter three). However, when a personalised and pictorial task was compared to one using words generally related to alcohol, both showed acceptable internal reliability but poor predictive validity, suggesting that the use of personalised and pictorial stimuli did not offer any significant advantages.

Regarding the control stimuli used in the app (environmental-features words like bog, ravine, and valley), respondents in the focus group did not find them to be an adequate reference category for the alcohol words. They reported that many of these words sounded unfamiliar, and that familiarity of the stimuli could have an impact on reaction times.

> *"words like ravine is not something I'd use or ever hear very often" – " So, maybe because you're not hearing it often, you're not as quick on it" - par 1.*

> *"there was something like crag or something and I just did not know what that word was, and I was looking at it thinking what is that?" - par 4*

Alternative categories of words that were suggested to be used as control stimuli were non-alcoholic drinks and furniture. Non-alcoholic drinks were found appropriate because of the similar semantics with alcohol (both are appetitive stimuli), and furniture were found to be a good alternative because of their emotional neutrality.

> *"something like drinking but non-alcoholic drinks, so like coffee, tea, soft drink" – "Then you've still got the drinking element but they're not alcoholic" - par. 1*
>
> *"everyday objects like a table, a chair" - par. 6, "I was just thinking: table" – "you could really like some form of food that pops up and you could hate something else. So, yeah, I think inanimate objects (like furniture) is quite a... (good option)" - par. 3*.
>
> *"People feel fairly neutral about furniture. As you've said, you don't want people to actually think about, almost to stop seeing the word" - par 5*.
>
> *"Because you would probably feel quite neural towards it" - par 6*

However, one participant noted that there might be some caveats with emotionally neutral control stimuli.

> *"Would you not want another condition though as well, where it is things you're not neutral about [...], ones that are the more interesting things, like the alcohol. Those words stand out but not just because, maybe, you want a drink, but also just because the words are more interesting" - par 4.*

The apparent advantage of using soft-drinks as control stimuli is that they would control for the appetitive properties of the stimuli in the task, thus offering a purer measure of AB for alcohol (e.g. participants pay attention to beer because it is alcohol, and not merely because it evokes more emotions compared to something as neutral as a furniture).

However, Christiansen and Bloor (2014) found no AB for alcohol when alcohol stimuli were contrasted to soft-drinks, and the same has been found in this thesis too (chapter three study one, and chapter four study two). Also, study one in chapter three showed that participants showed AB for soft-drinks when contrasted to neutral environment-related words, which indicates that soft-drinks may acquire some attention-grabbing properties themselves. Utilization of soft-drinks might dilute the magnitude of difference between reaction times, masking any AB for alcohol (e.g. participants' attention is automatically attracted by beer because it's alcohol, and by coke because it's a drink they generally like, so no difference will be observed between the two).

6.4.3.2 Subtheme two: General methodological remarks

Respondents made various general remarks regarding the methodological characteristics of the app. Often, respondents stressed the importance of counterbalancing the order of the blocks (seeing alcohol stimuli first and then the control stimuli, and vice versa), and the importance of potential time effects (e.g. getting better or losing interest as the task continues).

> *"Maybe a control would be good, were you have the alcoholic ones second, to see if there's any difference" - par. 1*
>
> *"You got used to it after a while. The words started to mean less."- par. 3*
>
> *"I don't know whether you could either mix them in completely (the alcohol and control stimuli) [...]. Because it was, like, by the second one, I was bored. [...] if you didn't know what was going to come up next, then maybe it would take your attention a bit more that way" - par 4.*
>
> *"Towards the end and on the second task, with the environmental ones, I don't remember any of those words. More so with the alcoholic ones, I was reading them, but I don't know if that's cos it was first in the order" - par. 5*

Previous studies have shown that a substance Stroop-task with a block format is more reliable than one with mixed stimuli (Ataya et al. 2012). A blocked format is also preferable to avoid carry over effects (the transition of the interference caused by an alcohol stimulus to the control stimulus following on the next trial; see Waters et al. 2005).

A methodological limitation identified by one respondent was that the app could increase chances of relapse.

> *"I could understand why people might relapse if they did it. Cos they've got that association back and they've got it in their mind. So, I can understand why, later on, after doing that, they may be like 'oh I want a drink', cos they've been looking at those words" - par. 1.*

However, an advantage mentioned by another respondent was that the app could be used to differentiate between different levels of involvement with alcohol. Indeed, previous research has used AB measures taken with the alcohol Stroop-task to differentiate between heavy and light drinkers (Fadardi & Cox 2009; Field, Christiansen, et al. 2007).

" this person who has previously had alcohol problems, and is not drinking, is taking a lot longer to answer than someone who hasn't. You'd be able to, like, compare" - par 3.

Table 6.2. Themes and subthemes identified in the thematic analysis

Themes	Subthemes		
	Subtheme one	Subtheme two	Subtheme three
1. Users as research participants	Repeated day-to-day use	App features	
2. Users as clients	Health-behaviour content	Interface	Commercialization-Marketing
3. Users as researches	Stimuli	General methodological remarks	

6.5 Recommendations for future redevelopment of the app.

The purpose of the focus group was to understand how AlSAp could be redeveloped to become a professional digital behaviour-assessment tool that would meet the needs and preferences of its potential users. According to the views expressed from all three perspectives (users as participants, clients, or researchers), the app should strike the right balance between scientific validity and user friendliness.

First and foremost, the app should be effective, which is closely related to the quality of the stimuli used in the Stroop-task. Further research is needed to understand the advantages and disadvantages of the available choices (e.g. pictures vs words, general vs personalised stimuli, appetitive vs neutral control stimuli, complex vs simple pictorial stimuli, and colour filters vs colour outlines in pictorial stimuli). Pilot studies and focus groups could be used prior to designing a substance Stroop-task, to understand which stimuli would work best with the target population.

The app should also promote autonomy. This means that the app should require minimum effort by the users, and not become intrusive in users' lives. For example, all preparation steps (e.g. assignment of participant numbers, selection of the saving path, etc.) should run automatically in the background, and users should engage only with the behaviour-assessment elements (e.g. the Stroop-task). Also, users should be able to use the app with one hand (same way they use their phones for texting and browsing), and exert some control over when and how often they would like to engage with the app.

Finally, the app should increase level of engagement. To facilitate initial engagement, the app should convince the user about its relevance, quality, and trustworthiness. The name of the app should summarize its purpose, without strong/sensitive language (e.g. alcoholics, alcohol misuse, etc). Its interface (e.g. colours and shapes) should be modern and inviting, and its scientific credibility should come across easily (e.g. use of university logo, references, supporting evidence). To increase chances of continuous engagement, the app should have clear benefits for the user, and maintain their interest. Users should be able to easily understand the association between the measures taken by the app (e.g. colour-naming reaction times) and their drinking behaviour. Also, gamification should be considered to make the app more fun to use and more rewarding. (e.g. points, level of difficulty, rewards, etc.) However, developers should always be careful not to jeopardize the scientific validity of the app when introducing gamification.

Overall, the challenge for future developers would be to ensure the scientific validity and reliability of an alcohol Stroop-app, while also meeting users' expectations. Modern smartphone users are used to sophisticated graphic designs and highly interactive apps that can substantially enhance the user-app interaction. Health behaviour-assessment apps should be able to live up to these expectations, while also serving as research tools to obtain measures in the real world.

Chapter Seven

7. General Discussion

7.1 Scope of the thesis

The scope of the thesis was to assess the state-motivational nature of AB. To that end, a novel smartphone application (the AlSAp) was developed to measure AB (with the alcohol Stroop-task) and craving in the real world. The thesis first examined whether the smartphone-based alcohol Stroop-task would show good psychometric properties when administered in the real world (EMA modality of administration). Then, the app was used to examine variations in AB and craving, as well as their association with alcohol consumption, in various motivational contexts in the real world. Specifically, the situations of high motivation to drink alcohol that were examined here were a) close temporal proximity to initiation of drinking, b) moments when alcohol was physically present and about to be consumed (total availability), and c) moments of temptation to drink alcohol.

There were three overarching hypotheses in this thesis, all based on the model regarding the role of AB in addiction (Franken 2003), and on the state-hypothesis of AB (Field et al. 2014; Field et al. 2016). First, it was hypothesized that AB and craving will be higher when measured in a context of high motivation to drink alcohol in the real world. Second, it was hypothesized that measures of AB and craving will be associated with each other. Third, it was hypothesized that AB and craving will predict alcohol use in the short-term, when measured in a situation when motivation to drink was expected to be high.

Finally, I also qualitatively evaluated AlSAp in a focus group. Users that used AlSAp for the first time were interviewed and their views were organized in themes which described their experience when interacting with the app.

7.2 Overview of findings

7.2.1 Psychometric properties of measures taken in the real world.

The psychometric properties of the smartphone-based alcohol Stroop-task were examined in the two studies described in chapter three. Study one compared a classic alcohol-words Stroop-task (using general alcohol-related words, and environmental features as controls) to an upgraded Stroop-task (using personalized pictures of alcohol, and pictures of soft-drinks as controls). These tasks were either administered on a computer in a neutral university room (traditional modality), or on a smartphone in participants' homes (EMA modality).

Results showed that both alcohol Stroop-tasks had acceptable internal reliability when administered on smartphones in participants' homes but not when administered on laptops in a university neutral room. The upgraded version of the alcohol Stroop-task did not show better psychometric properties compared to the classic, regardless of way of administration (EMA or traditional). However, there was no association between any measure of AB and alcohol use.

In study two, the smartphone-based classic alcohol Stroop-task was administered in the real world, a neutral laboratory, and a semi-natural environment (a Bar-lab). Results showed that the task had acceptable reliability when administered in the real world, but not in the neutral laboratory or the Bar-lab.

Three general conclusions were drawn from these studies. First, both alcohol Stroop-tasks (classic and upgraded) had acceptable internal reliability when administered on EMA modality. Second, increased reliability in the EMA modality should be attributed more to the environment of administration (real world), rather than the mean of administration (smartphone). Third, despite the acceptable internal reliability of some variations of the Stroop-task, its predictive validity was poor.

7.2.2 Effect of the motivational context

The effect of the motivational context was examined in the three EMA studies described in chapter four and five. Across studies, participants were lent a smartphone with the AlSAp installed and used it for one week. Participants also recorded their quantity of alcohol consumption (in units) in daily diaries.

In the first two studies, participants were asked not to drink alcohol before the late afternoon, and their AB and craving was measured in the morning, early afternoon, and late afternoon, so that the late-afternoon assessments were the most proximal to alcohol consumption. One study used the alcohol-Stroop task with general alcohol words, while the other used personalized pictures of alcohol.

Results from both studies showed that AB remained stable over the course of the day without rising in magnitude when measured proximally to initiation of alcohol consumption. The first study (with the Stroop-task with alcohol-related words) found that participants showed significant AB for alcohol-related cues and that AB predicted initiation of alcohol consumption when measured proximally to alcohol consumption (i.e. late-afternoon assessment) but not earlier (i.e. morning or early afternoon). However, these findings were not replicated in the second study (with the Stroop-task with personalised pictures of alcohol).

Regarding craving, both studies found that it was elevated when measured proximally to initiation of alcohol consumption, as well as that craving predicted initiation of alcohol consumption when measured proximally to it (i.e. in the late-afternoon session), but not earlier (i.e morning, or early afternoon). Finally, in both studies, global measures of craving (measures taken across the whole day) were not associated with global measures of AB. For example, in the late afternoon, although craving was at its highest, AB remained at the same level with earlier assessments.

In the third EMA study, AB and craving were measured in moments of total availability of alcohol (when alcohol was present and about to be consumed), moments of temptation to drink alcohol, and control moments. Results showed that participants showed significant AB for alcohol-related cues at all moments, that AB was not higher in moments of temptation or total availability (compared to control), and that AB did not predict alcohol consumption (regardless of moment of assessment). It was also found that craving was elevated on moments of temptation and total availability (compared to control), and that it predicted alcohol consumption when measured in moments of total availability but not in moments of temptation. Again, there was no association between global measures of AB and craving.

Overall, hypothesis one (AB and craving will be elevated in high drinking-motivation situations) and three (AB and craving will be better predictors of alcohol consumption when measured in high drinking-motivation situations) were partially supported (true for craving but not for AB). Results did not support the second hypothesis that global measures of AB and craving will be associated to each other. The general conclusions drawn from the studies were that craving is a better predictor of alcohol consumption than AB, and that there is limited evidence that AB is a motivational state.

7.2.3 Focus group

Chapter seven described a focus group where respondents used the AlSAp for the first time and expressed their views. The thematic analysis revealed that respondents interacted with the app in three ways; as if they would have to use it as participants in an EMA study (theme one), as if they would have to use it as independent clients (theme two), and as if they were researchers interested in using the app for their projects (theme three). The general conclusion drawn from the focus group was that any successful further development of the app, either as a research tool (i.e. digital portable AB assessments) or a behaviour-change tool (i.e AB modification programs), should ensure effectiveness, autonomy of use, and motivation to engage.

7.3. Theoretical implications

7.3.1 Is AB an automatic response to alcohol-related cues that is associated with craving?

The cognitive-motivational model of addiction described by Franken (2003) suggests that (1) substance users would show AB for substance-related cues, and (2) AB will have a reciprocal relationship with craving. Findings of the thesis partially supported this model for non-dependent drinkers of alcohol, providing evidence for the first suggestion but not the latter.

Across all studies that compared alcohol-related cues to neutral control cues, the alcohol-cues caused greater interference in the alcohol Stroop-task than control cues, indicating significant AB for alcohol stimuli. This is consistent with previous results showing that alcohol-related cues can grab and hold the attention of non-dependent alcohol drinkers (e.g. Christiansen, Mansfield, et al. 2015; Miller & Fillmore 2010; Monem & Fillmore 2017; for a review see Field and Cox,2008), although there are also numerous examples of failures to show these results (e.g. Cox et al. 2003; Field, Duka, et al. 2007; Manchery et al. 2017).

However, participants showed no AB in studies where alcohol-related cues were contrasted to soft-drinks cues, which is consistent with previous studies (Christiansen & Bloor 2014; Cox et al. 2003). As discussed in chapter three (see section 3.3.4), one interpretation could be that appetitive stimuli (e.g. soft-drinks) are more appropriate control stimuli, compared to neutral cues (e.g. environmental features), as they allow for purer measures of alcohol AB, controlling for the appetitive nature of stimuli. Incentive motivation models of addiction suggest that alcohol can attract attention above and beyond other natural rewards, so alcohol-related cues would be expected to attract attention more than other appetitive stimuli. However, the opposite could also be argued, that appetitive control stimuli may muddy the waters by contrasting two salient stimuli that both attract attention for different reasons (e.g. attention is attracted by beer because it is alcohol and by coke because it is a motivationally salient drink).

In contrast to the predictions of Franken's model, this thesis found that AB and craving functioned independently. Increases in craving over the course of the day did not translate to increases in AB (i.e. there was no co-fluctuation), and AB was not higher on moments when participants felt tempted to drink alcohol, compared to moments of no temptation. Findings of a previous meta-analysis on the association between AB and craving (Field et al. 2009) provide useful context for these findings. The meta-analysis found an existing but generally weak association between AB and craving, which was more evident in cases of illicit substances compared to alcohol, and when AB was assessed directly (e.g. eye-gaze indices of AB rather than indirectly through reaction times). As a result, in a population of non-dependent social drinkers a reaction time-based measure of AB may show little concordance with craving.

The lack of association between AB and craving also contradicts results from a real-world study (Waters et al. 2014), where AB of smokers attempting to quit was associated with craving. However, participants in this thesis were social drinkers not looking to reduce their alcohol consumption. As discussed in chapter four (see section 4.5.1.1), in the case of the smokers attempting to quit, the constant interruptions of the habitual substance-use behaviour might have induced strong craving (which is consistent with predictions from the cognitive process model; Tiffany & Conklin 2000). In the contrary, in the present samples, despite some momentary increases over the course of the day (e.g. in the late afternoon), average level of craving was low to medium. As a result, the association between AB and craving might have been more apparent in a sample of smokers attempting to quite due to higher levels of craving compared to the current samples (which is also consistent with findings from the meta-analysis discussed earlier).

7.3.2 Is AB and craving elevated when motivation to drink alcohol is expected to be high?

According to the state-hypothesis of AB, (Field et al. 2014; Field et al. 2016), AB would be higher when the underlying motivation to drink alcohol is also high. As motivation can change rapidly over time, AB would therefore be expected to fluctuate over time. However, there was no evidence that AB in the real world fluctuates over the course of a drinking day or that it was affected by motivational variables (temptation to drink or total availability of alcohol).

This contradicts previous findings from laboratory studies. In studies where the incentive value of alcohol has been experimentally manipulated (e.g. alcohol tasting aversive), AB was higher when alcohol had greater incentive value (Rose et al. 2013; Rose et al. 2018). It has also been found that AB was higher when craving was stronger (for a review see Field et al. 2009), when drinkers were exposed to alcohol-related cues (e.g. seeing or smelling alcohol, Ramirez et al. 2015) and when alcohol was perceived as readily available for consumption (for a review see Jędras et al. 2013).

However, this thesis examined AB in the real world, not in the laboratory where the drinking experience often differs considerably from real world conditions; for example participants being asked to consume beer that has been adulterated to taste aversively bitter in Rose et al. (2013), or consume 20ml of beer with a straw during an eye-tracking task in Field, Hogarth, et al. (2011). In general, it is not surprising that laboratory findings are not consistent with findings from the real world, given the much-reduced ecological validity of laboratory findings. For example, Begh et al. (2016) showed that measures of AB of smokers attempting to quit, taken in a clinical environment, were not indicative of how their attention would function in the presence of smoking cues in the real world. They found no association between AB scores in the clinic (on the VPT and smoking Stroop-task), and self-reported measures of how many smoking-related cues participants have spotted in the real world, or the extent to which participants felt that their attention was automatically attracted by those cues.

However, results of this thesis also contradict a previous real-world study (Waters et al. 2012). Waters et al., reported that the AB of drug dependent patients in rehabilitation (measured with a drug Stroop-task) was elevated when measured in moments of temptation to use drugs, compared to random assessments, whereas this thesis demonstrated that AB remained stable across control and temptation assessments. However, this discrepancy could be explained by the difference in the level of involvement with the substance in the two samples (dependent patients as opposed to social drinkers).

The reasons as to why AB might elevate in moments of temptation in dependent samples, but not in non-dependent, have been discussed in detail in chapter five (section 5.5.1.2). First, dependent samples might experience a greater preoccupation with the substance during an episode of temptation due to the compulsive nature of substance use in dependent individuals (Lubman, Yücel, & Pantelis 2004). Second, abstinence-oriented dependent patients might have a strong motivation to avoid substance-related cues and supress any substance-related thoughts during an episode of temptation. Both of these processes have been shown to induce AB for substance-related stimuli, especially when AB was measured with the Stroop-task (thought suppression: Klein 2007; motivation to avoid alcohol: Moss et al. 2013). Finally, substance-related cues might be perceived as threatening by abstinence-oriented patients during an episode of temptation due to the risk of relapse. Perceived threat of stimuli has been shown to inflate reaction times in the emotional Stroop-task (Algom et al. 2004), which in the context of the substance Stroop-task would be interpreted as greater AB for substance-related cues.

Regarding craving, it was found to be elevated when the motivation to drink alcohol was expected to be high (high temporal proximity to initiation of drinking, total availability of alcohol, and moments of temptation to drink). Previous studies in the real world have found comparable results in similar motivational situations. For example, craving of smokers in a plane was found to elevate towards landing (proximity to use; Dar et al. 2010), and craving of smokers was found to be higher on regular days, compared to days when smoking was prohibited for religious reasons (perceived availability; Dar et al. 2005). Also, craving for alcohol was higher when participants self-reported that they felt tempted to drink alcohol (Jones, Tiplady, et al. 2018). However, Kirchner, Cantrell, Anesetti-Rothermel, Ganz, Vallone, and Abrams (2013) analysed GPS data and found that number of tobacco sale points encounter on a day (presence of the substance), was not associated with daily levels of craving.

7.3.3 Summary of theoretical implications

Results of this thesis supported the prediction of AB theoretical models that the attention of alcohol drinkers would be automatically attracted by alcohol-related cues. However, results did not support the prediction that the extent to which attention is automatically attracted by alcohol-related cues is correlated with craving intensity, or that it would be dependent on the motivational context where measures are taken. Overall, the present results suggest that most of the predictions of the AB theoretical models might not apply within social drinkers of alcohol, not seeking to reduce their drinking.

7.4 Practical implications

7.4.1 Is AB a robust predictor of alcohol use, when measured in the right motivational context?

Another prediction of the state-hypothesis of AB (Field, Marhe, & Franken 2014; Field, Werthmann, Franken, Hofmann, Hogarth, & Roefs 2016) is that AB would be associated with alcohol use, but only when AB is measured in the same motivational context where alcohol use takes place, and only in the short term. However, findings from the thesis provided limited support for this hypothesis. Although one study found that AB was associated with alcohol consumption when measured proximally to initiation of alcohol consumption (study one, chapter four), this pattern was not replicated in a second study (study two, chapter four). Also, AB did not predict alcohol consumption when measured in moments of total availability of alcohol, or in moments of temptation to drink in the final EMA study (chapter five).

Findings of the thesis add to the large body of evidence from laboratory studies showing no association between AB and alcohol use, both in non-clinical (e.g. Christiansen & Bloor 2014; Cox et al. 2003; Field, Duka, et al. 2007) and clinical samples (for a review see Christiansen, Schoenmakers, et al. 2015). However, findings of the thesis are inconsistent with an EMA study (Marhe et al. 2013), who found that magnitude of AB for drugs, measured on moments of temptation, indicated whether a drug dependent patient would relapse within the first week of the study. For the reasons discussed earlier (stronger craving and compulsive, uncontrollable, substance use in dependent samples), it is not unreasonable that AB, during an episode of temptation, might have been a stronger driver of substance use in a dependent sample, compared to a non-dependent sample.

7.4.2 Is there a clinical utility in AB modification (ABM) interventions for alcohol use?

Although the thesis did not examine the effectiveness of ABM, the present findings have far-reaching implication for these programs. ABM aims to retrain AB away from alcohol-related cues and towards alternative targets (e.g. healthier choices like non-alcoholic drinks). Their clinical utility is based on the hypothesis that reductions in AB would cause reductions in alcohol-use behaviour (see chapter 1, section 1.8). However, this thesis provided limited evidence that AB can have a robust effect on alcohol-use behaviour, undermining the proposed mechanism of action of ABM.

This would be consistent with previous findings which robustly demonstrate that although ABM might be effective in re-training AB[18], this does not have an effect on alcohol-use behaviour (Boendermaker et al. 2016; Field, Duka, et al. 2007; Schoenmakers et al. 2007, for a meta-analysis see Cristea et al. 2016). This pattern has also been demonstrated in ecological momentary intervention (EMI) studies administering ABM to smokers (Kerst & Waters 2014, however see also Robinson et al. 2017). This lack of robust associations between AB and substance-use behaviour may also explain the equivocal evidence around the clinical utility of ABM in reducing chances of relapse to substance use (for a review see Christiansen, Schoenmakers, et al. 2015).

ABM should be used with caution in any alcohol use intervention, given the lack of evidence for its efficacy and the failure in this thesis to support any of the theoretical short-term associations between AB and alcohol use. This should become a main point of consideration for future studies, especially those attempting to further develop ABM into a portable behaviour-change tool to be used in the real world (see for example Crane et al. 2017 for ABM as component of an alcohol-reduction app, and Cox, Intrilligator, et al. 2015 for ABM as a gamified commercialized app). Any such attempts should follow, rather than precede, robust evidence that fluctuations in AB in the real world would have consequences for alcohol use in the short term.

However, it should be noted that the findings of the thesis are based on non-dependent drinkers who were not motivated to reduce their drinking. Although the present findings undermine the core mechanism of action of ABM, it would be important for future studies to examine variations in AB and its association with alcohol use in the real world in clinical samples, in order to draw definite conclusions for the clinical utility of ABM.

7.4.3 Would measures of craving have a greater predictive utility than AB?

Craving predicted initiation of drinking when measured proximally to alcohol consumption, and it also predicted quantity of alcohol consumption when measured in moments of total availability of alcohol. Comparable results have also been found in previous studies in the real world. For example Dulin and Gonzalez (2017) showed that craving induced by the presence of alcohol was one of the strongest predictors of alcohol use. Notably, a recent meta-analysis of EMA studies showed that craving is a better predictor of relapse when

[18] Notably there are suggestions that this effect could be attributed to response contingency learning rather than re-training of AB. For example, some of these studies have found no generalizability of ABM effects on novel stimuli or AB measured on different tasks.

measured proximally to substance use (Serre et al. 2015). It should be noted, though, that craving in the thesis did not predict alcohol consumption when measured in moments of temptation, although this is consistent with previous EMA studies (see Jones, Tiplady, et al. 2018; Marhe et al. 2013).

The present results showed that craving predicted alcohol use more robustly than AB. If one of the aims of the thesis was to identify a measure that can be used as a momentary signal of increased chances of alcohol use in the real world, the present results would suggest that craving is a better candidate compared to AB.

Measures of craving in the thesis also showed some practical advantages over measures of AB. Craving was measured with a single question which asked participants to rate the strength of their current urge to drink alcohol. In the context of real-word applications, where measures are taken multiple times in an uncontrolled environment, robustly predicting alcohol use with an easy and quick measure is an attractive possibility. In contrast, a smartphone-based alcohol-Stroop task to measure AB in the real world is more time consuming, both for the researcher to develop it, and for participants to complete. Findings of the focus group in this thesis (see chapter six) provided a nice overview of the complexity of achieving both scientific credibility and user acceptability of an alcohol Stroop-task smartphone application.

7.4.4 Summary of practical implications

Results of this thesis demonstrate that AB was not a robust predictor of alcohol consumption in social drinkers. The lack of association between AB and alcohol consumption also creates concerns regarding the extent to which ABM could be used effectively to help individuals reduce their alcohol consumption, although further exploration of variations of AB and its association with alcohol-use behaviour in the real world in individuals misusing alcohol is warranted. On the other hand, this thesis found that craving was a robust predictor of alcohol consumption in social drinkers, which suggests that craving could potentially be a target variable for programs aiming to identify antecedents of alcohol consumption in real time.

7.5 Limitations and Strengths

7.5.1 Limitations

There are some limitations in the methodology used across the studies in the thesis. AB was measured with an addiction Stroop-task, which has been widely used in the literature for that purpose (Cox et al. 2006). However, it has been suggested that performance on the task might not be indicative only of selective attention for addiction-related cues but also of increased cognitive load due to craving, or a general slowdown in all cognitive processes due to the emotional nature of stimuli (e.g. Algom et al. 2004, for a review see Field & Cox 2008). It should be noted though that there are not many alternatives available. For example, reaction-time indices of AB measured with the VPT (which is the other most commonly used task to measure AB) has been found to be less reliable (Ataya et al. 2012), and it would be practically difficult to adapt it as an app (i.e. both cues are visible on a small screen, without shifting the focus of attention, which is a core component of the VPT).

With regard to the assessment of craving, craving was assessed with a single Likert scale, asking participants to rate the strength of their current urge to drink alcohol. Although these measures have advantages, they also present some problems (see Field et al. 2009; Sayette, Shiffman, Tiffany, Niaura, Martin, & Shadel 2000). In the context of EMA, single-item measures of craving are useful in that they reduce reactivity to repeated measurements (e.g. increases in craving caused by prolonged rumination over craving experiences) and offer quick measures temporally close to the event of interest (e.g. an episode of temptation to drink alcohol). However, single-item measures of craving introduce a certain terminology (e.g. urges to drink) which might differ from how participants perceive their own cravings, and they might be less reliable than multi-item questionnaires and lack the breath of craving components assessed by multi-item questionnaires (e.g. desire to drink alcohol and inability to inhibit drinking). It should be noted thought that multi-item questionnaires of craving have distinct problems as well, as for example measuring constructs which are related to craving but also separable (e.g. expectations and intentions; for a review see Kavanagh, Statham, Feeney, Young, May, Andrade et al. 2013).

In addition, daily alcohol use was retrospectively recorded by participants at the end of each day or first thing the next morning. Retrospective reports, even after 24 hours, may underestimate real quantity of drinking, especially after heavier drinking (see for example Monk et al. 2015). Although reporting all alcohol consumption after it has happened in daily diaries reduces the number of assessments (more user-friendly) and have been previously shown to be effective (e.g. Boynton & Richman 2014), real-time logging of each drink might be preferable.

As discussed earlier, alcohol-cues reactivity and alcohol-use behaviour are expected to be different in social drinkers, compared to participants with alcohol use disorder, or participants seeking to reduce their consumption or abstain (e.g. social drinker are expected to have lighter cravings, more voluntary behaviour as opposed to compulsive, and no motivation to avoid alcohol). Even though the average quantity of typical weekly alcohol consumption reported by participants in this thesis was above the recommended healthy limits (14 units per week), and the average scores in the AUDIT questionnaire would qualify them as hazardous drinkers (AUDIT scores ≥ 8), individual participants were eligible for participation if they simply consumed at least one alcohol drink per day. For that reason, findings of the thesis might not be applicable to clinical or change-oriented samples. Present findings should be replicated with clinical populations before definite conclusion are drawn for the clinical utility of AB and craving in an EMA context.

Lastly, the AlSAp was programed more as a cognitive experiment that could be completed on smartphones, rather than as a commercial app. Although it satisfied the element of portability for an app, it fell short on the level of automaticity expected by an app. For example, the app did not automatically complete some procedural steps in the background (e.g. allocation of participant number and selection of experiment), it did not generate prompts for completion, and it did not automatically track participants' progress (e.g. detecting whether this was the first or the second session of the day). As a result, participants had to rely on their memory for completing the assessment and do several of the processes manually (e.g. select a number of options before initiation of the alcohol-Stroop task and enter the session number for each session). As it was stressed out in the focus group, this may have increased the burden for participants and reduced motivation and level of engagement with the app. Although compliance rates in the studies in this thesis were over 80% (which is the recommended minimum rate; Jones et al. 2019), they were lower compared to an EMA study where participants repeatedly completed a cognitive task on a smartphone and reminders were also provided (over 90%, Jones, Tiplady, et al. 2018).

7.5.2 Strengths

The thesis also has some overarching methodological strengths. The smartphone-based alcohol Stroop-task used to measure AB was tested for its psychometric properties and has been found to show acceptable internal reliability (see chapter three). This was an important assurance, given the common criticism for the unreliability of AB measures and its detrimental effects on AB research (Rodebaugh et al. 2016).

Furthermore, variations in AB and craving, and their association with alcohol use, were examined with EMA, which is considered the optimal choice when examining phenomena that are episodic in nature (Moskowitz & Young 2006; Shiffman 2009). Motivation to drink alcohol is a very transient concept that can change rapidly from one moment to the other. Also, using EMA, this thesis examined AB and craving as they were expressed in the real world. As such, I collected multiple rich datasets with multiple data-points, capturing the cognitive-motivational state of participants across different times of the day and across diverse motivational contexts.

Another strong point is that the rich datasets obtained from the EMA studies were analysed with advanced statistical analysis, such as multilevel modelling and ridge regression, which allows for a better understanding of the data and offer more robust results. For example, multilevel modelling took into account the clustered nature of the data (sessions within days within participants) and provided information on the contribution of each level of analysis to the observed variance in AB, craving and alcohol use. Moreover, ridge regression allowed examining highly correlated predictors (e.g. measures of craving taken within the same day, within the same participant) without risking bias (particularly inflated regression coefficients) caused by multicollinearity.

7.6 Future directions

Based on the present findings, a number of possibilities for future research have been revealed. Studies aimed at creating a program that would be able to detect increased chances of alcohol consumption (which in clinical populations could translate to high risk for relapse), and possibly deliver real-time interventions, might want to focus more on craving, rather than AB. These studies could explore how craving could best be measured in the real word (e.g. comparing a single-item measure with digital portable versions of multi-item craving questionnaires), as well as which alcohol-related outcomes (e.g. quantity or frequency of drinking, episodes of binge drinking, or relapse after a period of abstinence) would be more strongly correlated with craving.

Future studies could examine AB in the real world using eye-gaze indices, rather than reaction-time differences. Eye-tracking is a direct measure of AB which has been found to be more reliable than reaction-time measures (e.g. Christiansen, Mansfield, et al. 2015).

Although there will be many practical challenges (e.g. programing an app that supports eye-tracking technology and lending expensive portable eye-tracking equipment to participants to take it home), the appropriate technology (e.g. mobile eye-tracking glasses) is currently available to make this feasible.

It is also worth noting that this thesis examined AB and craving in situations in the real world when motivation to drink alcohol was expected to be high (i.e. temporal proximity to alcohol use, total availability of alcohol, temptation to drink). However, there are also other factors that have been found to affect participants' motivation to consume alcohol in the laboratory, and future studies could examine these factors in the real world. For example, studies could examine the effect on AB and craving of acute effects of alcohol, fluctuations in mood, and fluctuations in inhibitory control (see for example Jones, Tiplady, et al. 2018).

Another way forward would be for future studies to utilize to greater extends the sensory capabilities offered by smartphones. For example, geolocation sensors (GPS) could be used to examine the effect of the built environment (e.g. points of alcohol sale; see Kirchner et al. 2013) or advertising on participants' AB, craving, and alcohol use.

Finally, this thesis shows that it is feasible to explore the cognitive-motivational nature of alcohol use in the real world. It is true that laboratories, where many confounders can be controlled, would remain the gold standard for purer measures. However, replication (or not) of laboratory findings in the real world would enhance confidence on the generalizability of these findings in real life, where humans live and drink alcohol. This "reality-check" of laboratory findings should become a standard course of action in future research.

7.7 General conclusions

Findings of the thesis shed light to the way AB and craving are expressed in the real world, how they are affected by the motivational context, and the extent to which they are associated with alcohol consumption. First, findings did not support the hypothesis that AB is a motivational state. Therefore, the effect of the motivational context on AB might not be a valid explanation for the inconsistent results in the literature regarding the association between AB and alcohol use.

Arguably, rejection of the state hypothesis adds to a long list of findings undermining the relevance of AB in alcohol-use research (e.g. unreliability of common measures of AB, inconsistent association of AB with alcohol use, and limited evidence for the effectiveness of ABM). There might be a need to re-evaluate the usefulness of AB as a construct which could reliably provide information on the current motivation to drink alcohol and predict alcohol use.

However, it should be noted that present findings are based on non-clinical samples. Definite conclusions about the clinical relevance of AB could not be drawn until the state hypothesis of AB is also examined in dependent or help-seeking alcohol users in the real world. These studies would reveal whether present indications of the state hypothesis of AB in drug-dependent patients could be generalized to alcohol-dependent patients.

Second, it was found that simply asking about alcohol craving may be a better alternative to assessing AB, when attempting to predict alcohol use. The possibility to reliably predict alcohol use by just asking a single question about participants' current urges is an attractive idea, especially in the context of EMA. Implications are even more important in the clinical field, where a simple craving question could trigger real-time interventions to reduce chances of relapse.

Finally, this thesis is a testament to the usefulness of real-world behavioural research, EMA design, and smartphone-based behavioural assessment. It shows that cognitive-experimental investigations of alcohol use, which are traditionally considered laboratory based, can be conducted in the real world. It also showed that EMA offers a suitable methodological framework for real-world cognitive-experimental research. Lastly, this thesis highlights the usefulness of smartphones as a portable digital device able to host cognitive-behavioural measures.

Overall, previous research on AB, craving and alcohol use was extended beyond the laboratory and into the real world to further our understanding of the cognitive-motivational nature of alcohol use. Findings suggested that the nature of AB and it usefulness as a construct might have to be revaluated. They also suggested that craving might be an important component behind momentary decisions of individuals to seek and consume alcohol in the real world.

8. References

Abroms, L. C., Lee Westmaas, J., Bontemps-Jones, J., Ramani, R., & Mellerson, J. (2013). A Content Analysis of Popular Smartphone Apps for Smoking Cessation. *American Journal of Preventive Medicine, 45*(6), 732-736. doi:https://doi.org/10.1016/j.amepre.2013.07.008

Adams, S., Ataya, A. F., Attwood, A. S., & Munafo, M. R. (2012). Effects of acute alcohol consumption on alcohol-related cognitive biases in light and heavy drinkers are task-dependent. (2), 245-253. doi:10.1177/0269881111405355

Algom, D., Chajut, E., & Lev, S. (2004). A Rational Look at the Emotional Stroop Phenomenon: A Generic Slowdown, Not a Stroop Effect. (3), 323-338. doi:10.1037/0096-3445.133.3.323

Anderson, B. A. (2013). A value-driven mechanism of attentional selection. *Journal of vision, 13*(3), 1-16. doi:10.1167/13.3.7.

Anderson, B. A. (2015). Value-driven attentional priority is context specific. *Psychonomic bulletin & review, 22*(3), 750-756. doi:10.3758/s13423-014-0724-0

Anderson, B. A., Laurent, P. A., & Yantis, S. (2011). Value-driven attentional capture. *Proceedings of the National Academy of Sciences, 108*(25), 10367-10371. doi:10.1073/pnas.1104047108

APA. (2013). *Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition*. Arlingtion, VA: American Psychiatric Association.

Ataya, A. F., Adams, S., Mullings, E., Cooper, R. M., Attwood, A. S., & Munafò, M. R. (2012). Internal reliability of measures of substance-related cognitive bias. *Drug and alcohol dependence, 121*(1), 148-151. doi:10.1016/j.drugalcdep.2011.08.023

Atwell, K., Abraham, C., & Duka, T. (2011). A Parsimonious, Integrative Model of Key Psychological Correlates of UK University Students' Alcohol Consumption. *Alcohol and Alcoholism, 46*(3), 253-260. doi:10.1093/alcalc/agr016

Babor, T. F., Higgins-Biddle, J. C., Saunders, J. B., & Monteiro, M. G. (2001). *AUDIT: The alcohol use disorders identification test: Guidelines for use in primary health care* (2nd ed.). Geneva: Word Health Organisation.

Barkby, H., Dickson, J. M., Roper, L., & Field, M. (2012). To Approach or Avoid Alcohol? Automatic and Self-Reported Motivational Tendencies in Alcohol Dependence. *Alcoholism: Clinical and Experimental Research, 36*(2), 361-368. doi:10.1111/j.1530-0277.2011.01620.x

Bauer, D., & Cox, W. M. (1998). Alcohol-related words are distracting to both alcohol abusers and non-abusers in the Stroop colour-naming task. *Addiction, 93*(10), 1539-1542.

Behrendt, S., Wittchen, H.-U., Höfler, M., Lieb, R., & Beesdo, K. (2009). Transitions from first substance use to substance use disorders in adolescence: is early onset associated with a rapid escalation? *Drug & Alcohol Dependence, 99*(1), 68-78. doi:10.1016/j.drugalcdep.2008.06.014

Berridge, K. C., & Kringelbach, M. L. (2008). Affective neuroscience of pleasure: reward in humans and animals. *Psychopharmacology, 199*(3), 457-480. doi:10.1007/s00213-008-1099-6

Berridge, K. C., Robinson, T. E., & Aldridge, J. W. (2009). Dissecting components of reward:'liking','wanting', and learning. *Current opinion in pharmacology, 9*(1), 65-73. doi:10.1016/j.coph.2008.12.014

Boendermaker, W. J., Sanchez Maceiras, S., Boffo, M., & Wiers, R. W. (2016). Attentional Bias Modification With Serious Game Elements: Evaluating the Shots Game. *JMIR Serious Games, 4*(2), e20. doi:10.2196/games.6464

Boutelle, K. N., & Kirschenbaum, D. S. (1998). Further Support for Consistent Self-Monitoring as a Vital Component of Successful Weight Control. *Obesity Research, 6*(3), 219-224. doi:10.1002/j.1550-8528.1998.tb00340.x

Boynton, M. H., & Richman, L. S. (2014). An online daily diary study of alcohol use using Amazon's Mechanical Turk. *Drug & Alcohol Review, 33*(4), 456-461. doi:10.1111/dar.12163

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology, 3*(2), 77-101. doi:10.1191/1478088706qp063oa

Brown, H. R., Zeidman, P., Smittenaar, P., Adams, R. A., McNab, F., Rutledge, R. B., & Dolan, R. J. (2014). Crowdsourcing for Cognitive Science – The Utility of Smartphones. *PLOS ONE, 9*(7), e100662. doi:10.1371/journal.pone.0100662

Bruce, G., & Jones, B. T. (2004). A pictorial Stroop paradigm reveals an alcohol attentional bias in heavier compared to lighter social drinkers. *Journal of Psychopharmacology, 18*(4), 527-533. doi:10.1177/0269881104047280

Buckner, J. D., Schmidt, N. B., Lang, A. R., Small, J. W., Schlauch, R. C., & Lewinsohn, P. M. (2008). Specificity of social anxiety disorder as a risk factor for alcohol and cannabis dependence. *Journal of psychiatric research, 42*(3), 230-239. doi:0.1016/j.jpsychires.2007.01.002

Capon, H., Hall, W., Fry, C., & Carter, A. (2016). Realising the technological promise of smartphones in addiction research and treatment: An ethical review. *International Journal of Drug Policy, 36*, 47-57. doi:https://doi.org/10.1016/j.drugpo.2016.05.013

Carroll, K. M. (1998). *Therapy manuals for drug addiction, manual 1: a cognitive-behavioral approach: treating cocaine addiction*. Maryland: National Institute on Drug Abuse.

Charmaz, K. (1996). Grounded Theory. In S. J.A., H. R, & V. L. L (Eds.), *Rethinking methods in Psychology* (pp. 27-49). London: Sage Publications.

Christiansen, P., & Bloor, J. F. (2014). Individualised but not general alcohol Stroop predicts alcohol use. *Drug and alcohol dependence, 134*, 410-413. doi:http://dx.doi.org/10.1016/j.drugalcdep.2013.10.021

Christiansen, P., Cole, J. C., & Field, M. (2012). Ego depletion increases ad-lib alcohol consumption: Investigating cognitive mediators and moderators. *Experimental and Clinical Psychopharmacology, 20*(2), 118-128. doi:10.1037/a0026623

Christiansen, P., Mansfield, R., Duckworth, J., Field, M., & Jones, A. (2015). Internal reliability of the alcohol-related visual probe task is increased by utilising personalised stimuli and eye-tracking. *Drug and alcohol dependence, 155*(Supplement C), 170-174. doi:https://doi.org/10.1016/j.drugalcdep.2015.07.672

Christiansen, P., Schoenmakers, T. M., & Field, M. (2015). Less than meets the eye: reappraising the clinical relevance of attentional bias in addiction. *Addictive behaviors, 44*, 43-50. doi:http://dx.doi.org/10.1016/j.addbeh.2014.10.005

Christiansen, P., Townsend, G., Knibb, G., & Field, M. (2017). Bibi ergo sum: the effects of a placebo and contextual alcohol cues on motivation to drink alcohol. *Psychopharmacology, 234*(5), 827-835. doi:10.1007/s00213-016-4518-0

Clerkin, E. M., Magee, J. C., Wells, T. T., Beard, C., & Barnett, N. P. (2016). Randomized controlled trial of attention bias modification in a racially diverse, socially anxious, alcohol dependent sample. *Behaviour Research and Therapy, 87*, 58-69. doi:https://doi.org/10.1016/j.brat.2016.08.010

Cox, M., Intrilligator, J., & Hillier, C. (2015). Chimpshop and alcohol reduction – using technology to change behaviour. *Perspectives in Public Health, 135*(3), 126-127. doi:10.1177/1757913915580926

Cox, W. M., Blount, J. P., & Rozak, A. M. (2000). Alcohol abusers' and nonabusers' distraction by alcohol and concern-related stimuli. *The American journal of drug and alcohol abuse, 26*(3), 489-495.

Cox, W. M., Brown, M. A., & Rowlands, L. J. (2003). The effect of alcohol cue exposre on non-dependent drinkers' attentional bias for alcohol-related stimuli. *Alcohol and Alcoholism, 38*(1), 45-49. doi:10.1093/alcalc/agg010

Cox, W. M., Fadardi, J. S., Hosier, S. G., & Pothos, E. M. (2015). Differential effects and temporal course of attentional and motivational training on excessive drinking. *Experimental and Clinical Psychopharmacology, 23*(6), 445-454. doi:10.1037/pha0000038

Cox, W. M., Fadardi, J. S., & Pothos, E. M. (2006). The addiction-stroop test: Theoretical considerations and procedural recommendations. *Psychological bulletin, 132*(3), 443-476. doi:10.1037/0033-2909.132.3.443

Cox, W. M., Hogan, L. M., Kristian, M. R., & Race, J. H. (2002). Alcohol attentional bias as a predictor of alcohol abusers’ treatment outcome. *Drug and alcohol dependence, 68*(3), 237-243.

Cox, W. M., Pothos, E., & Hosier, S. (2007). Cognitive-motivational predictors of excessive drinkers’ success in changing. *Psychopharmacology, 192*(4), 499-510. doi:10.1007/s00213-007-0736-9

Crane, D., Garnett, C., Brown, J., West, R., & Michie, S. (2015). Behavior change techniques in popular alcohol reduction apps: content analysis. *Journal Of Medical Internet Research, 17*(5), e118-e118. doi:10.2196/jmir.4060

Crane, D., Garnett, C., Brown, J., West, R., & Michie, S. (2017). Factors Influencing Usability of a Smartphone App to Reduce Excessive Alcohol Consumption: Think Aloud and Interview Studies. *Frontiers in Public Health, 5*(39). doi:10.3389/fpubh.2017.00039

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. doi:10.1371/journal.pone.0162226

Cule, E., & De Iorio, M. (2013). Ridge regression in prediction problems: automatic choice of the ridge parameter. *Genetic epidemiology, 37*(7), 704-714. doi:10.1002/gepi.21750

Dar, R., Rosen-Korakin, N., Shapira, O., Gottlieb, Y., & Frenk, H. (2010). The craving to smoke in flight attendants: relations with smoking deprivation, anticipation of smoking, and actual smoking. *Journal of Abnormal Psychology, 119*(1), 248-253. doi:10.1037/a0017778

Dar, R., Stronguin, F., Marouani, R., Krupsky, M., & Frenk, H. (2005). Craving to smoke in orthodox Jewish smokers who abstain on the Sabbath: A comparison to a baseline and a forced abstinence workday. *Psychopharmacology, 183*(3), 294-299. doi:10.1007/s00213-005-0192-3

Dawson, D. A., Goldstein, R. B., Patricia Chou, S., June Ruan, W., & Grant, B. F. (2008). Age at first drink and the first incidence of adult-onset DSM-IV alcohol use disorders. *Alcoholism: Clinical and Experimental Research, 32*(12), 2149-2160. doi:10.1111/j.1530-0277.2008.00806.x

De Houwer, J., Crombez, G., Koster, E. H., & De Beul, N. (2004). Implicit alcohol-related cognitions in a clinical sample of heavy drinkers. *Journal of behavior therapy and experimental psychiatry, 35*(4), 275-286. doi:10.1016/j.jbtep.2004.05.001

De Luca, M. A. (2014). Habituation of the responsiveness of mesolimbic and mesocortical dopamine transmission to taste stimuli. *Frontiers in Integrative Neuroscience, 8*(21). doi:10.3389/fnint.2014.00021

de Meneses-Gaya, C., Zuardi, A. W., Loureiro, S. R., & Crippa, J. A. S. (2009). Alcohol Use Disorders Identification Test (AUDIT): An updated systematic review of

psychometric properties. *Psychology & Neuroscience, 2*(1), 83-97. doi:10.3922/j.psns.2009.1.12
Del Boca, F. K., & Darkes, J. (2003). The validity of self-reports of alcohol consumption: state of the science and challenges for research. *Addiction, 98*, 1-12. doi:10.1046/j.1359-6357.2003.00586.x
Dennison, L., Morrison, L., Conway, G., & Yardley, L. (2013). Opportunities and challenges for smartphone applications in supporting health behavior change: qualitative study. *Journal Of Medical Internet Research, 15*(4), e86-e86. doi:10.2196/jmir.2583
Di Chiara, G. (2002). Nucleus accumbens shell and core dopamine: differential role in behavior and addiction. *Behavioural Brain Research, 137*(1), 75-114. doi:https://doi.org/10.1016/S0166-4328(02)00286-3
Di Lemma, L. C. G., Dickson, J. M., Jedras, P., Roefs, A., & Field, M. (2015). Priming of conflicting motivational orientations in heavy drinkers: robust effects on self-report but not implicit measures. *Frontiers in Psychology, 6*. doi:10.3389/fpsyg.2015.01465
Donker, T., Petrie, K., Proudfoot, J., Clarke, J., Birch, M.-R., & Christensen, H. (2013). Smartphones for Smarter Delivery of Mental Health Programs: A Systematic Review. *J Med Internet Res, 15*(11), e247. doi:10.2196/jmir.2791
Duka, T., & Townshend, J. M. (2004). The priming effect of alcohol pre-load on attentional bias to alcohol-related stimuli. *Psychopharmacology, 176*(3/4), 353-362. doi:10.1007/s00213-004-1906-7
Dulin, P. L., & Gonzalez, V. M. (2017). Smartphone-based, momentary intervention for alcohol cravings amongst individuals with an alcohol use disorder. *Psychology of Addictive Behaviors, 31*(5), 601-607. doi:10.1037/adb0000292
Elfeddali, I., Bolman, C., Pronk, T., & Wiers, R. W. (2016). A Randomized Controlled Trial of Web-Based Attentional Bias Modification to Help Smokers Quit. *Health Psychology, 35*(8), 870-880. doi:http://dx.doi.org/10.1037/hea0000346
Ernst, L. H., Plichta, M. M., Dresler, T., Zesewitz, A. K., Tupak, S. V., Haeussinger, F. B., Fischer, M., Polak, T., Fallgatter, A. J., & Ehlis, A. C. (2014). Prefrontal correlates of approach preferences for alcohol stimuli in alcohol dependence. *Addiction Biology, 19*(3), 497-508. doi:10.1111/adb.12005
Fadardi, J. S., & Cox, W. M. (2008). Alcohol-attentional bias and motivational structure as independent predictors of social drinkers' alcohol consumption. *Drug and alcohol dependence, 97*(3), 247-256. doi:https://doi.org/10.1016/j.drugalcdep.2008.03.027
Fadardi, J. S., & Cox, W. M. (2009). Reversing the sequence: reducing alcohol consumption by overcoming alcohol attentional bias. *Drug and alcohol dependence, 101*(3), 137-145.
Fatseas, M., Serre, F., Alexandre, J.-M., Debrabant, R., Auriacombe, M., & Swendsen, J. (2015). Craving and substance use among patients with alcohol, tobacco, cannabis or heroin addiction: a comparison of substance- and person-specific cues. *Addiction, 110*(6), 1035-1042. doi:10.1111/add.12882
Field, A. (2013). *Discovering statistics using IBM SPSS statistics* (4th ed.). London: Sage.
Field, M., Caren, R., Fernie, G., & De Houwer, J. (2011). Alcohol Approach Tendencies in Heavy Drinkers: Comparison of Effects in a Relevant Stimulus-Response Compatibility Task and an Approach/Avoidance Simon Task. *Psychology of Addictive Behaviors, 25*(4), 697-701. doi:10.1037/a0023285
Field, M., & Christiansen, P. (2012). Commentary on,'Internal reliability of measures of substance-related cognitive bias'. *Drug and alcohol dependence, 124*(3), 189-190. doi:10.1016/j.drugalcdep.2012.02.009
Field, M., Christiansen, P., Cole, J., & Goudie, A. (2007). Delay discounting and the alcohol Stroop in heavy drinking adolescents. *Addiction, 102*(4), 579-586. doi:10.1111/j.1360-0443.2007.01743.x

Field, M., & Cox, W. M. (2008). Attentional bias in addictive behaviors: a review of its development, causes, and consequences. *Drug and alcohol dependence, 97*(1), 1-20. doi:10.1016/j.drugalcdep.2008.03.030

Field, M., Di Lemma, L., Christiansen, P., & Dickson, J. (2017). Automatic Avoidance Tendencies for Alcohol Cues Predict Drinking After Detoxification Treatment in Alcohol Dependence. *Psychology of Addictive Behaviors, 31*(2), 171-179. doi:10.1037/adb0000232

Field, M., Duka, T., Eastwood, B., Child, R., Santarcangelo, M., & Gayton, M. (2007). Experimental manipulation of attentional biases in heavy drinkers: do the effects generalise? *Psychopharmacology, 192*(4), 593-608. doi:10.1007/s00213-007-0760-9

Field, M., & Eastwood, B. (2005). Experimental manipulation of attentional bias increases the motivation to drink alcohol. *Psychopharmacology, 183*(3), 350-357. doi:10.1007/s00213-005-0202-5

Field, M., Hogarth, L., Bleasdale, D., Wright, P., Fernie, G., & Christiansen, P. (2011). Alcohol expectancy moderates attentional bias for alcohol cues in light drinkers. *Addiction, 106*(6), 1097-1103. doi:10.1111/j.1360-0443.2011.03412.x

Field, M., Kiernan, A., Eastwood, B., & Child, R. (2008). Rapid approach responses to alcohol cues in heavy drinkers. *Journal of behavior therapy and experimental psychiatry, 39*(3), 209-218. doi:https://doi.org/10.1016/j.jbtep.2007.06.001

Field, M., Marhe, R., & Franken, I. H. A. (2014). The clinical relevance of attentional bias in substance use disorders. *CNS spectrums, 19*(03), 225-230. doi:10.1017/S1092852913000321

Field, M., Mogg, K., & Bradley, B. (2005). Craving and cognitive biases for alcohol cues in social drinkers. *Alcohol and Alcoholism, 40*(6), 504-510. doi:10.1093/alcalc/agh213

Field, M., Mogg, K., Mann, B., Bennett, G. A., & Bradley, B. P. (2013). Attentional biases in abstinent alcoholics and their association with craving. *Psychology of Addictive Behaviors, 27*(1), 71-80. doi:10.1037/a0029626

Field, M., Mogg, K., Zetteler, J., & Bradley, B. P. (2004). Attentional biases for alcohol cues in heavy and light social drinkers: the roles of initial orienting and maintained attention. *Psychopharmacology, 176*(1), 88-93. doi:10.1007/s00213-004-1855-1

Field, M., Munafò, M. R., & Franken, I. H. (2009). A meta-analytic investigation of the relationship between attentional bias and subjective craving in substance abuse. *Psychological bulletin, 135*(4), 589-607. doi:10.1037/a0015843

Field, M., & Powell, H. (2007). Stress increases attentional bias for alcohol cues in social drinkers who drink to cope. *Alcohol and Alcoholism, 42*(6), 560-566. doi:10.1093/alcalc/agm064

Field, M., & Quigley, M. (2009). Mild Stress Increases Attentional Bias in Social Drinkers Who Drink to Cope: A Replication and Extension. (5), 312-319. doi:10.1037/a0017090

Field, M., Werthmann, J., Franken, I., Hofmann, W., Hogarth, L., & Roefs, A. (2016). The role of attentional bias in obesity and addiction. *Health Psychology, 35*(8), 767-780. doi:10.1037/hea0000405

Field, M., Wiers, R. W., Christiansen, P., Fillmore, M. T., & Verster, J. C. (2010). Acute alcohol effects on inhibitory control and implicit cognition: implications for loss of control over drinking. *Alcoholism: Clinical and Experimental Research, 34*(8), 1346-1352. doi:10.1111/j.1530-0277.2010.01218.x

Foran, H. M., & O'Leary, K. D. (2008). Alcohol and intimate partner violence: A meta-analytic review. *Clinical psychology review, 28*(7), 1222-1234. doi:https://doi.org/10.1016/j.cpr.2008.05.001

Franken, I. H. (2003). Drug craving and addiction: integrating psychological and neuropsychopharmacological approaches. *Progress in Neuro-Psychopharmacology and Biological Psychiatry, 27*(4), 563-579. doi:10.1016/S0278-5846(03)00081-2

Freedman, M. J., Lester, K. M., McNamara, C., Milby, J. B., & Schumacher, J. E. (2006). Cell phones for ecological momentary assessment with cocaine-addicted homeless patients in treatment. *Journal of Substance Abuse Treatment, 30*(2), 105-111. doi:10.1016/j.jsat.2005.10.005

Garland, E., Boettiger, C., Gaylord, S., Chanon, V., & Howard, M. (2012). Mindfulness is Inversely Associated with Alcohol Attentional Bias Among Recovering Alcohol-Dependent Adults. *Cognitive Therapy & Research, 36*(5), 441-450. doi:10.1007/s10608-011-9378-7

Garland, E. L., Franken, I. H., & Howard, M. O. (2012). Cue-elicited heart rate variability and attentional bias predict alcohol relapse following treatment. *Psychopharmacology, 222*(1), 17-26. doi:10.1007/s00213-011-2618-4.

Garland, E. L., Franken, I. H., Sheetz, J. J., & Howard, M. O. (2012). Alcohol attentional bias is associated with autonomic indices of stress-primed alcohol cue-reactivity in alcohol-dependent patients. *Experimental and Clinical Psychopharmacology, 20*(3). doi:10.1037/a0027199

Giesbrecht, N., Cukier, S., & Steeves, D. A. N. (2010). Collateral damage from alcohol: implications of 'second-hand effects of drinking' for populations and health priorities. *Addiction, 105*(8), 1323-1325. doi:10.1111/j.1360-0443.2009.02884.x

Gladwin, T. E. (2017). Attentional bias variability and cued attentional bias for alcohol stimuli. *Addiction Research and Theory, 25*(1), 32-38. doi:10.1080/16066359.2016.1196674

Gladwin, T. E., Mohr, S. E., & Wiers, R. W. (2014). The potential role of temporal dynamics in approach biases: delay-dependence of a general approach bias in an alcohol approach-avoidance task. *Frontiers in Psychology, 5*. doi:10.3389/fpsyg.2014.01398

Goldstein, H. (2011). *Multilevel statistical models* (4th ed.). West Sussex: John Wiley & Sons.

Grant, B. F., Goldstein, R. B., Saha, T. D., & et al. (2015). Epidemiology of dsm-5 alcohol use disorder: Results from the national epidemiologic survey on alcohol and related conditions iii. *JAMA Psychiatry, 72*(8), 757-766. doi:10.1001/jamapsychiatry.2015.0584

Grant, V. V., Stewart, S. H., & Birch, C. D. (2007). Impact of positive and anxious mood on implicit alcohol-related cognitions in internally motivated undergraduate drinkers. *Addictive behaviors, 32*(10), 2226-2237. doi:10.1016/j.addbeh.2007.02.012

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.

Groefsema, M., Engels, R., Kuntsche, E., Smit, K., & Luijten, M. (2016). Cognitive Biases for Social Alcohol-Related Pictures and Alcohol Use in Specific Social Settings: An Event-Level Study. *Alcoholism: Clinical and Experimental Research, 40*(9), 2001-2010. doi:10.1111/acer.13165

Gustafson, D. H., Shaw, B. R., Isham, A., Baker, T., Boyle, M. G., & Levy, M. (2011). Explicating an Evidence-Based, Theoretically Informed, Mobile Technology-Based System to Improve Outcomes for People in Recovery for Alcohol Dependence. *Substance Use & Misuse, 46*(1), 96-111. doi:10.3109/10826084.2011.521413

Harrison, N. R., & McCann, A. (2014). The effect of colour and size on attentional bias to alcohol-related pictures. *Psicologica, 35*(1), 39-48.

Hayes, A. F. (2006). A Primer on Multilevel Modeling. (4), 385-410. doi:doi:10.1111/j.1468-2958.2006.00281.x

Helf, C., & Hlavacs, H. (2016). Apps for life change: Critical review and solution directions. *Entertainment Computing, 14*, 17-22. doi:https://doi.org/10.1016/j.entcom.2015.07.001

Heron, K. E., & Smyth, J. M. (2010). Ecological momentary interventions: Incorporating mobile technology into psychosocial and health behaviour treatments. *British Journal of Health Psychology, 15*(1), 1-39. doi:10.1348/135910709X466063

Hester, R., Dixon, V., & Garavan, H. (2006). A consistent attentional bias for drug-related material in active cocaine users across word and picture versions of the emotional Stroop task. *Drug and alcohol dependence, 81*(3), 251-257. doi:10.1016/j.drugalcdep.2005.07.002

Hingson, R. W., Heeren, T., & Winter, M. R. (2006). Age at drinking onset and alcohol dependence: age at onset, duration, and severity. *Archives of pediatrics & adolescent medicine, 160*(7), 739-746.

Hingson, R. W., & Zha, W. (2009). Age of drinking onset, alcohol use disorders, frequent heavy drinking, and unintentionally injuring oneself and others after drinking. *Pediatrics, 123*(6), 1477-1484. doi:doi:10.1542/peds.2008-2176

Hoeppner, B. B., Stout, R. L., Jackson, K. M., & Barnett, N. P. (2010). How good is fine-grained Timeline Follow-back data? Comparing 30-day TLFB and repeated 7-day TLFB alcohol consumption reports on the person and daily level. *Addictive behaviors, 35*(12), 1138-1143. doi:https://doi.org/10.1016/j.addbeh.2010.08.013

Houben, K., Nosek, B. A., & Wiers, R. W. (2010). Seeing the forest through the trees: A comparison of different IAT variants measuring implicit alcohol associations. *Drug & Alcohol Dependence, 106*(2), 204-211. doi:10.1016/j.drugalcdep.2009.08.016

Houben, K., & Wiers, R. W. (2006). Assessing implicit alcohol associations with the Implicit Association Test: Fact or artifact? *Addictive behaviors, 31*(8), 1346-1362. doi:doi:10.1016/j.addbeh.2005.10.009

Houben, K., & Wiers, R. W. (2007a). Are drinkers implicitly positive about drinking alcohol? Personalizing the alcohol-IAT to reduce negative extrapersonal contamination. *Alcohol and Alcoholism, 42*(4), 301-307. doi:10.1093/alcalc/agm015

Houben, K., & Wiers, R. W. (2007b). Personalizing the alcohol-IAT with individualized stimuli: Relationship with drinking behavior and drinking-related problems. *Addictive behaviors, 32*(12), 2852-2864. doi:https://doi.org/10.1016/j.addbeh.2007.04.022

Houben, K., & Wiers, R. W. (2008a). Implicitly positive about alcohol? Implicit positive associations predict drinking behavior. *Addictive behaviors, 33*(8), 979-986. doi:https://doi.org/10.1016/j.addbeh.2008.03.002

Houben, K., & Wiers, R. W. (2008b). Measuring implicit alcohol associations via the Internet: Validation of Web-based implicit association tests. *Behavior research methods, 40*(4), 1134-1143. doi:10.3758/BRM.40.4.1134

Houben, K., & Wiers, R. W. (2009). Beer makes the heart grow fonder: single-target implicit attitudes toward beer but not alcohol are related to drinking behaviour in regular beer drinkers. *Netherlands Journal of Psychology, 65*(1), 10-21.

Hu, X., Oroszi, G., Chun, J., Smith, T. L., Goldman, D., & Schuckit, M. A. (2005). An expanded evaluation of the relationship of four alleles to the level of response to alcohol and the alcoholism risk. *Alcoholism: Clinical and Experimental Research, 29*(1), 8-16. doi:10.1097/01.ALC.0000150008.68473.62

Hufford, M. R., Shields, A. L., Shiffman, S., Paty, J., & Balabanis, M. (2002). Reactivity to ecological momentary assessment: An example using undergraduate problem drinkers. *Psychology of Addictive Behaviors, 16*(3), 205-211. doi:10.1037/0893-164X.16.3.205

Jędras, P., Jones, A., & Field, M. (2013). The role of anticipation in drug addiction and reward. *Neuroscience and Neuroeconomics, 2014*(1), 1-10.

Jędras, P., Jones, A., Stancak, A., & Field, M. (2018). The effects of reward and loss anticipation on attentional bias for reward-related stimuli. *Appetite, 133*, 93-100. doi:https://doi.org/10.1016/j.appet.2018.10.007

Johnsen, B. H., Laberg, J. C., Cox, W. M., Vaksdal, A., & Hugdahl, K. (1994). Alcoholic subjects' attentional bias in the processing of alcohol-related words. *Psychology of Addictive Behaviors, 8*(2), 111-115.

Jones, A., Christiansen, P., & Field, M. (2018). Failed attempts to improve the reliability of the alcohol visual probe task following empirical recommendations. *Psychology of Addictive Behaviors, 32*(8), 922-932. doi:http://dx.doi.org/10.1037/adb0000414

Jones, A., Hogarth, L., Christiansen, P., Rose, A. K., Martinovic, J., & Field, M. (2012). Reward expectancy promotes generalized increases in attentional bias for rewarding stimuli. *The Quarterly Journal of Experimental Psychology, 65*(12), 2333-2342. doi:10.1080/17470218.2012.686513

Jones, A., Remmerswaal, D., Verveer, I., Robinson, E., Franken, I. H. A., Wen, C. K. F., & Field, M. (2019). Compliance with ecological momentary assessment protocols in substance users: a meta-analysis. *Addiction, 114*(4), 609-619. doi:10.1111/add.14503

Jones, A., Tiplady, B., Houben, K., Nederkoorn, C., & Field, M. (2018). Do daily fluctuations in inhibitory control predict alcohol consumption? An ecological momentary assessment study. *Psychopharmacology, 235*(5), 1487-1496. doi:10.1007/s00213-018-4860-5

Jones, B. C., Jones, B. T., Blundell, L., & Bruce, G. (2002). Social users of alcohol and cannabis who detect substance-related changes in a change blindness paradigm report higher levels of use than those detecting substance-neutral changes. *Psychopharmacology, 165*(1), 93-96. doi:10.1007/s00213-002-1264-2

Jones, B. T., Bruce, G., Livingstone, S., & Reed, E. (2006). Alcohol-related attentional bias in problem drinkers with the flicker change blindness paradigm. *Psychology of Addictive Behaviors, 20*(2), 171-177. doi:10.1037/0893-164X.20.2.171

Kadden, R., Caroll, K., Donovan, D., Cooney, N., Monti, P., Abrams, D., Litt, M., & Hester, R. (1995). *Cognitive-behavioral coping skills therapy manual: A clinical research guide for therapists treating individuals with alcohol abuse and dependence* (Vol. 3). Maryland: National Institute of Health.

Kavanagh, D. J., Statham, D. J., Feeney, G. F., Young, R. M., May, J., Andrade, J., & Connor, J. P. (2013). Measurement of alcohol craving. *Addictive behaviors, 38*(2), 1572-1584. doi:http://dx.doi.org/10.1016/j.addbeh.2012.08.004

Kersbergen, I., Woud, M. L., & Field, M. (2015). The validity of different measures of automatic alcohol action tendencies. *Psychology of Addictive Behaviors, 29*(1), 225-230. doi:10.1037/adb0000009

Kerst, W. F., & Waters, A. J. (2014). Attentional retraining administered in the field reduces smokers' attentional bias and craving. *Health Psychology, 33*(10), 1232-1240. doi:10.1037/a0035708

Keyes, K. M., Hatzenbuehler, M. L., & Hasin, D. S. (2011). Stressful life experiences, alcohol consumption, and alcohol use disorders: the epidemiologic evidence for four main types of stressors. *Psychopharmacology, 218*(1), 1-17. doi:10.1007/s00213-011-2236-1

Kindt, M., & Brosschot, J. F. (1997). Phobia-related cognitive bias for pictorial and linguistic stimuli. *Journal of Abnormal Psychology, 106*(4), 644-648.

Kirchner, T. R., Cantrell, J., Anesetti-Rothermel, A., Ganz, O., Vallone, D. M., & Abrams, D. B. (2013). Geospatial exposure to point-of-sale tobacco: Real-time craving and

smoking-cessation outcomes. *American Journal of Preventive Medicine, 45*(4), 379-385. doi:10.1016/j.amepre.2013.05.016
Klein, A. A. (2007). Suppression-induced hyperaccessibility of thoughts in abstinent alcoholics: A preliminary investigation. *Behaviour Research and Therapy, 45*(1), 169-177. doi:10.1016/j.brat.2005.12.012
Kline, P. (1999). *The Handbook of Psychological Testing* (2nd ed.). LondonRoutledge.
Laslett, A.-M., Catalano, P., Chikritzhs, T., Dale, C., Doran, C., Ferris, J., Jainullabudeen, T., Livingston, M., Matthews, S., & Mugavin, J. (2010). *The range and magnitude of alcohol's harm to others*. Victoria: AER Centre for Alcohol Policy Research, Turning Point Alcohol and Drug Centre, Eastern Health.
Loeber, S., Vollstädt-Klein, S., Goltz, C. V. D., Flor, H., Mann, K., & Kiefer, F. (2009). Clinical Study: Attentional bias in alcohol-dependent patients: the role of chronicity and executive functioning. *Addiction Biology, 14*(2), 194-203. doi:doi:10.1111/j.1369-1600.2009.00146.x
Lubman, D. I., Yücel, M., & Pantelis, C. (2004). Addiction, a condition of compulsive behaviour? Neuroimaging and neuropsychological evidence of inhibitory dysregulation. *Addiction, 99*(12), 1491-1502. doi:10.1111/j.1360-0443.2004.00808.x
Lusher, J., Chandler, C., & Ball, D. (2004). Alcohol dependence and the alcohol Stroop paradigm: evidence and issues. *Drug and alcohol dependence, 75*(3), 225-231. doi:10.1016/j.drugalcdep.2004.03.004
Maas, C., & Hox, J. (2005). Sufficient Sample Sizes for Multilevel Modeling. *Methodology, 1*(3), 86-92. doi:10.1027/1614-2241.1.3.86
MacLeod, C., Mathews, A., & Tata, P. (1986). Attentional bias in emotional disorders. *Journal of Abnormal Psychology, 95*(1), 15-20. doi:10.1037/0021-843X.95.1.15
Mahmud, M. S., Fang, H., Carreiro, S., Wang, H., & Boyer, E. W. (2018). Wearables technology for drug abuse detection: A survey of recent advancement. *Smart Health*. doi:https://doi.org/10.1016/j.smhl.2018.09.002
Manchery, L., Yarmush, D. E., Luehring-Jones, P., & Erblich, J. (2017). Attentional bias to alcohol stimuli predicts elevated cue-induced craving in young adult social drinkers. *Addictive behaviors, 70*, 14-17. doi:10.1016/j.addbeh.2017.01.035
Månsson, K., & Shukur, G. (2011). On Ridge Parameters in Logistic Regression. *Communications in Statistics - Theory and Methods, 40*(18), 3366-3381. doi:10.1080/03610926.2010.500111
Marhe, R., Waters, A. J., van de Wetering, B. J. M., & Franken, I. H. A. (2013). Implicit and explicit drug-related cognitions during detoxification treatment are associated with drug relapse: an ecological momentary assessment study. *Journal of consulting and clinical psychology, 81*(1), 1-12. doi:10.1037/a0030754
Marks, K. R., Pike, E., Stoops, W. W., & Rush, C. R. (2014). Test–retest reliability of eye tracking during the visual probe task in cocaine-using adults. *Drug and alcohol dependence, 145*, 235-237. doi:https://doi.org/10.1016/j.drugalcdep.2014.09.784
Marks, K. R., Roberts, W., Stoops, W. W., Pike, E., Fillmore, M. T., & Rush, C. R. (2014). Fixation time is a sensitive measure of cocaine cue attentional bias. *Addiction, 109*(9), 1501-1508. doi:10.1111/add.12635
Martino, S. C., Kovalchik, S. A., Collins, R. L., Becker, K. M., Shadel, W. G., & D'Amico, E. J. (2016). Ecological momentary assessment of the association between exposure to alcohol advertising and early adolescents' beliefs about alcohol. *Journal of Adolescent Health, 58*(1), 85-91. doi:10.1016/j.jadohealth.2015.08.010.
Mathôt, S., Schreij, D., & Theeuwes, J. (2012). OpenSesame: An open-source, graphical experiment builder for the social sciences. *Behavior research methods, 44*(2), 314-324. doi:10.3758/s13428-011-0168-7

McClernon, F. J., & Choudhury, R. R. (2013). I am your Smartphone, and I Know you are about to Smoke: the application of mobile Sensing and Computing approaches to Smoking research and treatment. *nicotine & tobacco research, 15*(10), 1651-1654. doi:10.1093/ntr/ntt054

McGeary, J. E., Meadows, S. P., Amir, N., & Gibb, B. E. (2014). Computer-delivered, home-based, attentional retraining reduces drinking behavior in heavy drinkers. *Psychology of Addictive Behaviors, 28*(2), 559-562. doi:10.1037/a0036086

McManus, S., Bebbington, P., Jenkins, R., & Brugha, T. (Eds.). (2016). *Mental health and wellbeing in England: Adult psychiatric morbidity survey 2014*. Leeds: NHS Digital.

Miller, M. A., & Fillmore, M. T. (2010). The effect of image complexity on attentional bias towards alcohol-related images in adult drinkers. *Addiction, 105*(5), 883-890. doi:10.1111/j.1360-0443.2009.02860.x.

Miller, M. A., & Fillmore, M. T. (2011). Persistence of attentional bias toward alcohol-related stimuli in intoxicated social drinkers. *Drug & Alcohol Dependence, 117*(2), 184-189. doi:10.1016/j.drugalcdep.2011.01.016

Mohapatra, S., Patra, J., Popova, S., Duhig, A., & Rehm, J. (2010). Social cost of heavy drinking and alcohol dependence in high-income countries. *International Journal of Public Health, 55*(3), 149-157. doi:10.1007/s00038-009-0108-9

Monem, R. G., & Fillmore, M. T. (2017). Measuring heightened attention to alcohol in a naturalistic setting: A validation study. *Experimental and Clinical Psychopharmacology, 25*(6), 496-502. doi:http://dx.doi.org/10.1037/pha0000157

Monk, R. L., Heim, D., Qureshi, A., & Price, A. (2015). “I Have No Clue What I Drunk Last Night” Using Smartphone Technology to Compare In-Vivo and Retrospective Self-Reports of Alcohol Consumption. *PLOS ONE, 10*(5), e0126209. doi:10.1371/journal.pone.0126209

Moskowitz, D. S., & Young, S. N. (2006). Ecological momentary assessment: what it is and why it is a method of the future in clinical psychopharmacology. *Journal of Psychiatry and Neuroscience, 31*(1), 13-20.

Moss, A. C., Albery, I. P., Siddiqui, I., & Rycroft, N. (2013). Attentional Bias for Alcohol-Related Stimuli among Belief Based and Non-Belief Based Non-Drinkers. *19*, 299-302. doi:10.1159/000348426

NCCMH. (2011). *Alcohol use disorders: The NICE Guideline on the Diagnosis, Assessment and Management of Harmful Drinking and Alcohol Dependence*. Great Britain: The British Psychological Society and the Royal College of Psychiatrists

Noël, X., Colmant, M., Van Der Linden, M., Bechara, A., Bullens, Q., Hanak, C., & Verbanck, P. (2006). Time course of attention for alcohol cues in abstinent alcoholic patients: the role of initial orienting. *Alcoholism, Clinical And Experimental Research, 30*(11), 1871-1877. doi:10.1111/j.1530-0277.2006.00224.x

Nolen-Hoeksema, S. (2004). Gender differences in risk factors and consequences for alcohol use and problems. *Clinical psychology review, 24*(8), 981-1010. doi:10.1016/j.cpr.2004.08.003

Nutt, D. J., King, L. A., & Phillips, L. D. (2010). Drug harms in the UK: a multicriteria decision analysis. *The Lancet, 376*(9752), 1558-1565. doi:https://doi.org/10.1016/S0140-6736(10)61462-6

O'Brien, C. (2011). Addiction and dependence in DSM-V. *Addiction, 106*(5), 866-867. doi:10.1111/j.1360-0443.2010.03144.x

Oulasvirta, A., Rattenbury, T., Ma, L., & Raita, E. (2012). Habits make smartphone use more pervasive. *Personal and Ubiquitous Computing, 16*(1), 105-114. doi:10.1007/s00779-011-0412-2

Peeters, M., Wiers, R. W., Monshouwer, K., Schoot, R., Janssen, T., & Vollebergh, W. A. M. (2012). Automatic processes in at-risk adolescents: the role of alcohol-approach

tendencies and response inhibition in drinking behavior. *Addiction, 107*(11), 1939-1946. doi:10.1111/j.1360-0443.2012.03948.x
Perski, O., Blandford, A., Ubhi, H. K., West, R., & Michie, S. (2017). Smokers' and drinkers' choice of smartphone applications and expectations of engagement: a think aloud and interview study. *BMC Medical Informatics and Decision Making, 17*(25). doi:10.1186/s12911-017-0422-8
Peugh, J. L. (2010). A practical guide to multilevel modeling. *Journal of School Psychology, 48*(1), 85-112. doi:http://dx.doi.org/10.1016/j.jsp.2009.09.002
Phillips, K. T., Phillips, M. M., Lalonde, T. L., & Prince, M. A. (2018). Does social context matter? An ecological momentary assessment study of marijuana use among college students. *Addictive behaviors, 83*, 154-159. doi:https://doi.org/10.1016/j.addbeh.2018.01.004
Pieters, S., Burk, W. J., Vorst, H., Engels, R. C., & Wiers, R. W. (2014). Impulsive and reflective processes related to alcohol use in young adolescents. *Frontiers in psychiatry, 5*. doi:10.3389/fpsyt.2014.00056
Popova, S., Lange, S., Probst, C., Gmel, G., & Rehm, J. (2017). Estimation of national, regional, and global prevalence of alcohol use during pregnancy and fetal alcohol syndrome: a systematic review and meta-analysis. *The Lancet Global Health, 5*(3), e290-e299. doi:https://doi.org/10.1016/S2214-109X(17)30021-9
Quené, H., & van den Bergh, H. (2004). On multi-level modeling of data from repeated measures designs: a tutorial. *Speech Communication, 43*(1), 103-121. doi:https://doi.org/10.1016/j.specom.2004.02.004
Ramirez, J. J., Monti, P. M., & Colwill, R. M. (2015). Alcohol-cue exposure effects on craving and attentional bias in underage college-student drinkers. *Psychol Addict Behav, 29*(2), 317-322. doi:10.1037/adb0000028
Rasbash, J., Charlton, C., Browne, W. J., Healy, M., & Cameron, B. (2009). *MLwiN Version 2.1.* . UK: Centre for Multilevel Modelling, University of Bristol.
Rasbash, J., Steele, F., Browne, W. J., & Goldstein, H. (2012). *A user's guide to MLwiN: Version 2.26*. United Kingdom: Center for Multilevel Modelling, University of Bristol.
Rehm, J. (2011). The Risks Associated With Alcohol Use and Alcoholism. *Alcohol Research & Health, 34*(2), 135-143.
Rehm, J., & Shield, K. D. (2013). Global alcohol-attributable deaths from cancer, liver cirrhosis, and injury in 2010. *Alcohol Research: Current Reviews, 35*(2), 174-183.
Rehm, J., Zatonksi, W., Taylor, B., & Anderson, P. (2011). Epidemiology and alcohol policy in Europe. *Addiction, 106*, 11-19. doi:10.1111/j.1360-0443.2010.03326.x
Rensink, R. A., O'Regan, J. K., & Clark, J. (1997). To See or Not to See: The Need for Attention to Perceive Changes in Scenes. *Psychological Science, 8*(5), 368-373.
Robbins, S. J., & Ehrman, R. N. (2004). The role of attentional bias in substance abuse. *Behavioral and Cognitive Neuroscience Reviews, 3*(4), 243-260. doi:10.1177/1534582305275423
Robinson, C. D., Muench, C., Brede, E., Endrighi, R., Szeto, E. H., Sells, J. R., Lammers, J. P., Okuyemi, K. S., & Waters, A. J. (2017). Effect of attentional retraining on cognition, craving, and smoking in African American smokers. *Psychology of Addictive Behaviors, 31*(5), 636-646. doi:10.1037/adb0000286
Robinson, T. E., & Berridge, K. C. (1993). The neural basis of drug craving: an incentive-sensitization theory of addiction. *Brain research reviews, 18*(3), 247-291.
Robinson, T. E., & Berridge, K. C. (2001). Incentive-sensitization and addiction. *Addiction, 96*(1), 103-114. doi:10.1080/09652140020016996
Robinson, T. E., & Berridge, K. C. (2008). The incentive sensitization theory of addiction: some current issues. *Philosophical Transactions of the Royal Society B: Biological Sciences, 363*, 3137-3146. doi:10.1098/rstb.2008.0093

Rodebaugh, T. L., Scullin, R. B., Langer, J. K., Dixon, D. J., Huppert, J. D., Bernstein, A., Zvielli, A., & Lenze, E. J. (2016). Unreliability as a threat to understanding psychopathology: The cautionary tale of attentional bias. *Journal of Abnormal Psychology, 125*(6), 840-851. doi:http://dx.doi.org/10.1037/abn0000184

Roefs, A., Huijding, J., Smulders, F. T., MacLeod, C. M., de Jong, P. J., Wiers, R. W., & Jansen, A. (2011). Implicit measures of association in psychopathology research. *Psychological bulletin, 137*(1), 149-193. doi:10.1037/a0021729

Roerecke, M., & Rehm, J. (2012). Alcohol Intake Revisited: Risks and Benefits. *Current Atherosclerosis Reports, 14*(6), 556-562. doi:10.1007/s11883-012-0277-5

Rose, A. K., Brown, K., Field, M., & Hogarth, L. (2013). The contributions of value-based decision-making and attentional bias to alcohol-seeking following devaluation. *Addiction, 108*(7), 1241-1249.

Rose, A. K., Brown, K., MacKillop, J., Field, M., & Hogarth, L. (2018). Alcohol devaluation has dissociable effects on distinct components of alcohol behaviour. *Psychopharmacology, 235*(4), 1233-1244. doi:10.1007/s00213-018-4839-2

Roy-Charland, A., Plamondon, A., Homeniuk, A. S., Flesch, C. A., Klein, R. M., & Stewart, S. H. (2017). Attentional bias toward alcohol-related stimuli in heavy drinkers: evidence from dynamic eye movement recording. *The American journal of drug and alcohol abuse, 43*(3), 332-340. doi:https://doi.org/10.1080/00952990.2016.1209511

Ryan, F. (2002). Attentional bias and alcohol dependence: A controlled study using the modified Stroop paradigm. *Addictive behaviors, 27*(4), 471-482. doi:https://doi.org/10.1016/S0306-4603(01)00183-6

Saunders, J. B., Aasland, O. G., Babor, T. F., De la Fuente, J. R., & Grant, M. (1993). Development of the alcohol use disorders identification test (AUDIT). WHO collaborative project on early detection of persons with harmful alcohol consumption-II. *Addiction, 88*, 791-791.

Sayette, M. A., Shiffman, S., Tiffany, S. T., Niaura, R. S., Martin, C. S., & Shadel, W. G. (2000). The measurement of drug craving. *Addiction, 95*, S189-S210.

Scarborough, P., Bhatnagar, P., Wickramasinghe, K. K., Allender, S., Foster, C., & Rayner, M. (2011). The economic burden of ill health due to diet, physical inactivity, smoking, alcohol and obesity in the UK: an update to 2006–07 NHS costs. *Journal of Public Health, 33*(4), 527-535. doi:10.1093/pubmed/fdr033

Schmukle, S. C. (2005). Unreliability of the dot probe task. *European Journal of Personality, 19*(7), 595-605. doi:10.1002/per.554

Schoenmakers, T., & Wiers, R. (2010). Craving and attentional bias respond differently to alcohol priming: a field study in the pub. *European Addiction Research, 16*(1), 9-16. doi:10.1159/000253859

Schoenmakers, T., Wiers, R. W., & Field, M. (2008). Effects of a low dose of alcohol on cognitive biases and craving in heavy drinkers. *Psychopharmacology, 197*(1), 169-178. doi:10.1007/s00213-007-1023-5

Schoenmakers, T., Wiers, R. W., Jones, B. T., Bruce, G., & Jansen, A. T. M. (2007). Attentional re-training decreases attentional bias in heavy drinkers without generalization. *Addiction, 102*(3), 399-405. doi:10.1111/j.1360-0443.2006.01718.x

Schoenmakers, T. M., de Bruin, M., Lux, I. F. M., Goertz, A. G., Van Kerkhof, D. H. A. T., & Wiers, R. W. (2010). Clinical effectiveness of attentional bias modification training in abstinent alcoholic patients. *Drug and alcohol dependence, 109*(1–3), 30-36. doi:http://dx.doi.org/10.1016/j.drugalcdep.2009.11.022

Schuckit, M. A. (2009). Alcohol-use disorders. *The Lancet, 373*(9662), 492-501. doi:10.1016/S01406736(09)60009-X

Serre, F., Fatseas, M., Swendsen, J., & Auriacombe, M. (2015). Ecological momentary assessment in the investigation of craving and substance use in daily life: A systematic review. *Drug and alcohol dependence, 148*, 1-20. doi:http://dx.doi.org/10.1016/j.drugalcdep.2014.12.024

Sharma, D., Albery, I. P., & Cook, C. (2001). Selective attentional bias to alcohol related stimuli in problem drinkers and non-problem drinkers. *Addiction, 96*(2), 285-295. doi:10.1080/09652140020021026

Shen, J., & Gao, S. (2008). A Solution to Separation and Multicollinearity in Multiple Logistic Regression. *Journal of data science, 6*(4), 515-531.

Shiffman, S. (2009). Ecological momentary assessment (EMA) in studies of substance use. *Psychological Assessment, 21*(4), 486-497. doi:10.1037/a0017074

Shiffman, S., Stone, A. A., & Hufford, M. R. (2008). Ecological momentary assessment. *Annual review of clinical psychology, 4*, 1-32. doi:10.1146/annurev.clinpsy.3.022806.091415

Shin, E., Hopfinger, J. B., Lust, S. A., Henry, E. A., & Bartholow, B. D. (2010). Electrophysiological evidence of alcohol-related attentional bias in social drinkers low in alcohol sensitivity. *Psychology of Addictive Behaviors, 24*(3), 508-515. doi:10.1037/a0019663

Simons, D. J., & Rensink, R. A. (2005). Change blindness: Past, present, and future. *Trends in Cognitive Sciences, 9*(1), 16-20.

Simpson, T. L., Kivlahan, D. R., Bush, K. R., & McFall, M. E. (2005). Telephone self-monitoring among alcohol use disorder patients in early recovery: a randomized study of feasibility and measurement reactivity. *Drug and alcohol dependence, 79*(2), 241-250. doi:https://doi.org/10.1016/j.drugalcdep.2005.02.001

Snelleman, M., Schoenmakers, T. M., & Mheen, D. (2015). Attentional bias and approach/avoidance tendencies do not predict relapse or time to relapse in alcohol dependency. *Alcoholism: Clinical and Experimental Research, 39*(9), 1734-1739. doi:10.1111/acer.12817

Sobell, L. C., & Sobell, M. B. (1990). Self-report issues in alcohol abuse: State of the art and future directions. *Behavioral Assessment, 12*, 77-90.

Sobell, M. B., Sobell, L. C., Klajner, F., Pavan, D., & Basian, E. (1986). The reliability of a timeline method for assessing normal drinker college students' recent drinking history: utility for alcohol research. *Addictive behaviors, 11*(2), 149-161.

Spruyt, A., De Houwer, J., Tibboel, H., Verschuere, B., Crombez, G., Verbanck, P., Hanak, C., Brevers, D., & Noël, X. (2013). On the predictive validity of automatically activated approach/avoidance tendencies in abstaining alcohol-dependent patients. *Drug and alcohol dependence, 127*(1), 81-86. doi:https://doi.org/10.1016/j.drugalcdep.2012.06.019

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. doi:10.1146/annurev.clinpsy.121208.131444

Staugaard, S. R. (2009). Reliability of two versions of the dot-probe task using photographic faces. *Psychology Science Quarterly, 51*(3), 339-350.

Stetter, F., Ackermann, K., Bizer, A., Straube, E. R., & Mann, K. (1995). Effects of disease-related cues in alcoholic inpatients: Results of a controlled “alcohol Stroop” study. *Alcoholism: Clinical and Experimental Research, 19*(3), 593-599.

Stone, A. A., Shiffman, S., Schwartz, J. E., Broderick, J. E., & Hufford, M. R. (2003). Patient compliance with paper and electronic diaries. *Controlled Clinical Trials, 24*(2), 182-199. doi:https://doi.org/10.1016/S0197-2456(02)00320-3

Stone, A. L., Becker, L. G., Huber, A. M., & Catalano, R. F. (2012). Review of risk and protective factors of substance use and problem use in emerging adulthood.

Addictive behaviors, 37(7), 747-775. doi:https://doi.org/10.1016/j.addbeh.2012.02.014

Stormark, K. M., Field, N. P., Hugdahl, K., & Horowitz, M. (1997). Selective processing of visual alcohol cues in abstinent alcoholics: an approach-avoidance conflict? *Addictive behaviors, 22*(4), 509-519.

Stormark, K. M., Laberg, J. C., Nordby, H., & Hugdahl, K. (2000). Alcoholics' selective attention to alcohol stimuli: automated processing? *Journal of Studies on Alcohol, 61*(1), 18-23.

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of experimental psychology, 18*(6), 643.

Swendsen, J., Conway, K. P., Degenhardt, L., Glantz, M., Jin, R., Merikangas, K. R., Sampson, N., & Kessler, R. C. (2010). Mental disorders as risk factors for substance use, abuse and dependence: results from the 10-year follow-up of the National Comorbidity Survey. *Addiction, 105*(6), 1117-1128. doi:10.1111/j.1360-0443.2010.02902.x

Taylor, B., Irving, H. M., Kanteres, F., Room, R., Borges, G., Cherpitel, C., Greenfield, T., & Rehm, J. (2010). The more you drink, the harder you fall: A systematic review and meta-analysis of how acute alcohol consumption and injury or collision risk increase together. *Drug & Alcohol Dependence, 110*(1), 108-116. doi:10.1016/j.drugalcdep.2010.02.011

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

Tiffany, S., & Conklin, C. (2000). A cognitive processing model of alcohol craving and compulsive alcohol use. Addiction, 95 (Suppl2), S145-S153.

Townshend, J., & Duka, T. (2001). Attentional bias associated with alcohol cues: differences between heavy and occasional social drinkers. *Psychopharmacology, 157*(1), 67-74. doi:10.1007/s002130100764

Townshend, J. M., & Duka, T. (2007). Avoidance of Alcohol-Related Stimuli in Alcohol-Dependent Inpatients. *Alcoholism: Clinical and Experimental Research, 31*(8), 1349-1357. doi:10.1111/j.1530-0277.2007.00429.x

Vinci, C., Li, L., Wu, C., Lam, C. Y., Guo, L., Correa-Fernández, V., Spears, C. A., Hoover, D. S., Etcheverry, P. E., & Wetter, D. W. (2017). The association of positive emotion and first smoking lapse: An ecological momentary assessment study. *Health Psychology, 36*(11), 1038-1046. doi:10.1037/hea0000535

Vollstädt-Klein, S., Loeber, S., Richter, A., Kirsch, M., Bach, P., Goltz, C. v. d., Hermann, D., Mann, K., & Kiefer, F. (2012). Validating incentive salience with functional magnetic resonance imaging: association between mesolimbic cue reactivity and attentional bias in alcohol-dependent patients. *Addiction Biology, 17*(4), 807-816. doi:doi:10.1111/j.1369-1600.2011.00352.x

Waters, A. J., & Li, Y. (2008). Evaluating the utility of administering a reaction time task in an ecological momentary assessment study. *Psychopharmacology, 197*(1), 25-35. doi:10.1007/s00213-007-1006-6

Waters, A. J., Marhe, R., & Franken, I. H. (2012). Attentional bias to drug cues is elevated before and during temptations to use heroin and cocaine. *Psychopharmacology, 219*(3), 909-921. doi:10.1007/s00213-011-2424-z

Waters, A. J., Sayette, M. A., Franken, I. H. A., & Schwartz, J. E. (2005). Generalizability of carry-over effects in the emotional Stroop task. *Behaviour Research and Therapy, 43*(6), 715-732. doi:https://doi.org/10.1016/j.brat.2004.06.003

Waters, A. J., Szeto, E. H., Wetter, D. W., Cinciripini, P. M., Robinson, J. D., & Li, Y. (2014). Cognition and craving during smoking cessation: an ecological momentary assessment study. *nicotine & tobacco research, 16*(Suppl 2), S111-S118. doi:10.1093/ntr/ntt108

Watson, P., De Wit, S., Hommel, B., & Wiers, R. (2012). Motivational Mechanisms and Outcome Expectancies Underlying the Approach Bias toward Addictive Substances. *Frontiers in Psychology, 3*(440). doi:10.3389/fpsyg.2012.00440

Werthmann, J., Roefs, A., Nederkoorn, C., & Jansen, A. (2013). Desire lies in the eyes: attention bias for chocolate is related to craving and self-endorsed eating permission. *Appetite, 70*, 81-89. doi:http://dx.doi.org/10.1016/j.appet.2013.06.087

Wertz, J. M., & Sayette, M. A. (2001). A review of the effects of perceived drug use opportunity on self-reported urge. *Experimental and Clinical Psychopharmacology, 9*(1), 3-13. doi:10.1037/1064-1297.9.1.3

WHO. (2014). *Global status report on alcohol and health-2014*. Luxemburg: World Health Organization.

Wiers, C. E., Gladwin, T. E., Ludwig, V. U., Gröpper, S., Stuke, H., Gawron, C. K., Wiers, R. W., Walter, H., & Bermpohl, F. (2017). Comparing three cognitive biases for alcohol cues in alcohol dependence. *Alcohol and Alcoholism, 52*(2), 242-248. doi:10.1093/alcalc/agw063

Wiers, R. W., Houben, K., Fadardi, J. S., van Beek, P., Rhemtulla, M., & Cox, W. M. (2015). Alcohol Cognitive Bias Modification training for problem drinkers over the web. *Addictive behaviors, 40*, 21-26. doi:http://dx.doi.org/10.1016/j.addbeh.2014.08.010

Wiers, R. W., Rinck, M., Dictus, M., & Wildenberg, E. V. D. (2009). Relatively strong automatic appetitive action-tendencies in male carriers of the OPRM1 G-allele. *Genes, Brain and Behavior, 8*(1), 101-106. doi:10.1111/j.1601-183X.2008.00454.x

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

Zetteler, J. I., Stollery, B. T., Weinstein, A. M., & Lingford-Hughes, A. R. (2006). Attentional bias for alcohol-related information in adolescents with alcohol-dependent parents. *Alcohol and Alcoholism, 41*(4), 426-430. doi:10.1093/alcalc/agl026

Zhang, M., Ying, J., Song, G., Fung, D. S., & Smith, H. (2018). Attention and Cognitive Bias Modification Apps: Review of the Literature and of Commercially Available Apps. *JMIR mHealth and uHealth, 6*(5). doi:10.2196/10034

Zvielli, A., Bernstein, A., & Koster, E. H. (2015). Temporal dynamics of attentional bias. *Clinical Psychological Science, 3*(5), 772-788. doi:10.1177/2167702614551572

Appendix 1 (Word stimuli in the alcohol Stroop-task)

Table 1. List* of alcohol and control word stimuli

Alcohol stimuli	Matched control stimuli (environmental features)
PUB	BOG
LIQUER	RAVINE
WINE	VALLEY
COCKTAIL	BRIDGES
BREWERY	PEBBLE
BREW	COVE
CIDER	CRAGS
SPIRITS	LEAVES
STOUT	CANAL
BOOZE	INLET
DRUNK	HARBOUR

* List of words was based on Sharma et al. (2001)

Example of presentation in the alcohol Stroop-task

Appendix 2 (Alcoholic-drink preference screening)

What would be your most preferred alcoholic drinks? Please indicate what your first four preferences would be, starting from the most preferred (number 1).

1.

2.

3.

4.

Appendix 3 (TLFB – 2 weeks)

TLFB

Please read the instructions below and then fill in the following tables and answer the following questions.

To help me evaluate your drinking I need to get an idea of your alcohol consumption in the past fourteen days. Please fill out the table with the number of units of alcohol consumed on each day, being as accurate as possible. Please use the information given below to work out how many units you consumed on each day in the past week and fill in the number of units in the table. On days when you did not drink please write 0 (zero). I realise it isn't easy to recall things with 100% accuracy, but if you are not sure how many units you drank on a certain day please try to give it your best guess.

What is a unit of alcohol?

The list below shows the number of units of alcohol in common drinks:

Please now fill in the following table stating the total number of alcohol units you consumed for each day. Please start from whichever day it was yesterday and work backwards. For example if today is Monday start from Sunday and work backwards, with Monday being Monday a week ago. Please double check that you have filled in the number of units for all fourteen days.

Last week:

Monday	Tuesday	Wednesday	Thursday	Friday	Saturday	Sunday

Previous week:

Monday	Tuesday	Wednesday	Thursday	Friday	Saturday	Sunday

Appendix 4 (AUDIT)

AUDIT

Please read the instructions below and answer the following questions.

Because alcohol use can affect your health and can interfere with certain medications and treatments, it is important that we ask some questions about your use of alcohol. Your answers will remain confidential so please be honest. Circle the option that best describes your answer to each question.

1) **How often do you have a drink containing alcohol?**

Never   Less than monthly   2-4 times a month   2-3 times per week   4+per week

2) **How many drinks containing alcohol do you have on a typical day when you're drinking?**

1-2   3-4   5-6   7-9   10+

3) **How often do you have 6 or more drinks on one occasion?**

Never   Less than monthly   Monthly   Weekly   Daily or almost daily

4) **How often during the last year have you found that you were not able to stop drinking once you had started?**

Never   Less than monthly   Monthly   Weekly   Daily or almost daily

5) **How often during the last year have you failed to do what was normally expected from you because of drinking?**

Never   Less than monthly   Monthly   Weekly   Daily or almost daily

6) **How often during the last year have you needed a drink first thing in the morning to get yourself going after a heavy drinking session?**

Never Less than monthly Monthly Weekly Daily or almost daily

7) **How often during the last year have you had a feeling of guilt or remorse after drinking?**

Never Less than monthly Monthly Weekly Daily or almost daily

8) **How often during the last year have you been unable to remember what happened the night before because you had been drinking?**

Never Less than monthly Monthly Weekly Daily or almost daily

9) **Have you or someone else been injured because of your drinking?**

No Yes, but not in the last year Yes, during the last year

10) **Has a relative, friend, doctor or other health worker been concerned about your drinking or suggested you cut down?**

No Yes, but not in the last year Yes, during the last year

Appendix 5 (Intention to drink question)

Do you have any plans for today that involve consumption of alcohol?

YES

NO

Appendix 6 (Type of assessment question)

You are doing this AlSAp assessment now because

1. It is one of the fixed ones.

2. I experienced a craving to drink alcohol.

3. I am about to drink alcohol.

Appendix 7 (Craving Likert scale)

Please touch a number to indicate how strong your urge to drink alcohol is right now.

1	2	3	4	5	6	7	8	9	10	11
No urge		Slight		Moderate			Strong			Very strong

Appendix 8 (Focus group topic guide)

Exploration of app usability - Topic guide

Focus group objectives:

- To receive feedback for the experimental app AlSAp.
- To explore the features and characteristics the app should have if it was to be developed as a real professional app.
- To explore the app-user interaction experience.

Introduction of Study

- Name and position of Focus Group Lead.
- Name and position of supporting research staff
- Nature of study – reason for conducting study
- Aims of the study (see objectives)
- Check participants okay with focus group duration (ensure available for a one hour block).
- Inform participants about structure of focus group (Section A: Explaining the purpose and logic of the app, Section B: Guided use of the app, Section C: Semi-structured interview for feedback on the app, Section D: Opportunity for overall or additional comments.).
- Inform where toilets are.
- Explain fire evacuation procedure
- Provide voucher to reimburse for time ensuring participants sign to confirm receipt.

Ground Rules

- The focus group is a relaxed discussion exploring one another views on use of the app.
- This is not a debate on which choices are 'best' and there will be no judgement on one another's experiences.
- Respect each other's contributions and be polite – do not interrupt, undermine or pass inappropriate judgement on others' experiences.
- If the researcher feels the discussion is going off topic they will stop the discussion and bring it back to relevant discussion.

- Feel free to respond to other's comments should you have anything to add. The discussion is to be kept completely confidential by the researchers and participants involved and should not be discussed outside of the group.
- I will be audio recording the discussion for transcription afterwards (check all ok)
- Researchers may use direct quotes in the write up of the study but will ensure that the identification of the participant is protected by the use of pseudonyms.
- In the rare case that there are any disclosures about issues such as safeguarding, exploitation, harm or drug or alcohol abuse, researchers may have to break confidentiality and contact the participants Health Care Professional so that the appropriate course of action can be taken
- You are entitled to withdraw your information at any time until the point of anonymization (approximately 3 weeks after the focus groups has been completed) should you no longer wish to take part in the study.

Participant Introduction

- Name (introduction to other participants)
- A sentence about you (e.g. where are you from, what you do as a job)
- Hand out demographic questionnaires (allow 5 mins to complete)
- Hand out smartphones with the app installed.

START RECORDING

1. Section A: Introduction to the app

The app will be displayed on a big screen and it will be explained by the Focus group Lead.

First I would like to say a few thing about the app I will ask you to evaluate and explain its logic. What we know from previous research is that the extent to which your attention is automatically attracted by alcohol related cues in the environment might be related to your alcohol consumption. In general, people tend to spot in the environment things they like and their attention will be attracted by those things. For alcohol, this might give as an indication of whether that person will drink alcohol later.

For example, when you enter a room, there might be various alcohol related objects there, a wine glass, a bottle opener or anything of the like. To what extent your attention will automatically focus on those objects, might be able to tell us something about the chances *of you having a drink soon afterwards. So, what the app does is that it evaluates to what extent your attention is automatically attracted by words related to alcohol. It looks like a very simple game. It shows you on the screen various words, some of them have something to do with alcohol while the rest are not alcohol-related. The words will appear in a certain colour, blue, green or red, and there will also be some buttons on the screen with the names of these three colours. What you have to do is to click the correct button to say what the colour of the word is. For example, if the word is written in blue you need to click the button reading blue. If it is in red, you should click the button reading red and so on. But you need to be fast, you will only have a few seconds to respond before the next word appears. If you do a mistake a red X will show up. As long as you don't see the X, it means you are giving the correct answer. Before and after that colour game, the app will also ask you a few questions, like whether you have any plans to drink alcohol today or how strong your urge would be to drink alcohol right now.*

Would you have any questions before we proceed with trying the app?

2. Section B: Guided use of the app

The interviewer will guide participants step by step to launch the app on the phones that were given to them and go through some initial menus before the app starts. The interviewer will wait for all participants to finish the app before proceeding to the next section and will also assist any participant requiring for help.

3. Section C: User-app interaction

(Questions on this section are based on Crane (2017), with the addition of some extra questions)

- Was there anything you particularly liked or disliked? How do you find the length and difficulty level?
- If you were asked to use the app again, when on your own, without the researcher being present to provide help, was there anything you would find particularly difficult to do?
- If asked to use the app again, when on your own, without the researcher being present to provide help, was there anything you would find particularly easy to do?
- How many times per day would you be willing to use the app and for how long? Would you find repeated use of the app intrusive?

- How would you prefer to be reminded to use the app during the day? For example, would you prefer to rely on your memory, receive a text-message, the phone beeping, etc.
- If our aim is to examine how much your attention is automatically attracted by alcohol-relate cues, what type of cues would you like to see there? What would you think would be the best comparison cues (control cues)?
- How believable would it be to you that using this app could predict if/how much you will drink later?
- Do you have any suggestions for how the app could be improved?

Section D: **General comments**

- Are there any other comments you would like to make?

Appendix 9 (AlSAp participant's manual – three assessments per day on pre-defined time-periods)

AlSAp – Manual

This manual will help you to use the Alcohol Stroop App.

1. When to use

The app should be used THREE TIMES per day during the following times:

- Once between **10:30 – 12:30.**

- Once between **13:30 – 15:30.**

- Once between **16:30 – 18:30.**

ATTENTION: Use the application in a safe and secure location with no distractions and with no risk for accidents. DO NOT USE the application during driving or any other activity that requires high concentration and awareness of the environment.

IMPORTANT: You are kindly requested **NOT TO** consume any alcohol at any point before the last session (16:30 – 18:30). Consumption of alcohol is allowed after you have completed the last session. In any case, if consumption of alcohol does occur before the last session, please indicate that by circling the day-number in the table in the phone-case.

2. Subject number

Every time you use the app it will ask you to enter a subject number. This number is different for every session. The subject number consists of two numbers, the first one corresponding to the day of use and the second to the time of use.

Use the table below to identify the appropriate subject number for every day and time-slot.

Table of subject numbers

Day	Time	Subject Number
1	10:30 – 12:30	**11**
	13:30 – 15:30	**12**
	16:30 – 18:30	**13**
2	10:30 – 12:30	**21**
	13:30 – 15:30	**22**
	16:30 – 18:30	**23**
3	10:30 – 12:30	**31**
	13:30 – 15:30	**32**
	16:30 – 18:30	**33**
4	10:30 – 12:30	**41**
	13:30 – 15:30	**42**
	16:30 – 18:30	**43**
5	10:30 – 12:30	**51**
	13:30 – 15:30	**52**
	16:30 – 18:30	**53**
6	10:30 – 12:30	**61**
	13:30 – 15:30	**62**
	16:30 – 18:30	**63**
7	10:30 – 12:30	**71**
	13:30 – 15:30	**72**
	16:30 – 18:30	**73**

NOTE:

- In case you enter any other subject number that is not included in the table (e.g. 15), AlSAp will close automatically. Launch the app again and give the right number.

- In case you accidentally enter the number from another time-slot, please launch the app again and use the correct number. Then move on with the rest of the table as you would normally do.

-In case you skip a time-slot, skip also the corresponding subject number and move on with the rest of the table as you would normally do.

- In case the application crashes for any reason during a session, just exit the application (or even restart your phone if necessary), launch the application again and start again the interrupted session (use the same participant number as you entered in the session that crashed).

3. Step by step guide

Before you launch the application, make sure that:

1. The sound is activated in the mobile phone.
2. You have identified the appropriate subject_number for this session, as indicated in paragraph 2.

To start the app, follow these steps:

- Touch the "menu" icon  on the bottom of the main screen.
- Scroll the screen until you find the OpenSesame icon . Touch the icon to launch the application.
- A black screen with the path */sdcard_* will appear on the screen and a touch-keyboard will appear. DO NOT write anything in the path. In case that anything other than */sdcard* is written on the screen, delete it and type */sdcard*. Press ***ENTER.***
- From the list of available experiments select "*Alc stroop 0*".
- A black screen will appear asking for a subject number. Enter the appropriate subject number as indicated in paragraph 2. Press ***ENTER.***
- A black screen will appear indicating the path */sdcard/subject[subject number].csv_* **DO NOT** write anything in the path. Press ***ENTER.***

4. Problems or Questions

If you experience any problems using the app or if you have any questions, contact Mr. Panagiotis Spanakis at spanak87@liv.ac.uk.

Appendix 10 (AlSAp participant's Manual – up to five assessments per day, depending on the motivational situation)

AlSAp – Manual

1. When to use

Standard times: 1. One time between 10:30 – 12:30.

2. One time between 13:30 – 15:30.

Optional times: 1. When you experience craving (max two times).

2. One time, when you are about to have the first drink of the day.

[After you have the first drink of the day, you should not use the app again on that day.]

2. Subject number

Every time you log on the app, it will ask you to enter a subject number (sn). This number should be different every time. The sn should consist of two numbers; the first one corresponds to the day and the second to the time of use.

Example: On day 1, the 1st time you will use the app the sn would be 11. On the same day, the second time, the sn would be 12. The next day (day 2), the first time, the sn will be 21.

[In case you accidentally enter the wrong sn, exit the app, log-on again and use the correct number]

3. How to log on

- On the main screen tap the OpenSesame icon. .
- A black screen with the path */sdcard_* will appear on the screen and a touch-keyboard will pop up. DO NOT write anything in the path. In case that anything different than */sdcard* shows on the screen, delete it and type */sdcard*.

- Press ***ENTER***.
- From the list of available experiments select "*Alsap 1*".
- A black screen will appear asking for a subject number. Enter the appropriate subject number as indicated in paragraph 2.
- Press ***ENTER***.
- A black screen will appear indicating the path */sdcard/subject[subject number].csv_* **DO NOT** write anything in the path.
- Press ***ENTER***.

4. How to register your daily alcohol consumption.

- On the main screen tap the "*Notes*" icon.

- Tap anywhere in the pop-up window to maximize.

- Tap the pencil icon on the yellow menu bar on the top.

- Tap anywhere next to the day you want to register you consumption for.

5. Problems or Questions

If you experience any problems using the app or if you have any questions, contact Mr. Panagiotis Spanakis at spanak87@liv.ac.uk.