Serious gamification

Motivating adolescents to do cognitive training

Boendermaker, W.J.

SERIOUS GAMIFICATION

Motivating Adolescents to do Cognitive Training

Wouter J. Boendermaker

Serious Gamification

Motivating Adolescents to do Cognitive Training

Wouter J. Boendermaker

Cover design and layout by Wouter J. Boendermaker
Printing: Ridderprint BV, the Netherlands
ISBN: 978-94-6299-532-1

The research presented in this work was supported through funding received from the Netherlands Organization for Scientific Research (NWO; #056-34-015).

Serious Gamification
Motivating Adolescents to do Cognitive Training

Serious Gamification
Adolescenten Motiveren voor Cognitieve Training

ACADEMISCH PROEFSCHRIFT
ter verkrijging van de graad van doctor
aan de Universiteit van Amsterdam
op gezag van de Rector Magnificus
prof. dr. ir. K.I.J. Maex
ten overstaan van een door het college voor promoties ingestelde
commissie, in het openbaar te verdedigen in de Agnietenkapel
op donderdag 9 maart 2017, te 10.00 uur

door Wouter Jan Boendermaker
geboren te Smallingerland

Promotores:	Prof. dr. R. W. H. J. Wiers	Universiteit van Amsterdam
	Prof. dr. P. J. M. Prins	Universiteit van Amsterdam
Overige leden:		
	Prof. dr. H. M. Geurts	Universiteit van Amsterdam
	Prof. dr. I. Granic	Radboud Universiteit Nijmegen
	Prof. dr. L. Krabbendam	Vrije Universiteit Amsterdam
	Prof. dr. J. M. J. Murre	Universiteit van Amsterdam
	Dr. B. D. Ostafin	Rijksuniversiteit Groningen
	Prof. dr. G. Overbeek	Universiteit van Amsterdam
Faculteit:	Faculteit der Maatschappij- en Gedragswetenschappen	

Contents

Appendices

Chapter 1
General Introduction

Stay a while, and listen.

–Deckard Cain, Diablo II
(Blizzard Entertainment)

Parts of this chapter are based on:
Boendermaker, W. J., Peeters, M., Prins, P. J. M., & Wiers, R. W. (2017). Using serious games to (re)train cognition in adolescents. In M. Ma, & A. Oikonomou (Eds.), *Serious Games and Edutainment Applications, Volume II.* UK: Springer-Verlag. ISBN: 978-3-319-51643-1

Theoretical background

Although any game that is used for a more serious purpose than entertainment can be viewed as a serious game, many are developed with the specific purpose of improving motivation for a serious task by making it more enjoyable. For example, there has been a recent surge in the development of video games aimed at improving (mental) health (Kato, 2010). Many of these games for health, as they are often called, are aimed at adolescents and young adults, as gaming is very popular among those age groups (Entertainment Software Association, 2014). But adolescence is also a sensitive period when it comes to mental health, potentially making serious games especially effective at this age. The question central to this chapter is, how can the scientific evidence for cognitive training principles be used as a basis to effectively improve mental health in adolescents, using serious gaming techniques? In the following section, an overview will be presented detailing what is currently known about cognitive training, underlining the importance of using optimal principles or paradigms. In the third section, we will take a closer look at the role of motivation in adolescents. The fourth section will describe the current state of affairs of cognitive training games aimed at adolescents, and the final section will conclude with some recommendations regarding the development and study of these games.

Training Cognitive Processes

Adolescence is a developmental period characterized by a considerable increase in the prevalence of internalizing problems, such as anxiety and depression (Paus, Keshavan, & Giedd, 2008), as well as externalizing behavior, such as experimenting with risky behavior (Steinberg, 2007). The European School Survey Project on Alcohol and other Drugs (ESPAD; Hibell et al., 2012) showed that almost 60 percent of the 100.000 students in the survey reported to have consumed at least one glass of alcohol at the age of 13 or younger, 2 percent had already been drunk at that age, and 18 percent had tried illicit drugs at least once during their lifetime. Although this behavior does not necessarily lead to mental problems, excessive contact with psychoactive substances at this age can lead to misuse, school dropout (Singleton, 2007) and ultimately to addiction problems later in life. As such, early intervention is important to prevent escalation. While there are numerous prevention and treatment programs available, which tend to focus on explicit education, their efficacy is sometimes limited (Werch & Owen, 2002), which may in part be due to the fact that adolescents have a hard time reflecting on and making changes to their behavior in general (see Kopetz & Orehek, 2015).

Interestingly, recent evidence has emerged that many of these adolescent mental problems are associated with specific cognitive deficiencies, such as relatively weak Working Memory Capacity (WMC; Martinussen, Hayden, Hogg-Johnson, & Tannock, 2005), or a tendency to selectively attend to information or approach stimuli that strengthen the problematic behavior (Bar-Haim, Lamy, Pergamin, Bakermans-Kranenburg, & van IJzendoorn, 2007; Peckham, McHugh, & Otto, 2010; Wiers, Gladwin, Hofmann, Salemink, & Ridderinkhof, 2013). For example, many anxious people have a tendency to attend to negative or threatening information (Bar-Haim et al., 2007), whereas heavy alcohol users tend to selectively attend to alcohol-related information (Field & Cox, 2008). Similarly, an approach bias has been found for alcohol-related cues in heavy drinkers (Field, Kiernan, Eastwood, & Child, 2008; Wiers, Rinck, Dictus, & van den Wildenberg, 2009), and for cannabis cues in cannabis users (Cousijn, Goudriaan, & Wiers, 2011; Field, Eastwood, Bradley, & Mogg, 2006).

Training, or re-training, these cognitive processes can be effective in decreasing symptoms. We can distinguish between two closely related types of cognitive training (Wiers et al., 2013). First, there are training paradigms aimed at modifying the maladaptive cognitive biases (cognitive bias modification; CBM). These training procedures are usually disorder-specific, as they often involve stimuli (i.e., pictures, words) related to the disorder. For example, attentional bias modification in long time heavy substance users has been associated with a reduction in drinking (Fadardi & Cox, 2009) and a significantly longer time to relapse (Schoenmakers et al., 2010). Similarly, Amir, Beard, Burns, and Bomyea (2009) and Schmidt, Richey, Buckner, and Timpano (2009) showed a decrease in anxiety symptoms in clinically anxious patient groups (for a recent review, see: Kuckertz & Amir, 2015). Another type of CBM, aimed at modifying automatically activated action tendencies to approach or avoid disorder-related stimuli, was effective in reducing alcohol intake directly after training (Wiers, Rinck, Kordts, Houben, & Strack, 2010). When this type of re-training was added to regular therapy for alcoholism, relapse one year after treatment discharge was reduced with 13 percent in patients who received this type of training, compared with those who received sham-training or no training in addition to regular treatment (Wiers, Eberl, Rinck, Becker, & Lindenmeyer, 2011). This finding was recently replicated in a large study (N>500), with evidence for statistical mediation (the clinical change was mediated by the change in automatic action tendencies) and moderation (patients with a strong approach bias profited more from this training than those without a strong approach bias, Eberl et al., 2013). In addition to training attentional bias and approach bias, positive memory biases have also been targeted with different methods, such as evaluative conditioning (pairing the focal category alcohol with negative stimuli, Houben, Schoenmakers, & Wiers, 2010) and selective inhibition (pairing the focal category alcohol with an inhibition signal (Houben, Havermans, Nederkoorn, & Jansen, 2012; Houben, Nederkoorn, Wiers, & Jansen, 2011), with initial promising results in heavy drinkers.

Second, there are more domain-general cognitive control functions, such as WMC and impulse inhibition, collectively called Executive Functions (EF). Deficits in EF are implicated in many psychological disorders such as addiction, Attention Deficit Hyperactivity Disorder (ADHD) and autism spectrum disorders (ASD), but also in internalizing disorders such as anxiety and depression. These EF deficits are related to functional impairments and specific problem behaviors (Hosenbocus & Chahal, 2012). The three components of the executive system that have received the most research attention are working memory, inhibition and cognitive flexibility (Diamond, 2013). Meta-analyses have shown that working memory and inhibition are often impaired in school-age children and adolescents with ADHD (Martinussen et al., 2005; Wilcutt, Doyle, Nigg, Faraone, & Pennington, 2005). Executive impairments have also been reported in school-age children with ASD on measures tapping planning, inhibition of prepotent responses, and self-monitoring (Robinson, Goddard, Dritschel, Wisley, & Howlin, 2009). Although these EF deficits have been consistently found in ADHD and ASD, and are related to functional impairments and specific problem behaviors, they are not specific enough for use as clinical markers of these disorders (Geurts, Verte, Oosterlaan, Roeyers, & Sergeant, 2004; Hosenbocus & Chahal, 2012). In addiction these cognitive control functions are needed to regulate the impulsive reactions involved in the substance-related biases. There is some evidence that when inhibition (Houben & Wiers, 2009; Peeters et al., 2012) and WMC (Grenard et al., 2008; Thush et al., 2008) are weak, they can fail in their regulatory function, leading to an imbalanced cognitive system (Wiers et al., 2007). However, support for a causal relationship, where impaired control functions are a result of adolescent substance use, is weak (Wiers, Boelema, Nikolaou, & Gladwin, 2015).

Impairments in EF have also been linked to problems with self-regulation and the development of disruptive behavior problems such as found in Oppositional Defiant Disorder (ODD) and Conduct Disorder (CD; Schoemaker, Mulder, Dekovic, & Matthys, 2013). Results regarding working memory performance in children and adolescents with ODD and CD, however, are less consistent than found in youth with ADHD. Although cognitive deficits are widely recognized to be an important component of anxiety (Moran, 2016), fewer studies have assessed the relation between EF deficits and internalizing problems such as anxiety and depression. Anxiety, both self-reported as well as experimentally induced, is thought to restrict WMC by competing with task-relevant processes. According to cognitive theories, anxiety problems are related to impairments in attention processes.

Just as cognitive biases can be influenced through targeted training, cognitive control functions can also be strengthened through training, with the best results in children with relatively weak WMC (Holmes, Gathercole, & Dunning, 2009), such as children with ADHD (Klingberg, 2010). While cognitive training effects do not always generalize to other cognitive abilities (Shipstead, Redick, & Engle, 2012), increasing WMC can also lead to reduced drinking in problem drinkers with strong automatic positive associations with alcohol (Houben, Wiers, & Jansen, 2011). Training working memory in adolescents in order to improve their executive attentional control also resulted in positive changes in symptoms of trait and test-anxiety, increased inhibitory control, and reduced attention to threat (Hadwin & Richards, 2016).

Interestingly, these two types of cognitive processes are intimately intertwined in several disorders. For example, when cognitive control is low, the impulsive processes tend to better predict the maladaptive behavior (Wiers et al., 2013). Although there is ample evidence in favor of cognitive training (for more elaborate reviews, see Klingberg, 2010; Wiers et al., 2013), it is not without controversy. Some authors (e.g., Cristea, Mogoașe, David, & Cuijpers, 2015; Emmelkamp, 2012) note that the quality of the evidence in support of CBM is limited, with reported changes pertaining mainly to the targeted biases, but limited or no effects on mental health outcomes. Others (e.g., MacLeod & Clarke, 2015) have pointed out that when CBM succeeds in changing the targeted bias, a change in emotional outcomes is almost always found, while studies that did not succeed in changing the bias also hardly ever found effects on emotional outcomes, in line with the hypothesized mechanism of CBM (emotional change through a change in the targeted bias). Similarly, the notion of strengthening cognitive functions through cognitive training has also been subject of recent debate (e.g., Shipstead et al., 2012). While this chapter is not a place to repeat this debate on the efficacy of cognitive training, it may help to distinguish between the nature of the training studies that show or refute cognitive training effects with regards to the experimental settings. Importantly, while effects have generally been limited in single session studies with unmotivated participants (usually heavy drinking students who do not wish to reduced their drinking, e.g., Schoenmakers, Wiers, Jones, Bruce, & Jansen, 2007) or community smokers who do not wish to quit (Kerst & Waters, 2014), studies in which cognitive training has been delivered to clinical samples as add-on to regular treatment (Eberl et al., 2013; Elfeddali, de Vries, Bolman, Pronk, & Wiers, 2016; Schoenmakers et al., 2010; Wiers et al., 2011) have yielded significant improvements in clinical outcomes.

To conclude, although cognitive training can be a promising basis for the development of serious games, the specifics of the disorder, the target population, and the training paradigm are very important for the efficacy of the training. Another important aspect that can influence the efficacy of cognitive training is the participant's motivation, which is where serious game techniques may play an important role (Gladwin, Figner, Crone, & Wiers, 2011).

Motivation(s)

Similar to physical training exercises, most cognitive training paradigms rely on a substantial number of repeated actions over multiple sessions to reach a training effect. To sustain performance during these training sessions, it is necessary to reach and maintain a state of motivation high enough to continue training. But prior to the actual training, adolescents first need to be motivated to consider participating in a training. That is, they need to have a basic motivation to change their problematic behavior (Boffo, Pronk, Wiers, & Mannarini, 2015). Although the negative effects of adolescents' problematic behavior is often obvious to those around them, adolescents themselves sometimes lack the realization that they even have a problem. For example, many adolescents don’t think of their heavy alcohol use as problematic or harmful (Johnston, O'malley, Bachman, & Schulenberg, 2012). This is not strange when we look at the developmental function that many risk behaviors, such as alcohol use, can fulfill. Adolescents often engage in risky behaviors to attain or increase peer status or to receive positive evaluation of peers (Crone & Dahl, 2012; Sommerville, 2013; Steinberg, 2007). Peers and their perspectives become increasingly important in adolescence and adolescents’ behavioral decisions are often taken while confirming to peer norms and peer cultures (Baumeister, 1991; Forbes & Dahl, 2010), making it harder to provoke behavioral change. For that reasons it is important that intervention strategies are developed in line with the perceptions of adolescents and they can only be effective when aspects such as motivation and attention are taken into account. In contrast, most CBM studies feature adult patient samples, where most participants have a long history of substance use problems, and are thus more motivated to change their behavior. As such, there are important motivational differences between adolescents and adults that have to be taken into account when designing a motivating intervention, such as a serious game training. When this motivation to change is low, it may be best to combine cognitive training with other types of intervention (Wiers et al., 2013), such as motivational interviewing (Miller & Rollnick, 2002) or (cognitive) behavioral therapy. But even when participants are somewhat motivated to make a behavioral change, they often find cognitive training paradigms to be long and boring and have a hard time believing that a simple computer task can help them (Beard, Weisberg, & Primack, 2012). As such, they may still need to be motivated to complete the full training. Applying serious gaming techniques to evidence based training paradigms may help adolescents to increase their motivation to train. Specifically, serious games as an intervention strategy for adolescents have the ability to anticipate on two important cognitive developments that characterize adolescence; (1) development of behavioral control and (2) increased sensitivity for reward. Adolescents are (hyper) sensitive for reward, but at the same time have difficulties in controlling their behavior as behavioral control continues to develop into late adolescence (18-20 years; Blakemore & Choudhurry, 2006; Luna, Garver, Urban, Lazar, & Sweeney, 2004). Not yet fully developed behavioral control skills increase the chance that insufficient attention is giving to the task, specifically when the task is long and boring, subsequently resulting in incomplete or inefficient training. Adolescents with externalizing behavioral problems such as ADHD and CD are particularly at risk because of their deficits with behavioral and motivational control functions (Dovis, van der Oord, Wiers, & Prins, 2013; Krueger, Caspi, Moffitt, White, & Stouthamer-Loeber, 1996). Exactly this group of youngsters are at the greatest risk for developing addictive behaviors (Peeters et al., 2015) and therefore could benefit most from an efficient intervention strategy. As such, serious games can significantly advance the field of interventions for adolescents in general, and particularly for adolescents with externalizing behavioral problems. Compared to traditional intervention approaches, serious games can be better equipped to grasp attention and increase adolescents' motivation to complete the training (Dovis et al., 2013). Moreover, the competitive and arousing character of games can better connect to the perceptions of adolescents (Granic, Lobel, & Engels, 2014); status increase and competition are two often observed reasons of why adolescents engage in risk taking behavior (de Boer, Peeters, & Koning, 2016; Sommerville, 2013) and these factors could act as important reinforcers when efficiently processed in a serious game. The gameplay and the competitive character of serious games may increase the rewarding and motivating capacity of the intervention, however, notion should be taken in how rewarding elements are incorporated (Dovis et al., 2013; Boendermaker, Prins, & Wiers, 2015; **Chapter 2** in this thesis).

Project overview

The project that was the basis for this thesis was financed as part of the National Initiative Brain and Cognition (NIHC) by the Dutch National Science Foundation (NWO) and consisted of three main aims: (1) to produce game-versions of previously successful (re-)training programs that appeal to youth and have the same beneficial effects on the targeted cognitive process, motivation to control alcohol/drug use ultimately alcohol and drug use and academic performance; (2) study the effect on brain activity caused by a gaming context by comparing changes in brain functioning to training in a non-game context; and (3) implement the games into school-based prevention programs through collaboration with the Trimbos Institute and Arkin-Jellinek-

Prevention. The project was to include three stages: (I) Development of one shell-type game aimed at alcohol and one similar shell-type game aimed at cannabis, as well as several pilot studies; (II) Two large studies on effects of the fully developed training-program with a total of 14 training conditions (7 per substance), including a placebo control condition, five specific active training conditions (aimed at working memory, inhibition, task switching, attention bias, and approach bias), with one active module and the others in placebo mode, as well as one condition with all five modules in the active mode. Finally, in stage (III) the neurocognitive effects of the game training would be investigated. As the project was specifically aimed at the learning adolescent, all samples were to be school-based (i.e., non-clinical). Despite several attempted collaborations with professional game developers, the games were eventually designed and developed by the author of this thesis. The main reason for this decision was the fact that it turned out that none of the companies could afford developing the proposed games for the limited budget available, were not suited for development on this scale, or eventually dropped out because of other commitments.

Outline and aims of this thesis

This thesis describes a number of studies on the matter of using serious game techniques to motivate adolescents and young adults to do cognitive training aimed at restoring balance to the cognitive systems and in effect decrease their problematic alcohol use. As this area of research was and still is a fairly new area of research, the thesis starts with laying out a framework for the development and evaluation of serious games based on cognitive training paradigms. The following chapters will elaborate on the model by providing data collected in several pilot studies among school-based adolescent samples. The data in this thesis should therefore not be interpreted primarily as a validation of cognitive training as a form of treatment in general (especially in the context of prevention). Rather these are first explorations on how to use game elements to increase motivations to train among adolescents that can be used as proofs of concept for further development and study of serious games and cognitive training.

The remaining chapters in this thesis are ordered as follows. **Chapter 2** starts with providing a framework for the development and evaluation of serious games based on cognitive training paradigms. As at the point of writing almost no research was available on gamification for CBM, specifically, examples from the general cognitive training literature are used to build a model that distinguishes between seven steps for going from an original, evidence based CBM paradigm towards a commercial game. In the remainder of the thesis, closer looks are taken at these steps, where **Chapter 3** describes a first study using a shell-type game, called CityBuilder. This game was developed especially with adolescents in mind and incorporates a multitude of cognitive training paradigms. In this specific study working memory is trained among high school adolescents, and its effect on motivation is evaluated. This study shows that serious game elements can help to increase motivation to train working memory in adolescents. **Chapter 4** concerns a second study using the CityBuilder Game where the specific role of rewards was studied. The study presents a comparison between a regular inhibition training, one where correct responses are rewarded by (arbitrary) point rewards only, and one where the points are being given meaning as they can be used within the CityBuilder Game. **Chapter 5** explores other elements of fun by use of integrated gamification methods and swiping on mobile devices. The first study in this chapter describes the Cheese Ninja Game, an innovative game embedded inside a social media environment (www.facebook.com), based on a response matching training paradigm. The second study presents a mobile application based on an alcohol avoidance training paradigm, where participants were allowed to train as often and as long as they wanted during a period of two weeks. **Chapter 6** describes a critical evaluation of the Shots Game, an attentional bias retraining that mainly contains an elaborate point-rewards system and fancy graphics, but has no additional gameplay elements included. Finally, **Chapter 7** presents a summary of the results presented in this thesis, as well as a critical reflection on several practical aspects of the project that led to this work. In the general discussion, implications for future research are discussed.

Chapter 2
Cognitive Bias Modification for Adolescents with Substance Use Problems: Can Serious Games Help?

This chapter is based on:
Boendermaker, W. J., Prins, P. J. M., & Wiers, R. W. (2015). Cognitive bias modification for adolescents with substance use problems – Can serious games help? *Journal of Behavior Therapy and Experimental Psychiatry, 49*, 13-20. DOI: 10.1016/j.jbtep.2015.03.008.

Abstract

Background and Objectives: Excessive use of psychoactive substances and resulting disorders are a major societal problem, and the most prevalent mental disorder in young men. Recent reviews have concluded that Cognitive Bias Modification (CBM) shows promise as an intervention method in this field. As adolescence is a critical formative period, successful early intervention may be key in preventing later substance use disorders that are difficult to treat. One issue with adolescents, however, is that they often lack the motivation to change their behavior, and to engage in multisession cognitive training programs. The upcoming use of *serious games for health* may provide a solution to this motivational challenge. **Methods:** As the use of game-elements in CBM is fairly new, there are very few published studies in this field. This review therefore focuses on currently available evidence from similar fields, such as cognitive training, as well as several ongoing CBM gamification projects, to illustrate the general principles. **Results:** A number of steps in the gamification process are identified, starting with the original, evidence-based CBM task, towards full integration in a game. While more data is needed, some steps seem better suited for CBM gamification than others. Based on the current evidence, several recommendations are made. **Limitations:** As the field is still in its infancy, further research is needed before firm conclusions can be drawn. **Conclusions:** Gamified CBM may be a promising way to reach at risk youth, but the term “game” should be used with caution. Suggestions are made for future research.

Introduction

Alcohol- and drug misuse among adolescents pose a major societal problem. They predict school dropout and academic underperformance (Singleton, 2007; Wood, Sher, & McGowan, 2000) and, as they are relatively likely to escalate into more problematic use (Thatcher & Clark, 2008), may ultimately lead to later addiction problems. There are indications that young binge drinkers (i.e., adolescents who consume large amounts of alcohol, e.g., more than five drinks, within a short time period) are likely to develop atypical reactions to alcohol, which is reflected in their enhanced cue-reactivity (e.g., Tapert et al., 2003) and reduced ability to perform in executive cognitive tasks (Duka et al., 2004; Maurage et al., 2012). Similar neuroadaptations have been found for other popular substances in youth such as cannabis (e.g., Cousijn et al., 2012; 2013). As such, successful intervention during adolescence may help to prevent cognitive decline and substance use disorders later on. Several interventions exist that aim to persuade adolescents to abstain from or regulate their substance use. We can distinguish between explicit and implicit interventions. For example, explicit warning messages about the dangers of substance use are applied frequently (e.g., Drug Abuse Resistance Education, DARE), although their efficacy has been refuted on multiple occasions (e.g., Werch & Owen, 2002). Motivational interviewing (Miller & Rollnick, 2002) is another explicit, but more personalized technique, which has shown support in young adults (for review, see Larimer & Cronce, 2007), although its efficacy in adolescents has been questioned (e.g., Thush et al., 2009), showing mixed results (for review, see Barnett, Sussman, Smith, Rohrbach, & Spruijt-Metz, 2012). An alternative, more implicit intervention is cognitive training (introduced below).

Many adolescents do not consider their alcohol use as problematic or harmful (Johnston, O'Malley, Bachman, & Schulenberg, 2010). For example, Wiers, van de Luitgaarden, van den Wildenberg, and Smulders (2005) found that while 74 percent of their pre-screened sample of 96 late adolescents met diagnostic criteria for likely alcohol problems, only one of them actually self-indicated to have an alcohol problem. This lack of awareness may exist because adolescents tend to perceive more positive than negative effects of their alcohol use (National Institute on Alcohol Abuse and Alcoholism, 2005). As such, adolescents' motivation to change is often low and explicitly confronting them with their substance use may not be the most efficient way to prevent serious problems. Inspired by dual process models of addiction (e.g., Deutsch & Strack, 2006; Wiers et al., 2007), several varieties of cognitive training have been developed. These models posit that prolonged use of addictive substances leads to two important sets of cognitive changes. First there are several distinct impulsive or motivational reactions (biases) towards substances, such as attentional bias (e.g., Field et al., 2007; Schoenmakers et al., 2010), automatic memory associations (e.g., Stacy, 1997; Houben, Havermans, Nederkoorn, & Jansen, 2012) and approach bias (Wiers, Eberl, Rinck, Becker, & Lindenmeyer, 2011; Wiers, Rinck, Dictus, & van den Wildenberg, 2009). Second, it was posited that cognitive control processes that regulate these impulsive reactions, such as response inhibition (Houben & Wiers, 2009; Peeters et al., 2012) and working memory capacity (Grenard et al., 2008; Thush et al., 2008), may become weakened through prolonged use and eventually fail to fulfill their regulatory function. However, a recent review has shown that there is stronger support for enhanced motivational reactions to stimulus cues than for impaired control functions as a result of adolescent substance use (Wiers, Boelema, Nikolaou, & Gladwin, 2015). Meanwhile, there is evidence that premorbid weak control functions are predictive of later substance use escalation (de Wit, 2009; Verdejo-Garcia, Lawrence, & Clark), and the underlying mechanism may be that these individuals have more trouble in controlling their enhanced implicit motivational processes (Peeters et al., 2013; Wiers, Gladwin, Hofmann, Salemink, & Ridderinkhof, 2013; Wiers et al., 2015). The resulting imbalance between these stronger impulsive and relatively weak control processes can then lead to the development of addictive behaviors. Restoring balance may slow down this development, and eventually lead to a decline of substance use. In order to do so, Cognitive Bias Modification (CBM) techniques can be used to change these biased automatic, impulsive reactions by providing more time to make decisions regarding the use of a substance (Wiers et al., 2013). Additionally, cognitive control over the impulses may be strengthened through executive function training (for review, see Klingberg, 2010), and has shown promise in addiction (Houben, Nederkoorn, Wiers, & Jansen, 2011).

CBM is a collection of different training techniques aimed at changing relatively fast or impulsive reactions to disorder-relevant stimuli (Koster, Fox, & MacLeod, 2009). For example, heavy alcohol users often show selective attention (Field et al., 2007) or approach tendencies (Wiers et al., 2009) towards alcohol-related cues, resulting in cognitive biases. CBM is often applied through computer based reaction time tasks that aim to modify the bias through extensive practice, rather than explicit instruction (Koster et al., 2009). The efficacy of CBM remains subject of debate (Emmelkamp, 2012), but there certainly are indications that these processes can successfully be retrained, with positive clinical effects in addiction and related disorders (for review, see Wiers et al., 2013). In anxiety, Clarke, Notebaert, and MacLeod (2014) noted that out of 29 reviewed studies on CBM-Attention (CBM-A), 26 showed a clear link between achieved bias modification and observed change in emotional vulnerability: either both were observed (n = 16), or both were absent (n = 10). Hence, effects on behavior can only be expected when a change of bias has occurred.

Although CBM seems to be a promising new technique, the repetitive nature of the training tasks often makes them inherently boring (Beard, Weisberg, & Primack, 2012). Moreover, subjects often have a hard time believing that a simple computer task such as CBM training can really help them control their substance use (Beard et al., 2012). Therefore, an intrinsic motivation to change may be necessary for participants to follow through with the full training program. Most CBM studies trained adult patients, who often have a long history of substance use problems, and tend to be motivated to change their habits. Adolescents rarely have this insight, nor do they have a strong motivation to change their behaviors. And even when they do recognize that they have a problem, they may still need to be motivated to do the full training. Gladwin, Figner, Crone, and Wiers (2011) identified several ways to tackle this problem, one of which is to introduce game-elements. The products of such combinations are sometimes called *serious games for health*. In the next part of this review, an overview is given of several ways of including game-elements to improve adolescents' motivation to train using CBM techniques.

Serious games for health

To understand how applying serious gaming techniques may help motivate adolescents to complete CBM training, let us first look at what constitutes a *serious game*. Unfortunately, despite the recent surge in the number of studies about serious games, there is no consensus yet on what defining elements should comprise a serious game (Bedwell, Pavlas, Heyne, Lazzara, & Salas, 2012). One reason might be the very diverse application of gaming techniques for serious purposes. Granic, Lobel, and Engels (2014) provide a comprehensive review of the many types of serious games and their use in fields such as education, medicine and mental health. They conclude that, although very promising, there are still relatively few serious games specifically aimed at improving mental health. A quick online search on the term “serious games” also reveals that many diverse techniques are used, such as virtual reality and motion capture techniques, to increase physical exercise and activities through gaming (also known as *exergaming*), as well as online games and lab-based games. To narrow down this wide field of serious games towards identifying the useful elements for CBM training, we make several distinctions in the ways gaming techniques can be applied to intervention techniques, such as CBM.

First, there is the focus of the game-development. As a serious game ideally is a combination of a serious component (e.g., a training) and a fun component (i.e., a game), the development of a serious game will usually start from one of these two positions. Coming from the game perspective, one can start with a so called “Off The Shelf” (OTS) game, which often is commercially developed and primarily meant for entertainment, and use it for serious purposes, such as cognitive training. Several studies have looked at the effects of prolonged gaming on cognitive abilities (for review, see Granic et al., 2014), and there is growing support for the notion that, contrary to popular belief in recent years, gaming may also have positive effects. However, when examining OTS entertainment games in more detail, it is hard to disentangle which aspect of the game is responsible for the desired training effect. This may limit the scientific use of OTS entertainment games for developing specific training games (e.g., CBM). Alternatively, one may start with a training procedure or training concept and introduce game-elements to make it more fun and motivating. For example, Merry and colleagues (2012) developed an intervention game called *SPARX*, based on cognitive behavioral therapy (CBT) principles. This game, aimed at adolescents seeking help for depression, proved to be as effective in treating depression as a therapist-administered CBT program. As the field of CBM already possesses a relatively strong scientific evidence base, this would seem to be the option of choice. Interestingly, however, a review by Kharrazi, Lu, Gharghabi, and Coleman (2012) showed that the recent surge of health game publications still often lacks underlying theoretical frameworks.

Another distinction that can be made within the serious games domain, concerns the difference between the explicit messaging, versus a more indirect, implicit varieties of training. For example, Noble, Best, Sidwell, and Strang (2000) introduced game-elements to explicit drug education, which was evaluated as being more fun than the regular method. In contrast, Prins and colleagues (2013) used evidence-based executive function training principles as the basis for their *Braingame Brian*. In this cognitive training with game elements the participant trains executive functions, such as working memory and inhibition through a diverse set of puzzles, while walking around in an extensive virtual world.

Finally, and perhaps most importantly, serious games differ with regard to the intrinsic and extrinsic motivational elements used. Intrinsic motivation can be defined as "doing something because it is inherently interesting or enjoyable" (Ryan & Deci, 2000, p. 55). In terms of motivating game-elements, this would mean that the tasks in the game, such as exploring the level or immersing in the story-line, are motivating or rewarding on their own. In contrast, extrinsic motivation refers to "doing something because it leads to a separable outcome" (Ryan & Deci, 2000, p. 55). In games this is often reflected by various point-based reward systems, such as collecting coins or achieving bonus rewards. While extrinsic motivation is sometimes viewed as a less effective (even if powerful) form of motivation, Ryan and Deci's Self-Determination Theory suggests that the efficacy of extrinsic motivators may depend on a person's internal perceived locus of causality (Deci & Ryan, 1985).

CBM gamification: some examples

As there are as yet very few publications on CBM using game-elements, we will use examples from several ongoing projects to describe a number of steps to introduce game-elements to increase motivation to train in a typical CBM training. One of the most frequently applied CBM techniques is focused on selective attention (CMB-A[1]; MacLeod, Rutherford, Campbell, Ebsworthy, & Holker, 2002). To measure and train selective attention, most CBM-A interventions use varieties of the Visual Probe Task (VPT; MacLeod, Mathews, & Tata, 1986). VPT versions aimed at substance-related attentional bias (e.g., Field et al., 2007; Field & Eastwood, 2005; Schoenmakers et al., 2010) usually use a pair of two visually similar pictures, shown simultaneously on the screen.

This pair consists of a target and a contrast picture, for example, an alcohol-related stimulus, such as a bottle of beer and a neutral picture of a soda, like a bottle of Coke. After a short while, usually 500 ms, a small probe, for example, an arrow, is shown at the center of the position of one of the pictures (depending on the version of the task, the pictures also disappear at this point, showing only the arrow, or they may stay visible, and the arrow is superimposed on one of the pictures, see **Figure 1a**). The arrow may point upwards or downwards, and the participant is instructed to respond to the arrow's direction as quickly and as accurately as possible by pressing the corresponding key on the keyboard. The placement of the target and contrast stimuli (left or right) and the arrow direction (up or down) are random. To measure the attentional bias, the arrow appears in the target stimulus' spot equally as often as in the contrast stimulus' spot. The idea behind the task is that attention is drawn more quickly to and maintained longer at the spot where the object of one's selective attention is located. Thus if the arrow is shown at that same location as is the focus of one's attention, reaction times will be shorter, on average, than when the arrow is shown at the other location. The attentional bias towards the target can be calculated by subtracting the average reaction times on target trials (i.e., when the arrow appears in the same spot as where the target picture was shown) from those on neutral trials. When the VPT paradigm is used for CBM training, the location of the arrow is changed to always match the contrast stimulus' spot (instead of in half of trials in the measurement version). The participant then implicitly learns to focus attention away from the target stimuli, towards the contrast stimuli.

Step 1 – Adding game-elements to the evidence-based training task

As a first step towards incorporating motivating game-elements, different kinds of reward systems can be included in the training. First, *motivating feedback*, such as sounds or animations, can be given after each trial or after a block of trials, telling participants how well they are doing, and optionally how to improve their performance. Similarly, progress bars can be included to show

Footnote 1: For clarity, the following sections will focus on CBM-A examples, to make comparisons more easy. The described techniques do apply to different types of CBM.

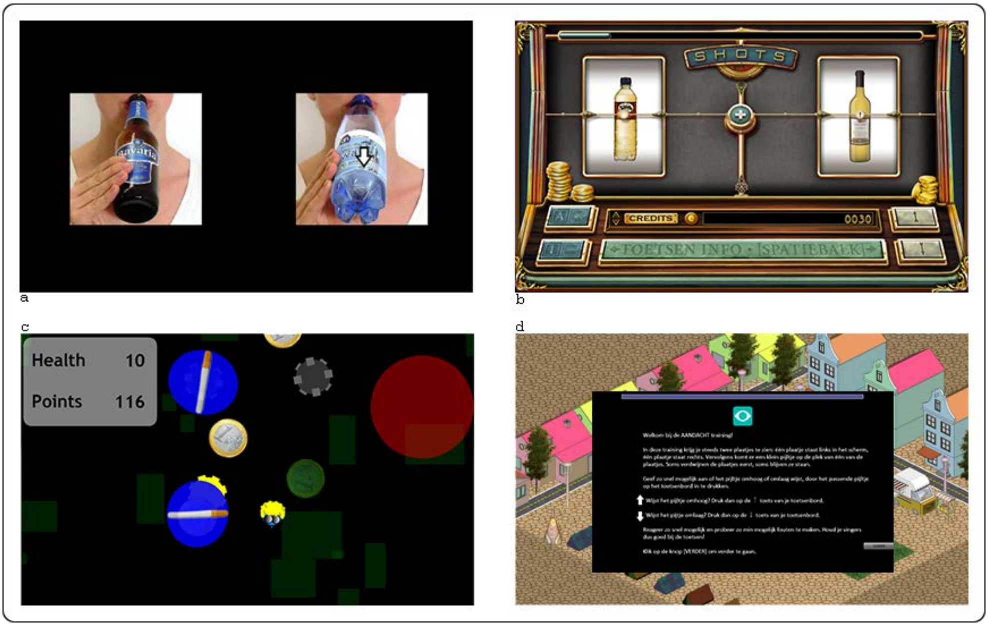

Fig. 1. Four versions of CBM-Attention using different types of game-elements. **a.** The original *Visual Probe Task* (VPT), without game-elements. **b.** *Shots Game*, a game version of the VPT, using the same set-up, but with various game-elements added. **c.** *BombDodger Game*, which is based on scientific attention and approach bias principles, but with a different form of presentation. **d.** *City Builder Game*, where the original VPT, with added progress bar and point reward system, is integrated within a game-shell.

how far along the training session they are. Second, a *point system* can be included, either based on participation (e.g., after completion of the training, the participant is awarded a prize, money or course credits; Anguera et al., 2012; Jaeggi, Buschkuehl, Jonides, & Shah, 2011) or performance (e.g., bonus points for doing well on the task, such as fast correct responses; van Deursen, Salemink, Smit, Kramer, & Wiers, 2013). While it has been suggested that extrinsic rewards, such as money, may hinder performance (Jaeggi, Buschkuehl, Shah, & Jonides, 2014) by undermining intrinsic motivation (Deci, Koestner, & Ryan, 1999), Dovis, van der Oord, Wiers, and Prins (2012) compared several types of rewards in a working memory task and found that children with ADHD were highly sensitive to performance based increases of the chance to win a (relatively large) monetary reward. This discrepancy may be explained by cognitive dissonance theory (Festinger, 1957), in that participants performing a training task without monetary rewards may justify their behavior (i.e., doing the boring task) by changing their conflicting cognition. For example, a participant may not be intrinsically motivated to train, but does so anyway. Then their cognitive dissonance may lead to reasoning along the lines of "why would I do this if it is not rewarding? – I'm doing it; therefore, it must be rewarding after all." In contrast, participants who do get rewarded may not be inclined to change their cognitions, or even worse, reason along the lines that "I am doing this for the rewards; therefore, the training itself is really not that much fun."

Step 2 – Intrinsic integration with the evidence-based training task as a basis

While such game-elements may work to motivate participants to continue training, their motivation will in principle be extrinsic. That is, they may still not like doing the training trials, but the external reward keeps them going. Both psychological theory and the game design literature agree that, although this technique may be effective and is indeed used in many games, the most direct and effective way to motivate is through intrinsic motivation, or in this context, making the participant enjoy doing the gamified training trials. So in order to minimize the distance between the task and the motivating elements, an evidence-based training can also be transformed into a game itself. This makes the training itself more fun, which should increase the intrinsic motivation to train. Interestingly, Dovis et al. (2012) also looked at a gamified version of their task, in which their original working memory task was fully integrated within a digital game. They found that both a regular version with a strong monetary incentive and their gamified version improved performance in children with ADHD, compared to the regular version with feedback only. A CBM-related example can be seen in **Figure 1b**, where the traditional VPT has been transformed into a game. In this game version of the VPT called *Shots* (van Schie & Boendermaker, 2014), participants watch two spinning wheels (visually like a slot machine, but without a gambling element), and when they stop an arrow appears, to which they have to respond, as in the original VPT. Doing well provides the participant with extra coins and the possibility to level up the machine. Another recent example comes from Dennis and O'Toole (2014), who developed a mobile app where participants watch two cartoon characters with different facial expressions (angry versus neutral), which after 500 ms simultaneously disappear into a field of grass. Only one of the faces leaves a trail, to which the participant should respond. Correct responses are rewarded with different jewels, based on speed. Both examples consist of a richer context for the points earned, very close to the actual task. Although there is no elaborate story line or character development (like in Dovis et al., 2012), performing a trial is more fun. An important aspect of this 'intrinsic integration technique' is that, in order to make the training more fun, changes are often made to some of the original features and task parameters. As these features may actually be essential to the workings of the training, removing them may very well render the training less effective. Hence, the adapted cognitive training should always be re-validated.

Step 3 – Intrinsic integration leaving the evidence-based training paradigm intact

To take the intrinsic integration technique a step further, instead of starting with an evidence-based task, such as the VPT, one may also start with the more fundamental principles of the theory behind the paradigm. For example, Notebaert, Clarke, Grafton, and Macleod (2015) used the popular card game 'snap' as a basis for their person-identity-matching (PIM) task. The task features virtual cards with happy and angry faces and requires participants to make matching judgements, based on the identities of the faces. While only loosely based on the attention paradigm, the task was demonstrated to effectively modify attention bias away from threat. In a similar project, T. Pronk (personal communication, July 11, 2014) developed a game called *BombDodger (***Figure 1c**), which has the participant selectively attend to and approach certain neutral stimuli, while disengaging from and avoiding others (in this case, cigarettes). While this theory driven game was praised for being fun to play, there is some discussion about which bias is affected, as both approach tendencies and attention processes are targeted. As a training game, it could thus very well be effective, but as a research tool, it would be hard to disentangle which bias modification led to the effect. Adjusting the game to target one bias at a time could help make the CBM game more specific.

Step 4 – Adding a game-shell around the original evidence-based training task

Instead of adding game-elements to the task, a full game may be added to the task. This involves taking the original training paradigm and leaving it structurally intact, while incorporating it into the look and feel of a surrounding *game-shell*. In these game-shell types of serious games, participants usually receive points for doing well on the original, unadjusted training tasks, which they may then spend during their actual play time within the shell-game surrounding the training, switching back and forth between training and playing. A key aspect of shell-games is that there are game aspects, such as a virtual world, that go beyond, and are unrelated to the training task. For example, by collecting points for doing well on the task, the player is allowed to progress in a story-based game world. Advantages of this design are that it allows the original, evidence-based training paradigm to remain intact and that it enables multiple training paradigms (e.g., both CBM and EF training paradigms) to be used within one game environment.

The *CityBuilder Game* (**Figure 1d**; Boendermaker, Prins, & Wiers, 2013) is an example of this shell-game technique, embedding CBM techniques into an engaging game world. This online shell-game features a virtual world where participants can use points earned through training to build a virtual city of houses, trees, roads, and so forth. The game also includes a social element by allowing the participant to view the cities of other participants, which they can rate with a "thumbs up". The incorporated training tasks can be switched on and off, or set to run as a placebo version. A typical training session takes approximately 30 minutes and consists of a training block, using one of the original tasks like the VPT, with only a point system and a progress bar added, and a subsequent period of game time. During the game breaks, participants are also allowed to do bonus training trials to collect more points. Each correct trial earns the participant points, with bonus points for speed. Initial results indicate that participants enjoy the training environment and are motivated to train more than using a regular training.

Step 5 – Combining intrinsic integration with a game-shell

While the shell-game technique works mostly as an extrinsic motivator, it would seem that a combination of intrinsic and extrinsic game elements could lead to optimal motivation. Although many serious games, including intrinsic integration versions, often do use some form of extrinsic motivators (e.g., a point system; cash or credits for participation), combining an intrinsic integration CBM game with a full shell-game has, to our knowledge, not been attempted before. Perhaps this is because integration of the core CBM elements with intrinsic motivators, as well as a motivating extrinsic reward system, which all match the feel of the game also make this option the hardest to realize. The cognitive control-training Braingame Brian (Prins et al., 2013) could arguably fit within this category, as the original training tasks on which the game-training is based remain intact, while they are also extensively integrated into the game-shell. Verbeken, Braet, Goossens, and van der Oord (2013) and van der Oord, Ponsioen, Geurts, Ten Brink, and Prins (2012) have used Braingame Brian and reported positive training effects in obese children and children with ADHD, respectively. These results in executive function trainings provide a good starting point for applying these techniques to CBM training principles.

Step 6 – CBM using OTS entertainment games

As a final step towards gamification, one may use an actual OTS entertainment game and just measure improvements on the selective attention of the players (e.g., Boot, Kramer, Simons, Fabiani, & Gratton, 2008; Green & Bavelier, 2003). While arguably the most fun for the participant, CBM often includes many disorder-specific stimuli, which may be difficult to incorporate into an existing commercial game. For example, the games used by Green and Bavelier (2003) were mainly action oriented shooters and race games. Adding a substantial number of alcohol pictures to these games, requiring quick and accurate responses, would seem practically impossible. Moreover, the fact that these games were not designed to incorporate the many stimuli used in CBM could render the game unplayable or much less motivating. As such, while it may not be impossible, most OTS entertainment games will be unsuited for CBM training.

Of course, combinations of the formats discussed above are possible, and it is hard to classify existing projects exactly into one of them. Nevertheless, taking the evidence-based VPT paradigm through several steps of gamification may give a good example of the practical possibilities. Given the notion that both ends of the spectrum may be seen as suboptimal for CBM in adolescents, being either too boring or insufficiently evidence based, the optimum may be found somewhere in the middle (e.g., steps 2 and 4, or their combination, step 5; see Figure 2).

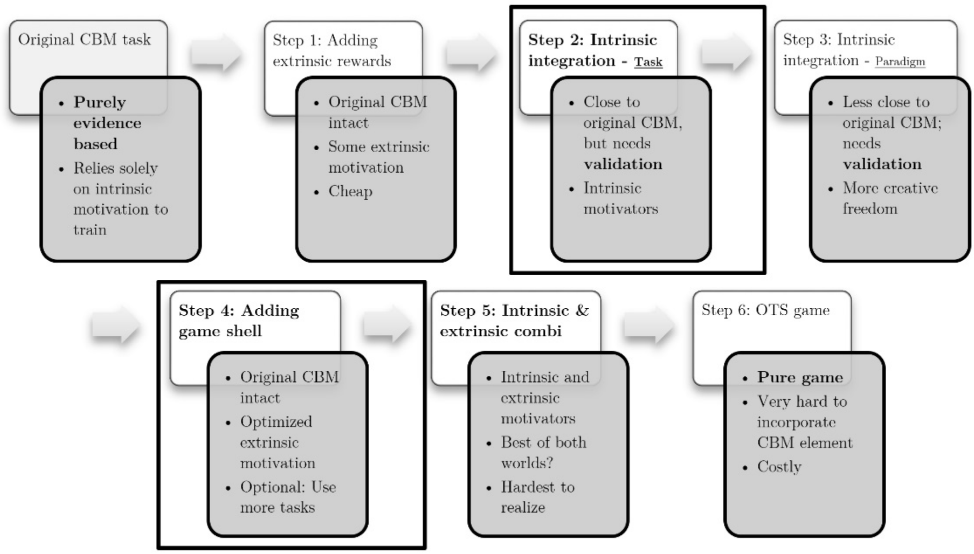

Fig. 2. Six gamification steps from evidence-based paradigm (CBM) to OTS game. For CBM, Steps 2 and 4, and perhaps their combination in Step 5, may be optimal for CBM gamification.

Recommendations

The G-word - Although game-elements may indeed enrich regular CBM training, the level of fun will probably never be comparable to an OTS entertainment game. This of course has to do with the serious nature of the games, for example the many repetitions needed for CBM, but also the often limited budgets for developing them. But the question is how much fun should serious games be expected to be? Indeed, Buday, Baranowski, and Thompson (2012) suggest that a direct comparison with entertainment games should perhaps be avoided. Given the fact that even expensive OTS entertainment games sometimes fail to interest players and are viewed as boring by the gaming community, what can realistically be expected of (relatively) low-budget games that also have to sacrifice fun for training purposes? As the word “game” undoubtedly creates certain expectations in youth, based on their previous gaming experience, perhaps the word should be used with caution when used to describe a serious game for health. Carefully using the word "game" may prevent users from having unrealistic expectations that can lead to disappointment and perhaps even *de*motivation, thus eventually to the opposite of what they should add to the training. It would be interesting to study specific participant expectations with regard to CBM, especially when it is accompanied by game-elements, and their possible effects on motivation and treatment outcome. A related problem that may occur when scientifically studying these training games is that after a game training, there usually are no game-elements during the post-measurement. This sudden lack of motivating elements might actually demotivate participants to do well, potentially even canceling out the possible training effect on the measure. Future research should study whether this is indeed the case, and whether including game-elements in measurement versions of cognitive tasks can solve this problem, without causing too much distraction to render the measurement unreliable.

Quantity and quality - Many cognitive training games aim to motivate participants to keep training as long as possible, as this may increase training efficacy. However, even the best OTS entertainment games eventually lose their appeal to most players. Therefore, an important aspect that should not be overlooked when evaluating a serious game training is whether the added game-elements are motivating enough to not only heighten initial motivation to train, but also to maintain that level of motivation throughout the multiple sessions of training. If the initial motivation is high, but it diminishes over sessions, it may actually start to work against the participant. Moreover, as participants may still have the expectation that doing the game training should be fun, when this is no longer the case, their motivation may even drop below the level of motivation that they would have had without game elements to begin with. As such, adding game elements may work better when training time is relatively short. Therefore, game elements should be carefully matched to the intended number of training sessions, as well as the average session duration. A related question is whether adding game-elements to CBM not only increases the quantity of trials that participants will be motivated to do, but also increases the quality of the performance, in terms of training effects on the underlying ability or bias. If this is the case, then perhaps fewer or shorter sessions are possible. To our knowledge, these issues have never been studied directly.

Critical notes

While the field of CBM games is still largely uncovered, evidence from similar fields shows promise. However, some critical notes also apply. First, preliminary data from Katz, Jaeggi, Buschkuehl, Stegman, and Shah (2014) on working memory training in children seem to suggest that some motivating elements, such as real-time scoring during play, may in fact distract from the training, and can actually lead to reduced task performance. It would therefore be wise to measure the degree to which motivating game-elements add to the cognitive load during task performance, which elements actually add to the training effect, and which are better left unused in the context of CBM games. To our knowledge, this has not been done systematically. Second, intrinsic motivators in games are often reported to be better than extrinsic motivators, and some data indeed seem to suggest this is the case (Habgood & Ainsworth, 2011). However, they are often harder to achieve than extrinsic motivators, in terms of both costs and design. The question therefore remains whether a set of extrinsic motivators might be good enough for CBM training purposes. Or perhaps a combination of both works best. As Deci and Ryan (1985) stated, the level of the perceived extrinsic versus intrinsic nature of a motivator may depend on a person's internal perceived locus of causality. An interesting hypothesis that could be tested is whether the efficacy of the extrinsic motivators in a game training in fact depends on the efficacy of its intrinsic motivators. So perhaps only if a game is intrinsically motivating to someone and immersion is relatively high, extrinsic motivators such as points are relatively more effective. More systematic research is needed to disentangle these two types of motivators before any definitive conclusions can be drawn as to which is more effective for CBM training. Third, we should perhaps more clearly differentiate between the two aforementioned types of motivation involved in CBM. Besides having a motivation to change one's behavior (e.g., maladaptive substance use), there is a related, but separate, motivation to complete a potentially tedious multi-session training in order to do so (Boffo, Pronk, Wiers, & Mannarini, 2015). While it remains unclear whether one needs both in order for CBM to be effective, participants may still need at least some degree of an intrinsic motivation to change their behavior in order for CBM to have any effect (Wiers et al., 2013; 2015). As it seems reasonable to assume that the use of game elements mainly affects *motivation to train*, the awareness of the problem itself may be targeted separately, in order to improve *motivation to change*. This means that, while using the term "game" with caution to prevent disappointment, trainings should also not hide the fact that they actually do have a serious purpose: helping to gain more control over one's substance use. Future research should take these considerations into account, especially when developing CBM games for prevention in younger adolescents. Specifically, the relationship between motivation to train and motivation to change could be further studied to see if and how one affects the other. For example, a very motivating CBM gamification could still fail to increase motivation to change, or it could even have a negative influence (cf. Deci et al., 1999). Finally, when introducing game-elements into CBM interventions, the core CBM mechanisms may become altered to some degree. Given the strong link between CBM-A efficacy and desired clinical outcome reported by Clarke and colleagues (2014), it is essential to validate these new gamified tasks and see how well they affect the targeted cognitive bias compared to the original CBM task.

Conclusions

To our knowledge this is the first review that considers the use of serious gaming techniques as a possible tool to motivate at-risk adolescents to use CBM. With this review we have attempted to give an overview of techniques that can be used to apply game-elements to motivate adolescents to follow through with their CBM training. Although many projects are currently in progress, our main conclusion has to be that at this point there is not yet enough evidence to draw any firm conclusions as to its efficacy. However, results from similar fields such as executive function training do show promise. Serious games may therefore be a promising new way to reach at risk youth (through prevention as well as intervention). While several interesting questions remain unanswered at this point, we feel confident that future studies will be able to address them in the coming years.

Acknowledgements

The authors wish to thank dr. Sebastiaan Dovis for his useful comments on an earlier draft of this article.

Chapter 3
Evaluating Working Memory Training with Serious Game Elements in Adolescents: A Randomized Controlled Trial

This chapter is based on:
Boendermaker, W. J., Gladwin, T. E., Peeters, M., Prins, P. J. M., & Wiers, R. W. (2016). Evaluating working memory training with serious game elements in adolescents: A randomized controlled trial. *Manuscript under review.*

Abstract

Background: Working memory capacity (WMC) has been found to be impaired in adolescents with various psychological problems, such as addictive behaviors. Training of WMC can lead to significant behavioral improvements, but is usually long and tedious, taxing motivation. Therefore, adding Serious Gaming techniques may help improve adolescents' motivation to train and eventually improve their mental health. **Methods:** Eighty-four high school students were allocated to one of three training conditions: a WMC-training, a gamified WMC-training or a placebo condition. WMC, motivation to train, and drinking habits were assessed before and after training. **Results:** Self-reported evaluations completed after the training did not show a preference for the game, but participants in the gamified WMC-training condition did train significantly longer. The game thus successfully increased motivation to train, but this effect faded over time. WMC increased equally in all conditions, but this may have been influenced by the fact that motivation to do the post-training WMC assessment was significantly lower in the gamified WMC-training condition only. The sudden lack of motivating rewards during the post-training WMC assessment after ten sessions of gameplay may have negatively affected participant motivation in this group, specifically. The training did not lead to significantly lower drinking, which is attributed to low drinking levels at baseline. **Conclusion:** This is the first study that shows that serious game elements can help to increase motivation to train WMC in adolescents. It is recommended that future studies attempt to prolong this motivational effect, as it appeared to fade over time.

Introduction

Psychological problems that occur during adolescence are often associated with deficiencies in self- regulation (Gladwin, Figner, Crone, & Wiers, 2011; Martinussen, Hayden, Hogg-Johnson, & Tannock, 2005; Wiers, Gladwin, Hofmann, Salemink, & Ridderinkhof, 2013). For example, working memory capacity (WMC; Baddeley, 1992) and inhibition are often impaired in adolescents with Attention-Deficit/Hyperactivity Disorder (ADHD; Martinussen et al., 2005; Wilcutt, Doyle, Nigg, Faraone, & Pennington, 2005). During adolescence, youngsters typically start experimenting with risky behaviors (Steinberg, 2007). For example, consumption of alcohol usually starts in early adolescence, and often at a much earlier age than is legally allowed (Hibell et al., 2012). Heavy use at this age can lead to school dropout (Singleton, 2007) and can escalate into more severe problems later on, such as substance dependence or addiction. Heavy drinking in youth has been associated with suboptimal cognitive control functions (e.g., de Wit, 2009; Verdejo-Garcia, Lawrence, & Clark, 2008). According to Dual Process Models of Addiction (e.g., Wiers et al., 2007), addictive behaviors emerge when an individual fails to self-regulate the impulsive reactions that develop with heavy substance use. Effects of these reactions on cognitive processing are termed cognitive biases, which can be detected using various implicit measures (Koster, Fox, & MacLeod, 2009; Wiers et al., 2007). Both inhibition (Houben & Wiers, 2009; Peeters et al., 2012) and WMC (Grenard et al., 2008; Thush et al., 2008) have been found to be weaker in heavy drinking youth, thus leading to an imbalanced cognitive system (Wiers, Boelema, Nikolaou, & Gladwin, 2015). As such, early intervention programs aimed at training cognitive control may play an important role in keeping these mental problems at bay. WMC has been considered the most central of cognitive control functions (Kane & Engle, 2002), the ability to adaptively update and monitor representations in working memory (Martinussen et al., 2005). WMC has been the target of many training studies aiming to improve WMC, with some moderate successes in children with relatively weak WMC (Holmes, Gathercole, & Dunning, 2009), such as children with ADHD (for review, see Klingberg, 2010; but see Sonuga-Barke et al, 2013). Increasing WMC has also led to reduced drinking in problem drinkers with strong automatic positive associations with alcohol (Houben, Wiers, & Jansen, 2011), as well as to positive changes in symptoms of trait and test-anxiety, increased inhibitory control, and reduced attention to threat in adolescents (Hadwin & Richards, 2016).

Despite its efficacy in specific adolescent groups, motivation is an important moderator of cognitive training efficacy (van Deursen et al., 2015; Wiers et al., 2007). As cognitive training paradigms can be very long and tedious, with as many as 25 separate sessions for WMC-training (e.g., Klingberg et al., 2005; van der Oord, Ponsioen, Geurts, Ten Brink, & Prins, 2012) motivation to train is likely to decline during training, which may impact the training's efficacy. Incorporating serious game elements into the cognitive training paradigms may help adolescents to endure, as such elements may be better at grasping and retaining adolescents' attention and increasing their motivation to complete the training (Dovis, van der Oord, Wiers, & Prins, 2012). There have been several attempts to gamify cognitive training paradigms. For example, Prins and colleagues (2013) developed an elaborate game world called Braingame Brian around multiple evidence-based executive function training principles. Positive training effects with this gamified training have been found in obese children (Verbeken, Braet, Goossens, & van der Oord, 2013) and children with ADHD (van der Oord et al., 2012). However, Braingame Brian is primarily aimed at school-aged children, and may be perceived as too childish by adolescents. For this reason we developed the CityBuilder Game (Boendermaker, Prins, & Wiers, 2013), which is specifically aimed at training cognitive functions and retraining substance-related cognitive biases in adolescents. The game is structured as a so-called 'game-shell' (Boendermaker, Prins, & Wiers, 2015; **Chapter 2** in this thesis), where the user receives points for doing well on several training tasks. The training tasks themselves are only minimally adjusted from the original evidence-based training paradigms to fit the game environment. The points collected during training can then be spent during separate periods of play time within the shell-game surrounding the training programs (Boendermaker et al., 2013). Finally, an element of alcohol-related context was added to the training task by briefly showing a picture of alcohol during the encoding phase of the task. This picture may be more distracting to heavier drinkers, potentially making the training a little more challenging. The current study describes the results of ten sessions of alcohol-related WMC-training using the CityBuilder Game. Three conditions are compared (all including the alcohol-related context): the gamified WMC-training using the CityBuilder Game (henceforth referred to as the 'gamified' condition); a non-gamified WMC-training (the ‘standard’ condition); and a non-gamified placebo training, not expected to improve WMC (the ‘placebo’ condition). The primary focus is on how the game helps to motivate adolescents to continue training. Adolescents in the gamified WMC-training condition are expected to show a higher motivation to train, compared to adolescents in the non-gamified conditions, as measured by explicit ratings and by the time spent training. A problem that can occur in an experimental comparison of a serious game training with an original training paradigm is that both conditions complete the same assessments before and after training, while only the gamified WMC-training condition has been rewarded during training. This may lead participants in the gamified WMC-training condition to be less motivated to complete the less rewarding assessment tasks, potentially affecting the assessment of training effects (Boendermaker et al., 2015; **Chapter 2** in this thesis). Therefore, motivation for doing the pre- and post-training assessments is also evaluated using self-report questions. Further, the training is expected to increase WMC, relative to placebo. As a secondary outcome we will look at potential transfer effects of WMC-training to drinking behavior, where participants are expected to drink less alcohol after the training. Finally, potential influence of the alcohol picture on performance will be analyzed, and it is expected that heavier drinkers in the sample may make more errors following the alcohol picture, as it may be more distracting to them than to lighter or non-drinkers.

Methods

Participants

Participants were 84 adolescents from a high school in the Netherlands between 13 and 16 years of age (M_{age} = 13.7 years; *SD* = 0.7; 40.4 percent boys). Participants trained during normal school hours in fourteen groups of six students. They were randomly assigned to one of the three training conditions stratified for age, gender and school class. Participants in each group were allocated to the same condition to prevent them from comparing the gamified and non-gamified versions among each other. There were 24 students (four groups) in the placebo condition, 30 students (five groups) in the standard WMC-training condition, and another 30 students (five groups) in the gamified WMC-training condition. The training took place in two cohorts: seven groups (two placebo, three standard WMC-training, and two gamified WMC-training) finished training before Christmas break; the other seven groups started after Christmas. The second cohort filled in an additional questionnaire assessing motivation to train after each session. Due to personal reasons, three students (two from the placebo and one from the standard WMC-training condition) dropped out during the study. The study was approved by the Ethics Committee of the University of Amsterdam (Protocol number 2012-COP-2449).

Design and Procedure

Before the study, parental consent was obtained for each adolescent, and at baseline, adolescents were informed about the training procedure and the reward for participation, which was a maximum of 15 Euros, consisting of 5 Euros for doing the baseline and post-training assessments and an additional 1 Euro for each completed training session. To keep the students motivated to continue training in all conditions, it was announced that the training money was only awarded when a minimum of eight training sessions was completed. The training was done on university laptops in groups of six adolescents, whereas the assessment sessions, which were the same in all three conditions, were done in groups of twelve students on school PCs. After the baseline assessment, participants performed ten daily training sessions on school days during the next two weeks. When a training session was missed because of an important school activity, an extra training session was planned for a total of ten training opportunities per participant. Finally there was a post-training assessment session.

Measures

WMC Assessment - WMC was assessed using the Self-Ordered Pointing Task (SOPT; Petrides & Milner, 1982). In the SOPT, the participant is shown a set of pictures with the instruction to click on a picture they have not clicked on before. Then the pictures in the set are shuffled and the instruction is repeated, until the number of responses equals the number of pictures &presented in the set. The current version used increasingly larger sets of pictures and alternated between sets of pictures of concrete objects (e.g., ball, umbrella) and sets of pictures of abstract objects (e.g., lines and figures), in the following order: 4 concrete (practice); 6 concrete, 6 abstract, 8 concrete, 8 abstract, 10 concrete, 10 abstract, 12 concrete, and finally 12 abstract pictures, respectively. This was done to gradually increase the difficulty of the task, in order to avoid a ceiling effect. The primary outcome measure of the SOPT was the total number of correct responses over all test blocks: a score between 8 and 72, with a higher score indicating better WMC (for reliability and validity, see Ross, Hanouskova, Giarla, Calhoun, & Tucker, 2007).

Standard WMC-Training - this training was based on the chessboard task by Dovis et al. (2012), but with the inclusion of several alcohol pictures. The alcohol picture was intended to slightly distract participants, with an expected greater effect on participants who drink more alcohol, as their attention can be biased towards alcohol pictures (Field et al., 2007; Schoenmakers et al., 2010), which can affect task performance on a working memory task (Gladwin & Wiers, 2011). Participants were presented with a four by four grid of green and blue squares (each 120 by 120 pixels large, presented in a chessboard pattern) that lit up in a specific sequence of three or more squares. The instruction was to remember this sequence, then mentally re-order the squares to reproduce first all green squares, and then all blue squares, in the order in which they appeared. To ensure re-ordering was necessary in each trial, each sequence showed at least one blue square before one or more green squares. During trials, the sequence length was first announced in the center of the screen for 1500 ms. Each square then lit up for 1500 ms, with a 1000 ms interval between squares, until the current number of squares was shown. A 540 by 540 pixel image of a beverage containing alcohol was shown for 600 ms during one of the inter-square-intervals (selected randomly). Different sets of ten unique pictures were used for this purpose during each training session. All alcohol stimuli were taken from the Amsterdam Beverage Picture Set (ABPS; Pronk et al., 2015). To prevent the use of memory strategies, the mouse cursor was invisible during the trials. After each trial there was always feedback about whether the answer was correct, followed by a self-paced button to go to the next trial. During feedback, a progress bar also indicated how far they were during the session. When two consecutive trials were answered correctly, the next sequence length was increased by one square. Similarly, when two consecutive trials were answered incorrectly, the next sequence length became one square shorter, with a minimum of three squares. Each training session lasted approximately 30 minutes and consisted of a minimum of 40 trials, with a first 3-minute break after the first block of 20 trials and a second 3-minute break after the second block of 20 trials. After the second break participants received the option to continue with another block of training trials, or wait for 5 minutes before going back to class collectively with the other participants in the group.

Placebo WMC-training - this version was exactly the same as the standard WMC-training, except that the sequence length was always kept at three, to prevent a training effect whilst presenting a visually similar experience (cf. Houben et al., 2011). As the overall duration of the task was shorter because of keeping the sequence length at a low level, participants in the placebo condition did a minimum of 50 trials per session (25 per training block).

Fig. 1. The CityBuilder Game. Left pane: The Game screen; Right pane: the WMC-training task is presented overlaying the game screen. During instructions, the game is shown in the background (as pictured); when the trials start, the background blacks out entirely.

Gamified WMC-training - this version was also similar to the standard WMC-training, but was embedded within a game context, the CityBuilder Game (Boendermaker et al., 2013; see Figure 1). As in the other conditions, each training session started with a block of training trials, but in the gamified WMC-training condition participants received points for correct trials. These points were saved up until the break and could then be spent as game money to buy houses, roads, trees, and other objects to build a virtual city. A social element was included in the game by letting participants view the cities built by other players, which they were also allowed to rate with a "thumbs up".

After the break (which lasted exactly 3 minutes), the game automatically reverted back to another training block, followed by the second break. As the final training block did not include any play time, the extra collected points could only be spent during the next training session.

TABLE 1. PROCEDURE DURING TRAINING SESSIONS

Version of WMC-training	*Standard*	*Placebo*	*Gamified*
Training block 1 [9 minutes]	20 trials	**25 trials**	20 trials
Break 1 [3 minutes]	Continue training Read magazine Enjoy break in silence	Continue training Read magazine Enjoy break in silence	Continue training Read magazine Enjoy break in silence **Play the game** [a]
Training block 2 [9 minutes]	20 trials	**25 trials**	20 trials
Break 2 [3 minutes]	Continue training Read magazine Enjoy break in silence	Continue training Read magazine Enjoy break in silence	Continue training Read magazine Enjoy break in silence **Play the game**
Optional extra training block [5 minutes] [b]	Continue training Read magazine Enjoy break in silence	Continue training Read magazine Enjoy break in silence	Continue training Read magazine Enjoy break in silence

a During the first session, participants in the gamified WMC-training condition always started the first break with a one-minute introduction to the game.
b During the last session, the second break lasted 8 minutes, and the extra training block was omitted, as there was no next session to spend the bonus points in.

As shown in Table 1, the breaks between the training blocks were introduced to match the time between participants in all conditions. All conditions were given three optional activities during the breaks: either continue training, read a magazine or spend the time in silence (cf. Prins, Dovis, Ponsioen, ten Brink, & van der Oord, 2011). Only participants in the gamified WMC-training condition were allowed to use this time to play the game. Training trials done during the break did not count towards the minimum of trials during the fixed training blocks. A final block of optional bonus trials was included as an additional behavioral measure of motivation to train. The same options were provided as during the breaks, but now also those in the gamified WMC-training condition were not allowed to play the game.

Motivation to train - Besides the number of bonus trials per session as a behavioral measure of motivation, two self-report questions were also added in the second cohort: “How much were you looking forward to this task?” and “How much did you like this task?”, both scored on a 10-point scale ranging from 1 (not at all) to 10 (very much). After the training, participants were asked about their previous game experience, as well as how much fun they thought the training had been, on a 5-point Likert scale from 1 (a lot of fun) to 5 (very boring), how difficult they thought the training had been, on a 5-point Likert scale from 1 (very difficult) to 5 (very easy) and how often they would continue doing the training if it would be made available at home, on a 5-point Likert scale from 1 (never) to 5 (very often).

Alcohol use - As heavy drinking does occur at this age in the Netherlands (Hibell et al., 2012), a brief personal interview version of the Alcohol Timeline Followback procedure (TLFB; Sobell, Sobell, Klajner, Pavan, & Basian, 1986; Wiers, Hoogeveen, Sergeant, & Gunning, 1997) was used to measure alcohol consumption per day over the past ten days. The personal interview was used to offer participants a more private and secure environment, compared to the computer room. Additionally, potential alcohol-related problems were assessed with the Alcohol Use Disorder Identification Test (AUDIT; Saunders, Aasland, Babor, de la Fuente, & Grant, 1993), the Rutgers Alcohol Problem Index (RAPI18; White & Labouvie, 2000) and the Five Shot Questionnaire (Seppa, Lepisto, & Sillanaukee, 1998). The AUDIT includes ten multiple-choice questions about alcohol consumption and alcohol-related problems. The overall score ranges between 0 and 40, with a score of 8 or higher indicating an increased risk of alcohol-related problems. The RAPI18 is an 18-item questionnaire for assessing problem drinking specifically among adolescents. Each item concerns a statement about the frequency of an alcohol related problem occurring during the past year, scores on a 4-point Likert scale ranging from 0 (never) to 3 (more than 5 times). The Five Shot Questionnaire contains five multiple-choice items about alcohol use. The maximum score is 7, with a score over 2.5 indicating alcohol misuse of alcohol dependence.

Additional baseline measures - To check for baseline differences in IQ, a sub-selection of 30 items from Raven's Standard Progressive Matrices (RPM; Raven, Raven, & Court, 2004) was assessed. Baseline differences in reward sensitivity were checked using the Dickman's Impulsivity Inventory (DII; Dickman, 1990), which contains 23 true-or-false questions divided over two subscales, and the Behavior Inhibition System / Behavior Approach System scale (BIS/BAS; Carver & White, 1994; Franken, Muris, & Rassin, 2005), which consists of 20 Likert-scale items over four subscale. Finally, basic family structure, family drinking habits and parental social economic status were also assessed.

Results

Before running the analyses, all dependent variables were screened for univariate outliers based on having a score removed more than 3 standard deviations from the group mean, which resulted in the exclusion of two outliers on the AUDIT, one on the Five Shots Questionnaire, four on the TLFB, one on the SOPT sum score, two on the RAPI18, one on the SOPT, one on the BAS Fun seeking and two on the BAS Reward responsiveness subscales, respectively. Due to technical problems, the data of four participants at baseline, TLFB data for three participants and RPM data for one participant were lost.

Baseline

Missing data - Due to various reasons (e.g., illness), some participants missed one or more sessions, but were allowed to continue training. Five participants, however, did not complete the full assessments and were therefore excluded from the relevant pre-post analyses. In total 29 participants finished in the gamified WMC-training condition; 27 in the standard WMC-training condition and 20 in the placebo condition. Levels of drinking were very low at baseline. The average sum score on the AUDIT was 1.2 (SD = 2.3), with 52 participants having a sum score of 0, and 0.4 (SD = 1.1) on the RAPI18. Therefore it was

TABLE 2. SAMPLE CHARACTERISTICS BY GROUP

	Standard	*Placebo*	*Gamified*	*Total*	*p*
Total (male)	30 (13)	24 (9)	30 (12)	84 (34)	0.908
Age (in years)	13.9 [0.7]	13.6 [0.7]	13.7 [0.7]	13.7 [0.7]	0.388†
AUDIT (sum score)	1.4 [2.5]	0.9 [1.9]	1.2 [2.4]	1.2 [2.3]	0.669†
Drinking in family	2.1 [1.0]	2.6 [1.3]	2.6 [1.2]	2.4 [1.2]	0.196†
Five Shot (sum score)	0.4 [0.8]	0.2 [0.4]	0.3 [0.5]	0.3 [0.6]	0.987†
RAPI18	0.4 [1.1]	0.1 [0.5]	0.7 [1.4]	0.4 [1.1]	0.440†
TLFB (drinks last 10 days)	0.4 [0.6]	0.3 [0.6]	0.1 [0.3]	0.2 [0.5]	0.204†
RAPI18	0.4 [1.1]	0.1 [0.5]	0.7 [1.4]	0.4 [1.1]	0.440†
SOPT sum unique pictures	55.0 [4.9]	55.1 [4.7]	56.2 [4.5]	55.4 [4.7]	0.551
DII: Dysfunctional subscale	-2.6 [4.5]	-4.9 [4.0]	-2.2 [4.4]	-3.1 [4.4]	0.074
DII: Functional subscale	2.8 [4.7]	1.6 [4.3]	2.9 [3.9]	2.5 [4.3]	0.438†
BIS/BAS: BAS Drive subscale	9.6 [2.2]	10.3 [2.3]	9.9 [2.4]	9.9 [2.3]	0.502
BIS/BAS: BAS Fun seeking subscale	10.4 [1.9]	10.8 [1.0]	10.5 [1.6]	10.5 [1.6]	0.648†
BIS/BAS: BAS Reward Responsiveness subscale	15.4 [2.3]	15.7 [1.7]	16.4 [1.6]	15.9 [1.9]	0.198†
BIS/BAS: BIS subscale	18.2 [3.4]	18.3 [3.3]	18.1 [3.3]	18.2 [3.3]	0.977
Raven Progressive Matrices	19.6 [2.8]	21.4 [3.2]	20.6 [3.4]	20.5 [3.2]	0.146

Table shows means, with [standard deviations between brackets].
† indicates a non-parametric Kruskal-Wallis test was applied due to violation of normality.
AUDIT, Alcohol Use Disorder Identification Test; TLFB, Timeline Followback, shows the number of standardized drinks during the ten days prior to the pre-training assessment; SOPT, sum score on the Self-Ordered Pointing Task; DII, Dickman Impulsivity Inventory; RAPI18, Rutgers Alcohol Problem Index; BIS/BAS, Behavior Inhibition System / Behavior Approach System scale.

TABLE 3. TRAINING OUTCOMES BY GROUP

	Standard	*Placebo*	*Gamified*	*Total*
SOPT sum score pre-training	55.4 [4.5]	55.1 [4.8]	56.2 [4.5]	55.6 [4.6]
SOPT sum score post-training	57.4 [5.3]	57.3 [4.3]	55.9 [4.8]	56.8 [4.9]
TLFB sum score pre-training	0.3 [0.6]	0.2 [0.5]	0.1 [0.2]	0.2 [0.5]
TLFB sum score post-training	0.3 [0.7]	0.1 [0.2]	0.0 [0.2]	0.1 [0.4]

Table shows means, with [standard deviations between brackets].
SOPT, sum score on the Self-Ordered Pointing Task; TLFB, Timeline Followback, shows the number of standardized drinks during the week prior to the assessment.

decided to include these two long-term measures again after training, to make sure this finding was stable. This was the case. There were no baseline differences in age, gender, IQ, impulsivity, or WMC between conditions. See Table 2 for a full overview of baseline measures.

Effects of training

There was a main effect of time on WMC as measured with the SOPT sum score, $F(1,72) = 6.033$, $p = 0.016$, $\eta_p^2 = 0.077$, but no effect of training condition, $F(2,72) = 0.052$, $p = 0.949$ (see Table 3). When a threshold of at least 8 out of 10 training sessions (cf. 20 out of 25 sessions in Houben et al., 2011) was used as a cut off for the effects analyses, resulting in the exclusion of two participants in the gamified WMC-training condition, three in the standard WMC-training condition, and two in the placebo condition, these effects did not change. There was no training effect on alcohol consumption as measured with the TLFB over time, $F(1,62) = 1.410$, $p = 0.240$.

Alcohol picture analysis

To determine the influence of the alcohol picture during the encoding phase of the training trials, we looked at the percentage of errors made specifically on squares that directly followed the alcohol picture versus the error percentage on squares that did not directly follow the alcohol picture. Overall, error percentages were different between the training conditions, but this was mainly because in the placebo condition, all sequences had exactly three squares, and thus fewer errors were made. When this condition was excluded, the standard and gamified WMC-training conditions did not differ (see Table 4). The average sequence length also did not differ between the standard and gamified WMC-training conditions (the placebo condition was not included as all sequences had exactly three squares). As the level of drinking was very low in this sample, no relationships between error percentage and alcohol consumption were found (all ps > 0.05).

TABLE 4. ERROR PERCENTAGES ON SPECIFIC SQUARES

	Standard	*Placebo*	*Gamified*	*Total*	*p*
Including placebo condition, n = 84					
Error percentage on squares directly following the alcohol picture	24.2 [5.8]	5.8 [3.8]	24.7 [4.8]	19.1 [9.8]	<0.001†*
Error percentage on squares not directly following the alcohol picture	24.3 [5.9]	6.8 [4.3]	24.2 [5.2]	19.3 [9.5]	<0.001†*
Ratio of errors directly following the alcohol picture over those that don't	1.00 [0.08]	0.85 [0.13]	1.03 [0.09]	0.97 [0.12]	<0.001†*
Without placebo condition, n = 60					
Error percentage on squares directly following the alcohol picture	24.2 [5.8]		24.7 [4.8]	24.5 [5.2]	0.329†
Error percentage on squares not directly following the alcohol picture	24.3 [5.9]		24.2 [5.2]	24.2 [5.5]	0.657†
Ratio of errors directly following the alcohol picture over those that don't	100.3 [8.2]		103.0 [8.7]	101.6 [8.5]	0.215
Mean sequence length	5.5 [0.8]		5.6 [0.7]	5.5 [0.7]	0.408

Table shows means, with [standard deviations between brackets].

† indicates a non-parametric Kruskal-Wallis test was applied due to violation of normality.

* indicates $p < 0.001$.

TABLE 5. MOTIVATIONS BY GROUP

	Standard	*Placebo*	*Gamified*	*Total*	*p*
How much fun was the training?	3.7 [0.7]	3.2 [0.9]	3.1 [0.7]	3.3 [0.8]	0.006†**
Would you like to have the training at home? (% yes)	7 %	0 %	17 %	9 %	0.103
How often would you train at home?	1.4 [0.7]	1.3 [0.5]	1.7 [0.8]	1.5 [0.7]	0.128†
How much were you looking forward to this task [the SOPT]? (change score)	1.3 [1.8]	1.0 [0.8]	-0.6 [1.2]	0.4 [1.6]	0.003**
How much did you like this task [the SOPT]? (change score)	1.1 [1.8]	0.5 [1.1]	-0.1 [1.8]	0.4 [1.7]	0.212
Number of training sessions completed	8.8 [1.1]	8.4 [1.1]	9.1 [0.8]	8.8 [1.0]	0.040*

Table shows means, with [standard deviations between brackets].

† indicates a non-parametric Kruskal-Wallis test was applied due to violation of normality.

* indicates $p < 0.05$; ** indicates $p < 0.01$.

Motivation

Table 5 features several measures of motivations by group. Despite a slight trend that suggests more participants preferred to have the game at home compared to the non-gamified versions, it was the standard WMC-training that participants rated as more fun to do, $H(2) = 10.093$, $p = 0.006$, compared to both the gamified ($U = 233.0$, $z = 3.145$, $p = 0.002$) and the placebo version ($U = 410.5$, $z = 2.128$, $p = 0.033$). Motivation to do the SOPT assessment increased over time in the non-gamified conditions, but it decreased in the gamified WMC-training condition, a difference that was significant, $F(2,28) = 7.363$, $p = 0.003$, $\eta_p^2 = 0.345$. Post hoc comparisons using the Tukey HSD test indicated that the mean score for the gamified WMC-training condition ($M = -0.6$, $SD = 1.2$) was significantly lower than both the standard WMC-training ($M = 1.3$, $SD = 1.8$) and the placebo condition ($M = 1.0$, $SD = 0.8$). A similar pattern of results was observed for the change in the level of fun on the SOPT, but these did not reach significance. Finally, there was a difference in the average number of training sessions completed between conditions, where adolescents in the gamified WMC-training condition completed significantly albeit slightly more sessions on average than participants in the two non-gamified conditions.

3

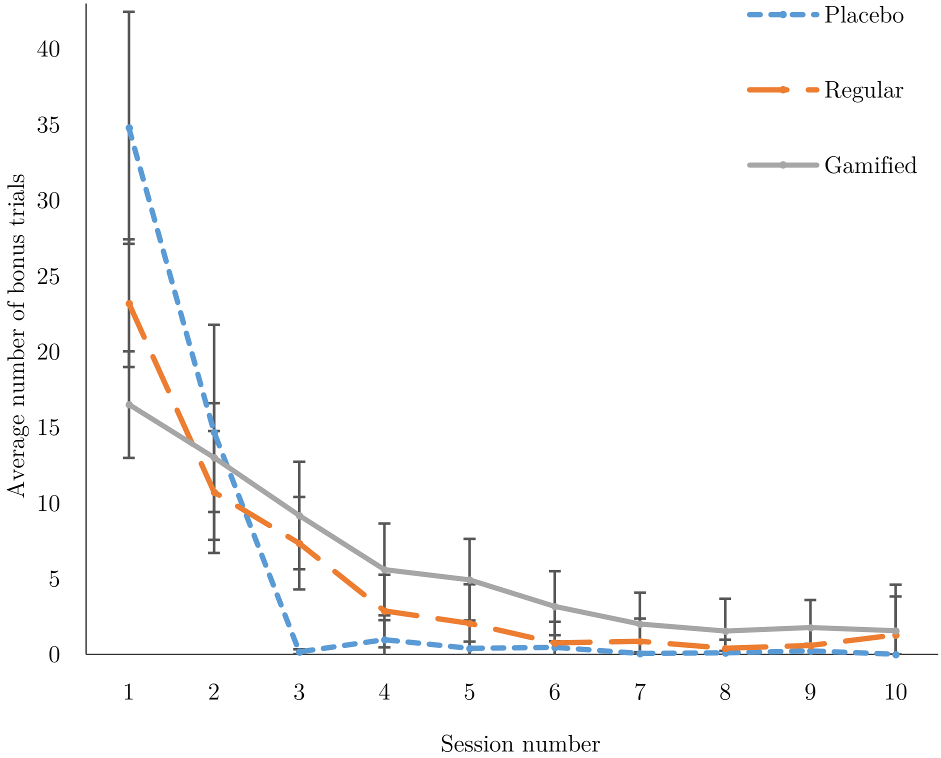

Fig. 2. Average number of bonus trials per session. Error bars indicate 95 % confidence interval.

TABLE 6. AVERAGE NUMBER OF BONUS TRIALS PER SESSION

Session	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*
Placebo	34.8	14.7	0.2	1.0	0.4	0.5	0.1	0.1	0.2	0.0
	[19.2]	[17.8]	[0.4]	[3.2]	[1.1]	[1.9]	[0.2]	[0.3]	[0.8]	[0.0]
Standard	23.2	10.7	7.3	2.9	2.0	0.8	0.9	0.4	0.6	1.3
	[11.8]	[11.1]	[8.4]	[6.6]	[7.0]	[3.8]	[3.9]	[1.4]	[2.5]	[3.4]
Gamified	16.5	13.0	9.2	5.6	4.9	3.2	2.0	1.5	1.8	1.6
	[9.8]	[10.0]	[9.9]	[8.5]	[7.5]	[6.5]	[5.8]	[5.7]	[4.7]	[4.7]

Table shows means, with [standard deviations between brackets].

As another measure of motivation to train, we looked at the total number of bonus trials done during each session (i.e., during both breaks as well as in the final, optional training block), where we numbered the sessions per participant (see Figure 2 and Table 6). For this analysis we used a multiple step approach. As the count variable (number of bonus trials) had a skewed distribution, but not all sessions had many zeros, a Poisson distribution was used rather than zero inflation. Robust Maximum Likelihood (MLR) was used as an estimator to account for the non-normality. The first step taken was a Confirmatory Factor Analysis (CFA) on the total number of bonus trials during each session. As session 1 showed much higher numbers of bonus trials in all conditions, compared to the following sessions, the CFA did not converge when session 1 was included and it was therefore excluded from the analysis. The resulting CFA on sessions 2 through 10 showed that all factor loadings were significant. Due to the nature of the Poisson model, using numerical integration, no standardized factor loadings are available. The second step involved looking at the overall effect of condition on the latent session factor, which was significant: $B = .444$, $SE = .088$, $p < .001$, indicating more bonus trials were done in the gamified WMC-training condition, compared to the other conditions.

In the final step we looked at change over time using a growth model of sessions 2 through 10, again with the bonus trials count variables as latent growth indicators. Several models were compared, first constraining groups to be equal or not (i.e., assuming there were, or there were no group differences), and subsequently constraining only the slopes to be equal or not (i.e., assuming there were, or there were no differences in the decrease of bonus trial counts), and the intercepts to be equal or not (i.e., assuming there were, or there were no baseline differences in bonus trial counts). The best model fit in terms of Akaike's Information Criterion (AIC; Akaike, 1973), as well as the Bayesian Information Criterion (BIC; Schwarz, 1978), was found for the model with free (decreasing) slopes, but with constrained (equal) intercepts for the standard and gamified WMC-training conditions, $AIC = 2758$; $BIC = 2782$. In this model the placebo training's intercept is at 0.667, while both the standard and gamified WMC-training's intercepts are at 1.219; and slope coefficients are -2.855 for the placebo, -1.782 for the standard, and -0.859 for gamified WMC-training, respectively. Note that due to the nature of the count model used here, these coefficients do not represent the actual number of bonus trials, but should rather be interpreted relative to each other, for example, the decrease is much steeper in the placebo condition, compared to the gamified WMC-training condition.

Discussion

In this study we investigated the beneficial effects of a serious game environment on adolescents' motivation to do cognitive training. While no relevant differences were found in the primary outcome measure (WMC), several interesting findings were obtained regarding motivation to train. First, the self-reported motivation measures after the training was completed showed mixed results, with participants having a slight preference for the standard WMC-training. This may indicate that participants did not like the game more than they liked the standard WMC-training, but it can also mean that they merely lost interest over time. Other than the non-gamified training versions, the gamified WMC-training, being presented as a game, likely increased participants' expectations of its entertainment value. If the game then did not fully satisfy these expectations over the ten sessions of training, this may have influenced the motivation-assessment after the training. As such, it is advisable to assess motivation to train at multiple points in time, to see if there might be an initial effect that fades over time. This can be achieved with the behavioral measure of motivation: the number of training trials done beyond the minimum amount required. This number was higher in the gamified WMC-training condition than in the non-gamified conditions, but it also declined over time in all conditions. As such, the gamified WMC-training version was found to motivate adolescents to train more intensively over the course of the ten training sessions, compared to the non-gamified versions. This is important when adolescents need to participate in a relatively long and tedious cognitive training paradigm, such as WMC-training. An important question for future research is how this motivational effect can be extended.

The second motivational finding concerns participants' motivation to perform well on the study's main cognitive outcome measure: the pre- and post-training WMC assessments (SOPT). Although WMC was found to increase over time in all training conditions, which could indicate a test-retest effect, where participants' performance increased due to having done the task before, motivation to complete the task had increased after the training in the non-gamified conditions, but had decreased in the gamified WMC-training condition. This finding is in line with our hypothesis that the rewarding nature of the gamified WMC-training condition may negatively affect motivation to complete assessment tasks afterwards. Although it is unclear if, and to what degree, this motivational effect may have influenced the assessment of the actual training gain, it does have important implications for future research aiming to validate serious games, compared to their non-rewarding, original counterparts. Incorporating the assessment task in the game and having a mini-assessment at the start of every training session (cf. van Deursen, Salemink, Smit, Kramer, & Wiers, 2013) is one option to prevent decline in motivation for the post-assessment in the gamified WMC-training condition. However, this may also intensify the entire training program by prolonging its overall duration.

The results presented in this paper do have to be interpreted with some caution, due to several limitations. First, no training effects were found on drinking behavior, however alcohol use was very low at baseline in this sample. As it obviously could not get much lower through training, no inferences on the effects of (gamified) cognitive training on drinking behavior should be made based on the current study. It would be interesting for future intervention research to include adolescents with cognitive deficits and at risk for problematic alcohol use (Peeters et al., 2015). Second, when comparing the active training conditions, there were no discernable effects of the alcohol pictures presented during training trials on the percentage of errors made during these trials, nor did they affect the average sequence length. When the active training conditions were compared to the placebo condition, the latter showed a notably lower percentage of errors on squares directly following the alcohol picture. This could be due to the easiness of trials in the placebo condition, so that presentation of a distractor resulted in a more optimal level of arousal, but further research is necessary to disentangle this effect. Although the alcohol pictures may have inadvertently introduced a priming effect, which was not assessed separately, they were presented in the same manner in all conditions, and no effects on drinking were found. Nevertheless, future studies that incorporate alcohol pictures in their WMC-training should consider assessing, for example, attentional bias towards alcohol before and after exposure, especially if a future training study is done in heavier drinkers. Third, despite the fact that we did find a WMC-training effect, several studies report optimal cognitive and behavioral training results (e.g., reduced alcohol intake; Houben et al., 2011) with around 15-25 sessions of training (Klingberg, 2010; Morrison & Chein, 2011), rather than the ten sessions presented here. As the game's benefit to participants' motivation to train had already faded over ten sessions, future studies are encouraged to design motivating games aimed at adolescents that incorporate training for at least that many sessions.

Despite these limitations, to the best of our knowledge the current study is the first to demonstrate that WMC-training in adolescents can benefit from the use of game elements by increasing motivation to train. It follows that the challenge for future research will be in trying to prolong this effect, for example by making bigger, more immersive games that last longer (although this is quite a challenge, even in commercial gaming). By closely monitoring the levels of motivation throughout the study, as well as by managing participants' expectations about the entertainment value of the training, which may still be an important factor in determining the training outcome, more insight may be acquired into the specific effectiveness of the use of game elements in cognitive training. Finally, future research could also apply gamified WMC-training in specific at-risk groups, such as adolescents who have specific difficulties with traditional training approaches due to attention or motivation related problems. Moderation analyses can then be used to reveal individual differences in the effectiveness of the gamified training, identifying those who could benefit the most from these motivational features.

Acknowledgements

The authors wish to thank the students involved in the data collection for this study: Nathan Bleijenberg, Bouwien Westerhuis, Emmily Harwig, Marcia Dominicus, and Wendy Kuijn.

Chapter 4
Cognitive Training in Adolescents: A Comparison of Training Reinforcements

This chapter is based on:
Boendermaker, W. J., Nikolaou, K., & Wiers, R. W. (2016). Cognitive training in adolescents: A comparison of training reinforcements. *Manuscript in preparation for publication.*

Abstract

Heavy drinking remains popular among young adults and is an important source of health-related problems. Cognitive training aimed at increasing inhibitory control may help to keep their drinking habits in check. Eighty heavy alcohol drinking students were trained to consistently withhold a behavioral response in the presence of alcohol-related stimuli. To see whether the training could benefit from specific reinforcement techniques, four training variants were compared. The regular version of the alcohol-specific inhibition training included only minimal feedback on performance. A point-reward system was added to the training in the second condition, whereas a digital game world in which these points has monetary value was used in the third condition. Finally, the fourth condition was a control condition in which also minimal feedback was provided, but the inhibition training was general in nature, and not alcohol-specific. Alcohol-specific inhibition was found to increase in all training conditions, but more so in the alcohol-specific training than in the general inhibition condition. The point-rewarded and game variants of the training did not increase this training effect. A significant decrease in alcohol intake was found after the training in all conditions.

Introduction

Despite the well-known adverse effects of prolonged heavy alcohol use (e.g., Kuntsche et al., 2013), heavy drinking remains a popular activity during late adolescence and early adulthood (Hibell et al., 2012). This may relate to the fact that adolescents seldom see their drinking habits as problematic or harmful (Johnston, O'malley, Bachman, & Schulenberg, 2012; Wiers, van de Luitgaarden, van den Wildenberg, & Smulders, 2005), focusing on short-term rewards, rather than the long-term negative effects. Research has shown that heavy alcohol users indeed tend to overvalue alcohol-reinforcers, relative to other available rewards (Goldstein & Volkow, 2002). Alcohol-related stimuli can acquire motivational salience through classical conditioning, and start to attract more attention, initiating approach behavior that can lead to the consumption of alcohol. These changes in attention and approach tendencies are also referred to as cognitive biases (for reviews, see Field & Cox, 2008; Wiers, Gladwin, Hofmann, Salemink, & Ridderinkhof, 2013). In addition to augmented cognitive biases associated with the processing of alcohol-predicting cues, heavy alcohol use is also characterized by suboptimal top-down cognitive control processes, that under normal conditions would regulate cognitive and behavioral responses associated with alcohol and alcohol cues (Wiers et al., 2013).

Recently, research has focused on developing training paradigms that aim to modify cue-elicited cognitive biases, either by reducing the salience of alcohol-predictive-cues, or by increasing top-down cognitive control processes. One such procedure used a variant of the Go/No-Go paradigm to train participants to consistently withhold a behavioral response in the presence of alcohol-related stimuli (Houben, Havermans, Nederkoorn, & Jansen, 2013; Houben, Nederkoorn, Wiers, & Jansen, 2011). They showed that repeatedly pairing No-Go cues with images depicting alcoholic beverages reduced alcohol consumption in the week following training and increased negative implicit attributions assigned to these stimuli in heavy drinking students. The training did not affect behavioral inhibition in the presence of alcohol-related stimuli, as assessed using a modified Alcohol Stop-Signal Task (ASST), nor did it affect behavioral approach responses to alcohol-related stimuli, assessed using an Alcohol Approach Avoidance Task (AAT). Interestingly, the reduction in post-training weekly alcohol-use was mediated by the increase in the negative implicit attributions and not via an effect on response inhibition on the ASST (Houben et al., 2013).

The aim of the current study was to examine how different ways of reinforcing correct alcohol-specific inhibition during training may impact response inhibition. Heavy alcohol drinking students were presented with a standard Go/No-Go training paradigm with images of alcohol-containing drinks and pictures of drinks that do not contain alcohol, where we distinguished between three levels of reinforcement. In the Minimal-Feedback alcohol-specific inhibition training condition (MFB), directed feedback was given only after incorrect responses. In the Point-reinforced alcohol-specific inhibition training condition (PTS), points were given as rewards for correct responses, with additional bonus points for speed. Finally, as these points alone were relatively arbitrary, a Game-reinforced alcohol-specific inhibition training condition (GAM) was used where the same point-rewards gained value in a game (CityBuilder; Boendermaker, Prins, & Wiers, 2013) that was presented after each block of training. Finally, a general inhibition control training condition (CON) was added. This variant was similar to the MFB condition, but here the images of alcohol- and non-alcohol related beverages were evenly paired with Go- and No-Go-cues. While participants did practice general inhibition in this condition, it was not aimed at alcohol, and it therefore served as a control condition. These expectations are in line with a recent meta-analysis by Allom, Mullan, and Hagger (2016) on the effects of inhibitory control training. They observed with regard to effects on health behavior, (1) behavior-specific training tasks based on the Go/No-Go paradigm resulted in medium effect sizes, whereas behavior-specific training methods based on the Stop-Signal paradigm resulted in small effect sizes; and (2) effects of neutral Stop Signal training methods were not significant. Although neutral Go/No-Go training studies were not included in the meta-analysis, this finding does suggest that the effects of the neutral Go/No-Go training condition on health behavior can be expected to be inferior to that of the alcohol-specific training conditions.

Interestingly, Dovis, van der Oord, Wiers, and Prins (2012) made a similar comparison of types of rewards in a working memory assessment task among children with ADHD and found them to be highly sensitive to performance-based increases of the chance to win a (relatively large) monetary reward, as well as rewarding in terms of play time in an accompanying computer game.

In the current study, however, we will look directly at using different reward systems as specific training reinforcers, and at potential benefits they may have to training effects.

There were three main outcome measures in this study. First, to directly assess training effects, inhibition was assessed using a similar Go/No-Go assessment task both before and after training. Similarly, levels of weekly alcohol intake before, during and after the training were assessed. We expected to find stronger training effects for both of these measures the stronger the reinforcement: GAM > PTS > MFB, with minimal or no training effects in the CON condition. Finally, Blood-Oxygen-Level-Dependent (BOLD) responses were assessed before and after training using a visual cue-reactivity task. These data will not be presented in this chapter but will be presented in a separate paper.

Methods

Participants

Eighty moderate-to-heavy social drinkers (defined below) between the ages of 18 and 23 participated in this study (40 females, $M_{age} = 20.0$ years, $SD = 1.5$). They were all undergraduate university students who were recruited to the study via online advertisements placed on the University of Amsterdam's (UvA) participant-recruitment website, and via paper advertisements placed in faculty buildings. Advertisements stated that social alcohol drinkers were required for an fMRI experiment designed to test the mechanisms by which computer-based cognitive training affects people's attentional processes. To prevent the risk of disappointing participants who were not allocated to the GAM condition, the game aspect of the study was not mentioned here. Participation was reimbursed with either course credits or a monetary reward of 50 euros.

Study inclusion required: (a) being between 18 and 23 years of age; (b) reporting drinking an average of ten or more standard units of alcohol per week over the past six months (assessed using the Alcohol Use Questionnaire, AUQ; Mehrabian & Russell, 1978); and (c) scoring above 8 on the Alcohol Use Disorder Identification Test (AUDIT; Saunders, Aasland, Babor, de la Fuente, & Grant, 1993), indicating hazardous or problematic drinking. To make sure any heavy users in the sample also received feedback on their levels of alcohol use, a printed handout was given to all participants on the last day of training. This document featured personal feedback tailored to their own reported use, and was accompanied by verbal feedback explaining heavy users that prolonged drinking of more than 14 and 21 (females/males, respectively) units per week is highly dangerous for their physical health. Additional inclusion criteria were normal or corrected-to-normal visual acuity, being right-handed, and having a BMI within the normal range (18-30). All participants fulfilled the inclusion criteria during screening.

Exclusion criteria were a history of psychiatric or neurological problems, and being on any medication for any psychological or physical condition at the time of the study (including paracetamol and antibiotics, but excluding the contraceptive pill; assessed using an adapted version of the Nuffield Hospitals Medical History Questionnaire (www.nuffieldhealth.com) which covers past and present physical and psychiatric health status, including any current medication). Additional exclusion criteria were pregnancy, trying to conceive or breastfeeding; and MRI contraindications (i.e., having any metal implants, teeth braces or bridges, tattoos above the shoulder, or a cardiac pacemaker). All participants provided written informed consent at the start of the study, and the study was approved by the Ethics Committee of the University of Amsterdam (Protocol Number: 2014-DP-3620).

Design/Procedure

Participants were randomly assigned to one of the four training conditions, stratified by gender. A digital platform designed for the administration of experimental procedures (LOTUS) was used to make sure this allocation was double-blind, and that the experimental procedure was the same for all participants. The experiment comprised of five phases, presented in the following order: (a) baseline assessments; (b) pre-training assessments; (c) three separate training sessions; (d) post-training assessments; and (e) a final follow-up assessment. These phases were completed by each participant during three separate visits to the experimental facilities, with the exception of the follow-up session, which was completed over the internet. The baseline and pre-training assessments, and the first training session took place during the first visit to the lab. Participants were then invited back twice, within three to four, and seven to eight days later, to complete the second and third training sessions, respectively. The post-training assessments were administered during the third visit, immediately after the third training session. The pre- and post-training assessments were identical to one another, and included the (f)MRI scanning procedure (not reported here). Finally, seven to eight days after the final lab visit, participants were asked to undergo a short follow-up session online to examine the effects of the training on alcohol use and alcohol craving a week after the end of the training.

Baseline assessments phase (a)

Upon arrival to the experimental facilities of the UvA, participants provided written informed consent, and underwent a dummy scan to familiarize themselves with the scanning environment. They then logged on to the LOTUS environment, for the first time, using their email address and a unique password that they themselves provided.

A series of baseline assessments was then administered to ensure that the training groups were well matched at baseline on: demographic information (age, gender, years in education, etc.), lifetime and past-month substance use (including nicotine, alcohol, and cannabis), assessed using the CORE Drug and Alcohol Survey (Core Institute; Presley, Meilman, Cashin, & Lyerla, 1996), and degree of nicotine dependence, using the Modified Fagerström Tolerance Questionnaire (MFTQ; Prokhorov et al., 2000). Participants also completed self-report indices assessing their motives for drinking alcohol using the Drinking Motives Questionnaire (DMQ; Cooper, 1994; Grant, Stewart, O'Connor, Blackwell, & Conrod, 2007), and their motivational state related to reducing their alcohol use (Readiness to Change Questionnaire, RCQ; Rollnick, Heather, Gold, & Hall, 1992). Four additional measures were administered at baseline to ensure that the groups did not differ on self-report indices of trait impulsivity (Barratt Impulsiveness Scale; BIS; Patton, Stanford, & Barratt, 1995) and attentional control (Attentional Control Scale, ACS; Derryberry & Reed, 2002), and on cognitive indices of stopping behavior (Visual Stop-Signal Task, VSST; Nikolaou Critchley, & Duka, 2013) and working memory capacity (Reversed Digit Span Task, RDST; e.g., Wechsler 1981). The VSST and RDST were administered outside the LOTUS environment (VSST was run on Cogent 2000; www.vislab.ucl.ac.uk/cogent.php; and RDST was a pen and paper task).

Pre (b) and post-training (d) assessments phases

Both before and after training, participants completed the Desire for Alcohol Questionnaire (DAQ; Kramer et al., 2010; Love, James, & Willner, 1998), a short version of the Timeline Followback Questionnaire (TLFB; Sobell & Sobell, 1990; short version: Wiers, Hoogeveen, Sergeant, Gunning, 1997), and an alcohol Go/No-Go task (e.g., Boendermaker, Boffo, & Wiers, 2015; **Chapter 5** in this thesis; Houben et al., 2011; 2013), in order to assess training-induced effects on alcohol craving, weekly alcohol use, and inhibition in the presence of alcohol-related images, respectively. These assessments were followed by the scanning procedure, which consisted of performing an alcohol cue-reactivity task (CRT; see below for details) during acquisition of T2*-weighted images covering the whole brain. Finally, an Alcohol-Valence Implicit Associations Task (AV-IAT; e.g., Greenwald, McGhee, & Schwartz, 1998) was completed. Technical problems on this task resulting from the use of Adobe Flash led to unexplained, less accurate response times (measured in milliseconds) for most participants: an unusually large percentage of reaction times (93.2 percent at pre- and 95.2 percent at post training) appeared to be ending in zeroes, where roughly ten percent would be expected. Consequently, data derived from this task were deemed unreliable and were thus not analyzed, and will not be discussed further.

Follow-up assessments phase (e)

The DAQ and TLFB were completed for the third and final time seven days after the post training assessment.

Training phase (c)

The Alcohol Go/No-Go training task was based on the alcohol Go/No-Go task used by Boendermaker and colleagues (2015; **Chapter 5** in this thesis), see Figure 1 (top pane). During each of the three training sessions, participants in all conditions received the same basic task instructions, followed by a series of 352 trials, divided over two training blocks: The first block started with four neutral filler trials (as a start-up buffer), followed by 172 training trials with a short break halfway to 'take a breath and continue'. Halfway the session, a three-minute break was taken, followed by the second block, again starting with four neutral buffer trials and then another 172 training trials with a short halfway break. A progress bar was included to show progress through the task. Each trial started with a 50 by 50 pixel white fixation cross, presented in the center of the black screen for 500 ms. Following the fixation cross, a centered 500 by 500 pixel picture was presented for 1500 ms, with an 80 by 80 pixel visual Go or No-Go cue superimposed in one of its corners. Finally, a black screen interval of 500 ms was used before the next trial started. Each picture presented a photograph of a beverage that either contained alcohol ('alcohol picture', 78 trials per block) or did not contain alcohol ('non-alcohol picture', 78 trials per block) or a neutral filler picture of office supplies (16 trials per block). The filler pictures were used to slightly mask the contingency between image content and Go/No-Go cue (cf. Houben, Schoenmakers, & Wiers, 2010; Schoenmakers et al., 2010). The superimposed visual cue was a black letter 'p' or 'f', presented inside a black circle on an 80 by 80 pixel white square, indicating whether to press the spacebar as quickly as possible (i.e., a 'Go cue'), or do nothing and wait for the image to disappear on its own (i.e., a 'No-Go cue'). The cue location was determined randomly, to avoid creating a predictable cue location, and the matching between cue type (letter 'p' or 'f') and cue indication (Go-signal or No-Go-signal) was counterbalanced over participants. Whenever a key other than the default response key (spacebar) was pressed, or whenever a mistake was made (i.e., a response was given following a No-Go cue or no response was given following a Go cue), this was indicated, followed by a brief repetition of the task instructions. Following this feedback, the same trial was repeated until a correct response was given.

In the **CON** condition, specifically, each picture was presented equally often with both cues, meaning that participants had to give a Go-response to alcohol pictures equally often as to non-alcohol pictures in 50 percent of the beverage trials, and a No-Go-response to alcohol pictures equally often as to non-alcohol pictures in the other half of the beverage trials. In the **MFB** condition, all alcohol pictures were paired with a No-Go cue, whereas all non-alcohol pictures were paired with a Go cue. The **PTS** condition was similar to the MFB condition with the addition of points as a reward for correct responses. For correct Go-trials, the faster the response, the more training points were awarded: one training point for a reaction time (RT) over 1200 ms; two points for an RT between 1200 and 900 ms; three points for an RT between 900 and 600 ms; four points for an RT between 600 and 300 ms, and five points for a

Fig. 1. The Go/No-Go task. Top pane: presentation during assessment and training (full screen). Middle pane: instructions for the embedded Go/No-Go training task over-laying the CityBuilder Game screen in the GAM condition. After the instructions, the game in the background faded out to black. Bottom Pane: Impression from the CityBuilder Game.

reaction time below 300 ms. As correct No-Go responses (i.e., no response) do not have a reaction time, the last amount of points awarded for a correct Go-response (i.e., between one and five points) was repeated until the next correct Go-response. A specific explanation of the point system, including a visual example, was added to the basic task instructions at the start of the task. Finally, the **GAM** condition incorporated the PTS condition inside an online game shell called The CityBuilder Game (Boendermaker et al., 2013). The game allowed participants to spend their training points as coins to build a virtual city of little houses, trees, roads, and so forth (see Figure 1, bottom pane). Participants could also view cities made by other participants, which they were also allowed to rate with a "thumbs up". When the training was started, participants were logged into the game, greeted with a welcoming message explaining the game was part of the training and they could earn points by doing well. Each session had a target amount of points to be reached (e.g., 875 points in the first session), and an explanation of the game was provided after the first training block. Then immediately, training started (i.e., no playing was allowed before training and points were acquired). Then the training screen was superimposed on the game screen, where the training-specific instructions were shown as in the PTS version. During the three-minute break between the first and second training block, the total number of points was shown. In the first session, a tutorial for the game was provided, after which three minutes of play time were awarded (matching the break time in the other conditions), where participants could spend their points to build a virtual city and compare theirs with that of others. After the break, the second training block automatically commenced. After the second training block, there was another three to six minutes of play time allowed, after which the application was closed and the participant returned to the project screen.

Measures - Baseline assessments

Demographic information - Nine multiple choice questions enquired about participants' age, gender, nationality (including those of their mother and father), mother tongue, marital status, years in education and income.

Alcohol and drug use and related problems - The Alcohol Use Questionnaire (AUQ; Mehrabian & Russell, 1978) gives an estimate of the average number of weekly alcohol-units consumed over the previous six months (a glass of wine is measured as 1.5 units; a pint of beer/cider as 2.4 units; a shot of spirit as 1 unit; and a bottle of alcopops as 1.7 units). An overall score was calculated based on the sum of weekly alcohol-unit consumption. Townshend and Duka (2002) have previously demonstrated that the AUQ is a reliable measure of drinking quantity. High scores indicate increased average weekly alcohol use.

The Alcohol Use Disorders Identification Test (AUDIT; Saunders et al., 1993) was designed to identify individuals with harmful or hazardous alcohol consumption. It consists of 10 questions measuring alcohol use, and an individual's assessment of other's feelings towards their alcohol consumption. The present study used the total AUDIT score, with high scores reflecting high severity of alcohol use.

We administered the sections of the CORE Drug and Alcohol Survey (Core Institute; Presley et al., 1996) that assess the number of occasions in a person's lifetime, and in the past 30 days, in which they had consumed a particular substance. The questionnaire assessed use of 22 substances (derived from the main drug categories, including: alcohol, tobacco, marijuana, cocaine, ecstasy etc.). Responses are coded in a categorical way and range from 0 = 0 occasions to 10 = More than 91 occasions. High scores on a particular substance are indicative of more frequent use.

A version of the Fagerström Tolerance Questionnaire (Fagerström, 1978) specifically modified for younger populations (the Modified Fagerström Tolerance Questionnaire; MFTQ, Prokhorov et al., 2000) was used to assess the level of nicotine dependence. It includes seven items, and uses a four-point Likert scale for all items, except for one item on smoking during the first two hours of the day. A total score was calculated, and high total scores were indicative of greater nicotine dependence. The modified FTQ scale was reported to be a valid measure for adolescent smokers (Prokhorov et al., 2000). The Modified Drinking Motives Questionnaire - Revised (Modified DMQ-R; Grant et al., 2007) consists of 28 items that assess the reasons why individuals consume alcohol. Responses on each item are made on five-point Likert scale ranging from 'Almost never/never' to 'Almost always/always. Each item loads on one of five factors that include: social, coping-anxiety, coping-depression, enhancement, and conformity. A score was calculated for each factor. The

DMQ-R was reported to be a reliable and valid measure of undergraduates' drinking motives (Grant et al., 2007).

Other baseline measures - The Readiness to Change Questionnaire (RCQ; Rollnick et al., 1992) is a twelve-item questionnaire designed to measure the "stage of change" that an excessive alcohol-drinker is in. Responses are on a five-point Likert scale ranging from 'Strongly disagree' to 'Strongly agree'. Here we only calculated an overall "Readiness to change" score, whereby high scores were indicative of greater willingness to modify drinking behavior. The Barratt Impulsiveness Scale (BIS; Patton et al., 1995) is a 30-item questionnaire designed to measure three aspects of impulsivity: (a) non-planning impulsivity or the inability to plan and think carefully; (b) motor impulsivity or acting on the spur of the moment; and (c) attentional impulsivity or the inability to focus on the task at hand. Items are rated on a four-point Likert-type scale ranging from "rarely/never" to "almost always". Higher scores on each factor loading represent greater levels of impulsive behavior. The Attentional Control Scale (ACS; Derryberry & Reed, 2002) is a twenty-item questionnaire that consists of nine items assessing the ability to focus attention, and eleven items that measure the ability to shift attention. Responses are given on a four-point Likert scale ranging from 'almost never' to 'always'. High scores on each factor loading index improved attentional ability. The Reversed Digit Span Task (RDST; e.g., Wechsler 1981) is a pen and paper assessment of verbal working memory. On each trial of the task, participants listened to a sequence of digits, presented once. They were then required to repeat the sequence in the reversed order (e.g., 4125, becomes 5214). If they made an error, a different sequence of the same length was presented. Two consecutive errors on the same sequence-length terminated the task. Conversely, upon correct reversed recall, the sequence-length of the following trial increased by one digit. The sequence-length of the first trial was three digits. The task consisted of seven sequence-lengths in total (with two attempts at each sequence length). One point was awarded for correct reversed recall on each sequence-length, yielding a highest obtainable score of seven. A percentage of correct sequence lengths out of the possible seven is used for analyses. The Visual Stop-signal task (VSST; Nikolaou et al., 2013) is based on the standard auditory stop-signal task, but in the visual modality, and as such, measures a persons' ability to stop an initiated behavioral response. Participants were instructed to respond, as quickly and as accurately as possible, to the direction of a frequently occurring green arrow (Go-stimulus), and to be as accurate as possible at withholding their response when the green arrow turned red (Stop-stimulus). Each trial began with the presentation of a central fixation cross on a grey background for a jittered duration of 1200-1500 ms. The stimulus-display that followed, always began with the presentation of the Go-stimulus on a grey background, which either remained on screen for a total stimulus-display duration of 800 ms (Go-trials), or was replaced, after a variable stimulus onset asynchrony (SOA), by the Stop-stimulus (Stop-trials). Initially, the SOA was set at 200 ms and increased by 50 ms following successful Stop-trials (i.e., Stop-trials in which participants successfully inhibited their response), or decreased by 50 ms following unsuccessful Stop-trials (i.e., Stop-trials in which participants were not able to inhibit their response). This staircase procedure resulted in a relatively equal number of successful and unsuccessful stops (Verbruggen & Logan, 2009). Stop-signal reaction time (SSRT; Logan & Cowan, 1984) was calculated by subtracting the mean SOA from the average reaction time to correct Go-trials (Verbruggen & Logan, 2009). Further dependent measures from the VSST included Go-latency and Go-accuracy. Participants completed a total of 160 trials of the VSST (120 Go-trials and 40 Stop-trials) in eight minutes.

Measures - Pre/Post-training and Follow-up assessments

The Timeline Followback (TLFB; Sobell & Sobell, 1990; Wiers et al., 1997) was used to assess the number of alcohol units participants drank on each day for seven days. It was administered during both the pre/ and post/training assessment phases, as well as during the online Follow-up assessment session. A total TLFB score was computed for each assessment phase representing the total number of alcohol units consumed (a) during the week before the training; (b) the week during the training; and (c) the week after the training. Higher scores indicate heavier use. The shortened Desires for Alcohol Questionnaire (DAQ; Love et al., 1998) is a fourteen-item questionnaire that assesses subjective feelings of alcohol craving. Each question is rated on a five-point scale ranging from “strongly disagree” to “strongly agree” (cf. Kramer et al., 2010), and is classified under one of four factors: “positive and negative reinforcement”; “strong desires and intentions”; “mild desires and intentions”;

and "control over drinking". Each factor was treated as an independent measure, and we additionally computed a total DAQ score for each assessment phase (i.e., pre/post-training and follow-up). The Alcohol Go/No-Go assessment task was similar to the CON condition of the training task, except that it and consisted of one block with ten practice trials with neutral filler images, which was followed by an assessment block which started with four neutral buffer trials and then 84 assessment trials, each showing either a picture of a beverage containing alcohol or one without alcohol. A unique set of pictures was used for the assessment, to ascertain the effect of training on untrained pictures. The Alcohol Cue Reactivity Task (CRT) was an event-related task participants completed while inside an fMRI scanner, where they were presented with pictures of alcohol, non-alcohol and animal pictures. Outside of the scanner, once scanning was completed, participants were asked to rate the images of alcohol-related and non-alcohol-related beverages they had just seen during the CRT, on a ten-point scale, based on: (a) "Willingness-to-pay" (WTP; "How much money would you be willing to pay for the item depicted in the picture"); (b) "Degree-of-Craving Alcohol" (DCA; "How strong is your desire to drink alcohol right now?); and (c) "Degree-of-Craving the depicted Drink" (DCD; "How much would you like to drink the drink in the glass?"). This rating task was administered via the E-Prime software (Psychology Software Tools, Pittsburgh, PA), outside of the LOTUS environment. Results of this task are not included in this chapter, as they will be presented alongside the fMRI data.

Picture Stimuli

During training, a total of eighteen unique alcohol pictures, eight-teen unique non-alcohol pictures, and six unique neutral filler pictures were presented in random order during the first block of each session, while a second, unique set of the same size was used in the second block of each session. These two sets of pictures were used during each training session. Another, unique set of six alcohol pictures, six non-alcohol pictures, and five filler pictures was used for both the pre- and post-training assessment, to ascertain the effect of training on untrained pictures. All pictures were 500 by 500 pixels in size and depicted either the beverage alone or with an actor drinking it. All pictures originated from the Amsterdam Beverage Picture Set (ABPS; Pronk, van Deursen, Beraha, Larsen, & Wiers, 2015).

Statistical analyses

Baseline assessments - Participant's scores were deemed outliers and were thus removed if they deviated by more than three standard deviations from the group mean. This led to the exclusion of one outlier in the RTs on VSST-GO trials; three outliers on the VSST-GO errors; five outliers on the VSST-STOP errors and two outliers on the VSST-SSRT. In order to assess the entire range of substance use, outliers on any of the substance use measures were not excluded from the analyses. Next, all baseline measures were checked for normality. Those that were not normally distributed were analyzed using non-parametric tests. Measures that were normally distributed were analyzed using analyses of variance (ANOVA), using condition (4) as the between-subject factor. Significant effects were explored with Bonferroni-corrected t-tests or, when test assumptions were violated, non-parametric tests.

For each group the number of smokers and non-smokers were determined, based on having an MFTQ score higher than 0 and then this ratio was compared between the groups, indicating they were divided equally over conditions, $\chi^2(3) = 0.417$, $p = 0.937$. In addition, in order to make sure that smokers in each group did not differ in their degree of dependence, we compared the smokers only based on the total MFTQ score, which indicated there was no difference found between the groups: $H(3) = 0.164$, $p = 0.983$. Accuracy scores on the GNG training and assessment data were checked to see if there were any indications that participants did not do the training or the assessment task properly. Participants were marked when their response accuracy deviated by more than 3SDs from the sample mean, and their accuracy was less than 60 percent correct. Based on these criteria, none of the participants were deemed outliers (the lowest accuracy scores were 91 percent correct during training and 92 percent correct during the assessments).

Pre/Post-training and Follow-up assessments - To determine the training induced effects on the ability to withhold a behavioral response in the presence of alcohol-related stimuli, the Go/No-Go data were analyzed using two indices. The first index involved calculating the ratio of the percentage of correct alcohol-No-Go trials over the percentage of correct non-alcohol-No-Go trials for the pre- and post-assessments. The second index concerned the total percentage of correct alcohol-No-Go trials during the baseline assessment, each training session and the post-training assessment. Percentages were used as the CON condition had fewer alcohol-No-Go trials compared to the other conditions, and the assessment version also had fewer alcohol-No-Go trials than the training sessions. Training effects on drinking behavior were compared between conditions from pre- to post-training and follow-up. For this we used the total number of standardized units of alcohol consumed during the last seven days, as measured by the TLFB, as well as the sum scores on the total DAQ, as well as its four subscales.

Due to a technical problem, data for the Go/No-Go post-training assessment was not saved for one of the participants in the MFB condition. This participant was therefore excluded from all GNG-related analyses. Prior to the analyses, these dependent variables were screened for univariate outliers. However, as there were notable ceiling effects in these data, for example, due to some participants making few or no mistakes on the Go/No-Go assessments, the Analysis of Variance (ANOVA) assumption of normality was violated. Therefore, an Aligned Rank Transformation procedure (ART; Wobbrock, Findlater, Gergle, & Higgins, 2011) was applied to all dependent variables, which is robust against violation of normality and makes it possible to avoid excluding outliers. As part of the ART-procedure, factorial repeated-measures ANOVAs were applied to the generated estimated means using time (with two levels for the Go/No-Go data and three for the TLFB and DAQ data) as a within-subjects factor and condition (4) as a between subjects factor. Significant interactions were analyzed further by applying Mann-Whitney tests on the change scores over time between each time point compared to baseline, using the following three planned contrasts: (1) comparing the training conditions (MFB, PTS, and GAM) with the control condition (CON); (2) the reward-reinforced training conditions (PTS and GAM) with the minimal feedback training condition (MFB); and finally, to look at the added effect of reinforced training with a meaningful (game) reward system, (3) the Point-rewarded training condition (PTS) with the Game-rewarded training condition (GAM).

Results

Baseline assessments

The mean AUDIT score in this sample was 14.9 (SD = 5.5), indicating heavy drinking levels were present. The four groups were matched on all demographic information, drug and alcohol use indices, trait characteristics, and baseline cognitive ability (see Table 1), except on the ACS total score, $F(3,76) = 3.837$, $p = 0.013$, $\eta_p^2 = 0.132$ (no significant Bonferroni-corrected contrasts); on the ACS focusing subscale, $F(3,76) = 3.248$, $p = 0.026$, $\eta_p^2 = 0.114$, where the PTS condition scored higher than the CON condition ($p = 0.048$); the ACS shifting subscale, $F(3,76) = 3.976$, $p = 0.011$, $\eta_p^2 = 0.136$, where the MFB condition scored higher than the GAM condition ($p = 0.009$); on the BIS non-planning subscale, $F(3,76) = 3.382$, $p = 0.022$, $\eta_p^2 = 0.118$, where the GAM condition scored higher than the CON condition ($p = 0.018$); and finally the DMQ Coping-Depression subscale also showed a significant difference between conditions, $H(3) = 8.401$, $p = 0.038$, but then there were no significant Bonferroni-corrected contrasts.

Effects of training on behavior

Mean accuracy scores on the alcohol Go/No-Go assessment task and average scores on the DAQ and TLFB, for each training-group, at each assessment period, are presented in Table 2. The ratio of the percentage of correct alcohol-No-Go trials over the percentage correct non-alcohol-No-Go trials differed significantly between conditions, $F(3,75) = 4.335$, $p = 0.007$, $\eta_p^2 = 0.148$, and also increased over time in all conditions, $F(1,75) = 37.233$, $p < 0.001$, $\eta_p^2 = 0.332$, indicating that overall alcohol specific inhibition increased. The interaction between time and condition was also significant, $F(3,75) = 3.852$, $p = 0.013$, $\eta_p^2 = 0.133$. Planned contrasts indicated that the training conditions combined showed a greater increase in alcohol inhibition over time than the CON condition, $U = 823.5$, $z = 2.504$, $p = 0.012$, $r = 0.280$. The other planned contrasts did not yield significant results.

The percentage of correct alcohol-No-Go trials per session changed increased significantly over time, $F(4,300) = 26.062$, $p < 0.001$, $\eta_p^2 = 0.258$, but this did not differ between conditions, $F(3,75) = 2.113$, $p = 0.106$. The interaction between time and condition was also not significant, $F(12,300) = 1.363$, $p =$

TABLE 1. BASELINE CHARACTERISTICS

	Control (CON)	*Minimal feedback (MFB)*	*Points (PTS)*	*Game (GAM)*	*P*
Total N (male)	20 (10)	20 (10)	20 (10)	20 (10)	
Age (years)	20.5 [1.6]	19.8 [1.2]	20.1 [1.6]	19.7 [1.5]	
AUDIT	13.6 [6.1]	16.2 [6.8]	14.4 [4.6]	15.3 [4.1]	0.500†
AUQ (units)	23.3 [15.0]	25.4 [12.1]	18.9 [6.3]	19.1 [7.3]	0.208†
ACS					
Focusing	21.1 [4.2]	22.6 [3.8]	24.4 [3.7]	21.2 [3.6]	0.026*
Shifting	28.5 [4.2]	31.6 [3.3]	29.5 [4.3]	27.5 [3.6]	0.011*
Total	49.6 [7.1]	54.1 [6.2]	53.9 [6.9]	48.7 [5.7]	0.013*
BIS					
attention	18.5 [4.6]	17.1 [3.0]	17.5 [3.3]	18.0 [3.4]	0.665
motor	23.1 [6.0]	24.0 [5.0]	21.6 [3.6]	22.8 [3.5]	0.485†
non-planning	22.8 [4.1]	24.1 [4.5]	23.9 [4.5]	26.8 [3.7]	0.022*
CORE					
Life: Alcohol	10.8 [0.9]	10.1 [2.5]	10.9 [0.4]	11.0 [0.2]	0.578
Life: Marijuana	4.8 [3.7]	6.3 [4.0]	4.2 [3.1]	5.0 [3.1]	0.419
Life: Cigarettes	7.0 [4.4]	7.3 [4.2]	7.2 [4.4]	6.6 [4.0]	0.961
Month: Alcohol	3.2 [1.2]	3.4 [2.2]	3.2 [2.2]	3.4 [1.4]	0.442
Month: Marijuana	1.5 [0.5]	2.0 [1.7]	1.6 [2.0]	1.5 [0.5]	0.100
Month: Cigarettes	3.3 [3.2]	2.3 [1.8]	2.7 [2.7]	2.8 [2.8]	0.904
Modified DMQ-R					
Social	16.3 [3.5]	15.1 [4.1]	14.8 [3.7]	16.5 [4.1]	0.443
Coping-Anxiety	8.3 [3.0]	8.0 [2.4]	7.6 [1.9]	8.9 [3.3]	0.747†
Coping-Depression	11.9 [3.7]	13.3 [4.9]	11.6 [1.8]	15.2 [5.3]	0.038†*
Enhancement	15.2 [4.2]	16.2 [3.7]	15.4 [4.2]	14.7 [3.2]	0.662
Conformity	5.8 [1.5]	6.4 [1.7]	5.7 [1.0]	6.6 [1.6]	0.279†

Table shows means, with [standard deviations between brackets].

† indicates a non-parametric Kruskal-Wallis test was applied due to violation of normality.

* indicates $p < 0.05$.

AUDIT, Alcohol Use Disorder Identification Test; AUQ, Alcohol Use Questionnaire; ACS, Attentional Control Scale; BIS, Barrett Impulsiveness Scale; CORE, CORE Drug and Alcohol Survey; Modified DMQ-R, Modified Drinking Motives Questionnaire - Revised.

0.183. Planned contrasts on the change scores over time between each time point compared to baseline indicated that the MFB condition performed better, compared to the combined GAM and PTS conditions, at T1-T2, $U = 210.5$, $z = -2.974$, $p = 0.003$, $r = 0.384$, T1-T3, $U = 264.0$, $z = -2.136$, $p = 0.033$, $r = 0.276$, and at T1-T4, $U = 244.0$, $z = -2.449$, $p = 0.014$, $r = -0.316$.

TABLE 1. BASELINE CHARACTERISTICS (CONTINUED)

	Control (CON)	*Minimal feedback (MFB)*	*Points (PTS)*	*Game (GAM)*	*P*
MFTQ					
All participants	0.7 [1.3]	0.6 [0.9]	0.5 [0.6]	0.6 [0.9]	0.983†
Smokers only	2.1 [1.4]	1.5 [1.0]	1.2 [0.3]	1.4 [0.7]	0.413†
Number smokers	7/20	8/20	8/20	9/20	0.937
RCQ	37.0 [5.3]	38.9 [9.0]	38.9 [6.5]	38.1 [6.6]	0.787
RDST	0.5 [0.2]	0.5 [0.2]	0.6 [0.2]	0.5 [0.2]	0.563†
VSST					
GO RT (ms)	460.6 [79.7]	456.7 [88.7]	512.3 [89.0]	449.4 [69.9]	0.071
GO Accuracy (%)	97.8 [1.9]	98.3 [1.9]	98.7 [0.9]	97.9 [2.2]	0.430†
STOP Accuracy (%)	50.4 [4.1]	51.0 [3.8]	52.6 [3.5]	49.1 [6.8]	0.163†
SSRT (ms)	243.1 [40.7]	240.5 [31.3]	236.8 [27.2]	235.4 [34.3]	0.890

Table shows means, with [standard deviations between brackets].

† indicates a non-parametric Kruskal-Wallis test was applied due to violation of normality.

* indicates $p < 0.05$.

MFTQ, Modified Fagerström Tolerance Questionnaire; RCQ, Readiness to Change Questionnaire (a lower score mean a higher readiness to change); RDST, Reversed Digit Span Task; VSST, Visual Stop-Signal Task; GO RT, average reaction times on Go-trials; GO Accuracy, percentage correct Go-trials; STOP Accuracy, percentage correct stop trials; SSRT, Stop-Signal Reaction Time.

Effects of training on drinking measures

The TLFB data indicated a significant decline in the number of drinks over time as measured over the last seven days at baseline, after training and at follow-up, $F(2,152) = 5.150$, $p = 0.007$, $\eta_p^2 = 0.063$, but this was the same in all conditions, $F(3,76) = 1.505$, $p = 0.220$. There was also no significant interaction, $F(6,152) = 0.444$, $p = 0.849$. Results on the DAQ subscales and total score indicated no significant effects over time, between conditions, or interactions (all ps > 0.05).

TABLE 2. PRE/POST-TRAINING AND FOLLOW-UP ASSESSMENTS PRESENTED SEPARATELY FOR EACH TRAINING GROUP

	Control (CON)	*Minimal feedback (MFB)*	*Points (PTS)*	*Game (GAM)*
Ratio of percentage correct alcohol-No-Go trials over percentage correct non-alcohol No-Go trials				
Pre-training assessment	0.999 [0.059]	0.988 [0.071]	0.993 [0.040]	1.006 [0.036]
Post-training assessment	1.011 [0.035]	1.039 [0.063]	1.101 [0.111]	1.060 [0.071]
Percentage of correct alcohol-No-Go trials				
Pre-training assessment	97.9 [3.9]	95.5 [5.8]	97.4 [3.6]	98.1 [2.8]
Training session 1	97.6 [2.5]	98.1 [1.4]	96.9 [2.2]	96.2 [3.0]
Training session 2	97.9 [2.6]	98.2 [1.5]	97.9 [1.5]	97.0 [2.2]
Training session 3	98.4 [2.2]	98.8 [0.7]	97.8 [1.6]	97.4 [1.8]
Post-training assessment	99.0 [2.0]	98.7 [2.7]	98.8 [2.1]	98.1 [3.2]
Percentage of correct non-alcohol-No-Go trials				
Pre-training assessment	98.1 [3.2]	96.7 [3.9]	98.1 [2.8]	97.6 [3.3]
Post-training assessment	98.1 [2.8]	95.2 [4.5]	90.5 [7.9]	92.9 [6.3]
DAQ				
Total sum score				
Pre-training assessment	26.4 [9.0]	26.0 [9.6]	28.3 [7.2]	29.0 [9.2]
Post-training assessment	23.5 [6.5]	27.4 [8.0]	28.4 [6.7]	27.1 [9.3]
Follow-up	27.4 [9.5]	23.7 [9.6]	28.0 [11.9]	25.7 [8.9]
Positive and negative reinforcement subscale				
Pre-training assessment	6.9 [3.1]	6.6 [2.8]	8.3 [2.9]	8.2 [4.1]
Post-training assessment	6.8 [3.2]	7.9 [2.5]	8.6 [2.6]	8.6 [3.7]
Follow-up	7.7 [3.4]	6.8 [2.4]	8.3 [3.5]	8.4 [3.8]
Strong desires and intentions subscale				
Pre-training assessment	5.3 [1.6]	5.5 [2.8]	5.2 [2.7]	5.6 [2.3]
Post-training assessment	4.8 [1.1]	5.7 [2.7]	5.2 [1.5]	5.4 [2.2]
Follow-up	5.7 [2.3]	5.2 [3.1]	6.1 [4.1]	5.5 [2.0]
Mild desires and intentions subscale				
Pre-training assessment	10.1 [5.0]	10.2 [5.0]	11.6 [4.9]	11.1 [3.9]
Post-training assessment	8.9 [3.6]	10.4 [4.0]	11.2 [5.0]	9.8 [4.5]
Follow-up	9.9 [5.0]	8.6 [4.2]	10.2 [5.9]	8.7 [4.0]

Table shows means, with [standard deviations between brackets].

DAQ, Desire for Alcohol Questionnaire; TLFB, Timeline Followback questionnaire.

TABLE 2. PRE/POST-TRAINING AND FOLLOW-UP ASSESSMENTS PRESENTED SEPARATELY FOR EACH TRAINING GROUP (CONTINUED)

	Control (CON)	*Minimal feedback (MFB)*	*Points (PTS)*	*Game (GAM)*
DAQ				
Control over drinking subscale				
Pre-training assessment	4.3 [2.6]	3.8 [2.1]	3.4 [1.5]	4.1 [1.9]
Post-training assessment	3.1 [1.5]	3.6 [2.0]	3.5 [1.8]	3.5 [1.6]
Follow-up	4.1 [2.9]	3.2 [1.7]	3.4 [1.7]	3.3 [1.3]
TLFB: Number of drinks				
Pre-training assessment	18.7 [16.0]	22.2 [15.1]	15.9 [9.8]	19.8 [16.9]
Post-training assessment	15.8 [16.5]	23.3 [17.5]	15.5 [11.0]	20.4 [16.1]
Follow-up	16.1 [17.9]	20.7 [16.2]	12.8 [10.2]	17.9 [20.1]

Table shows means, with [standard deviations between brackets].
DAQ, Desire for Alcohol Questionnaire; TLFB, Timeline Followback questionnaire.

4

Discussion

This study investigated several ways of increasing alcohol-specific inhibition by comparing different levels of reinforcement. Alcohol-specific inhibition increased in all training conditions, but was found to increase the most when the inhibition training was also alcohol-specific, as evidenced by an increase in the ratio of correct alcohol-specific over non-alcohol-specific inhibition. When correct alcohol-specific inhibition was also rewarded with point-rewards, both without the points having any further value, or when these point-rewards held value in a Serious Game accompanying the training, the training effect did not increase significantly beyond giving minimal feedback only. Moreover, the level of correct alcohol-specific inhibition actually increased significantly more over time when there were no point-rewards given.

Although participants reported drinking less alcohol after the training in all conditions, they did not show a change in terms of their self-reported desires for alcohol after the training. As this effect was also visible in the general, non-alcohol-specific inhibition training condition, the decrease in alcohol intake cannot be attributed solely to alcohol-specific inhibition training. As such, despite a stronger cognitive effect in the alcohol-specific inhibition training conditions, this finding could suggest that a general inhibition training may be sufficient to help young adults gain more control over their drinking habits, which is surprising given the lack of a significant effect size found with neutral Stop Signal training methods reported by Allom et al. (2016). In any case, it should be noted that without an experimental condition where participants do not train at all (i.e., a true placebo condition), this notion remains speculation, based on the current data.

The fact that specific point-rewards did not appear to reinforce the training effect, and even seemed to make the training less effective in some aspects, may have to do with the participants' motivation. Adding the wrong, or too many motivating elements can have a distracting, rather than beneficial effect (Boendermaker, Sanchez Maceiras, Boffo, & Wiers, 2016; **Chapter 6** in this thesis), for example by introducing additional cognitive demands. For example, Katz, Jaeggi, Buschkuehl, Stegman, and Shah (2014) found working memory training improvement to be higher when only minimal motivational features were used, rather than fully gamified versions. It could also be argued that adding points to the training was experienced as an extrinsic, rather than an intrinsic motivator, which may have negatively affected participants' overall motivation. As the training became something to do because it lead to a separable outcome (points), rather than for the more personal goal of gaining more behavioral control, it may have had a negative effect on the training outcome, as extrinsic motivation is typically seen as inferior to intrinsic motivation (Ryan & Deci, 2000). However, Lumsden, Skinner, Woods, Lawrence, and Munafò (2016) found point-rewards to be a suitable motivator in a Go/No-Go assessment. As such, more research is necessary to explore under what conditions the addition of game elements, and points in particular, have a beneficial, rather than detrimental effect on training outcomes. Nevertheless, their value as reinforcers of alcohol-specific inhibition may be limited.

There are several limitations to these findings that should be taken into account when interpreting these data. First, the fairly high percentages of correct alcohol-specific inhibition on the Go/No-Go tasks may indicate that participants did not respond as fast as they could have. This may have negatively impacted their training benefit, but this remains a speculation. Second, the fact that there was a higher percentage of alcohol-No-Go trials in the training task (50 percent) compared to the assessment task (25 percent) could have an influence on the comparison of correct alcohol-No-Go trials over time. Although we used percentages here instead of actual trial counts, this may still have skewed the effect to some degree. Third, the generalizing cognitive effects of Go/No-Go training paradigms are usually evaluated on the IAT, which was not possible in this study due to unreliability of the IAT data. This should of course be prevented in future research. Finally, future studies could include more extensive measures of motivation to train and motivation to change drinking behavior to determine the specific effects of these reinforcing (game) elements.

To our knowledge this study was the first to systematically compare the effects of different forms of reinforcement on alcohol-specific inhibition training. As the varying reinforcers did not seem to have an added effect in terms of cognitive gain in this study, future research may benefit from the notion that even without such elements, Go/No-Go training can be effective in increasing inhibitory control and reducing alcohol intake in a heavy drinking student sample. Interestingly, the effects found in this study we robust even when the training was not focused on alcohol-specific inhibition. The current study could be followed-up by investigating these effects in a clinical sample, by looking at generalization effects on other cognitive measures, such as the IAT, but also by further adjusting the game elements used to motivate participants.

Acknowledgements

The training procedure and the majority of tasks and questionnaires used in the present study were administered through the use of an online platform called LOTUS, which was developed by the Technical Support Unit of the University of Amsterdam. The authors wish to thank Bruno Boutin and Tim de Jong for their help in setting up the experimental procedure in LOTUS. We would also like to thank Annelieke Hagen and Sam Prinssen for their help during participant recruitment and testing.

Chapter 5
Exploring Elements of Fun to Motivate Youth to do Cognitive Bias Modification

This chapter is based on:
Boendermaker, W. J., Boffo, M., Wiers, R. W. (2015). Exploring elements of fun to motivate youth to do cognitive bias modification. *Games for Health Journal, 4*(6), 434-443. DOI: 10.1089/g4h.2015.0053

Abstract

Objective: Heavy drinking among young adults poses severe health risks, including development of later addiction problems. Cognitive re-training of automatic appetitive processes related to alcohol (so-called Cognitive Bias Modification [CBM]) may help to prevent escalation of use. Although effective as a treatment in clinical patients, the use of CBM in youth proves more difficult, as motivation in this group is typically low, and the paradigms used are often viewed as boring and tedious. This paper presents two separate studies that focused on three approaches that may enhance user experience and motivation to train: a serious game, a serious game in a social networking context and a mobile application. **Materials and Methods**: In the Game Study, 77 participants performed a regular CBM training, aimed at response matching, a gamified version, or a placebo version of that training. The gamified version was presented as a stand-alone game or in the context of a social network. In the Mobile Study, 64 participants completed a different CBM training, aimed at approach bias, either on a computer or on their mobile device. **Results:** Although no training effects were found in the Game Study, adding (social) game elements did increase aspects of the user experience and motivation to train. The mobile training appeared to increase motivation to train in terms how often participants trained, but this effect disappeared after controlling for baseline motivation to train. **Conclusions:** Adding (social) game elements can increase motivation to train, and mobile training did not underperform compared with the regular training in this sample, which warrants more research into motivational elements for CBM training in younger audiences.

Introduction

Despite the risks involved, use of alcohol among adolescents and young adults remains commonplace (Danielsson, Wennberg, Hibell, & Romelsjö, 2011). Heavy alcohol use at this age can induce significant health problems, school dropout (Singleton, 2007) and increases the risk of alcohol dependence later in life (Bonomo, Bowes, Coffey, Carlin, & Patton, 2004). While the focus of many prevention and treatment programs is on explicit drug education, their efficacy appears to be limited (Foxcroft & Tsertsvadze, 2012; Werch & Owen, 2002). An alternative and less explicit approach comes from the field of cognitive psychology, through Dual Process Models of addiction (e.g., Deutsch & Strack, 2006; Wiers et al., 2007). These models posit that an imbalance between two types of cognitive processes can lead to the development of addiction problems. On the one hand, repeated use leads to the development of automatic reactions towards alcohol-related objects over time. For example, heavy alcohol users often show a biased attention towards alcohol-related stimuli (Field et al., 2007) and are quicker to approach them, relative to non-alcoholic stimuli Wiers, Rinck, Dictus, & van den Wildenberg, 2009). On the other hand, heavy users also tend to have weaker cognitive control abilities, such as working memory (Grenard et al., 2008) and inhibition (Houben & Wiers, 2009). These abilities then fail to regulate the relatively strong automatic tendencies, which can explain the problematic drinking behavior. To restore balance to the system, cognitive control can be trained (Bickel, Christensen, & Marsch, 2011; Houben, Nederkoorn, Wiers, & Jansen, 2011) and the biased automatic processes can be re-trained through Cognitive Bias Modification (CBM), showing a decrease of symptoms in long time heavy users (for review, see Wiers, Gladwin, Hofmann, Salemink, & Ridderinkhof, 2013).

While effective in long time heavy users and clinical patients (Bickel et al., 2011; Eberl et al., 2013; Field et al., 2007), these training paradigms are often long and tedious and could be viewed as boring (Beard, Weisberg, & Primack, 2012). As youth often perceive more positive than negative effects of their alcohol use (National Institute on Alcohol Abuse and Alcoholism, 2005), any motivation to train they may have may decline during the training, which can lead to smaller training effects and drop out before completion. Adding elements of fun to the training paradigms may therefore be the key to increase the chances of success in this population. In this study, we looked at three elements that may increase the user experience of the training, and hence increase motivation to train. The first element we studied was *gamification* of the training task (cf. Gladwin, Figner, Crone, & Wiers, 2011). Indeed, Dovis, van der Oord, Wiers, and Prins (2012) showed that the inclusion of game elements in a working memory task increased motivation to train in children with ADHD. Building on this first element, the second element concerned the inclusion of a *social* game-aspect to the training (cf. Adams, 2013). In the first pilot study we focused on a social and non-social gamified version of an Alcohol/No-Go training, aimed at altering positive associations with alcohol in memory (Houben, Havermans, Nederkoorn, & Jansen, 2012).

The third element examined is the way the training's *ease of use* affects user experience. Usually CBM training paradigms are done on a computer, either in a lab setting or at home, through a web-based interface (van Deursen, Salemink, Smit, Kramer, & Wiers, 2013). The necessity of computer access may, however, still inhibit people from training as frequently as they can. Two studies (Dennis & O'Toole, 2014; Enock, Stefan, & McNally, 2014) showed that mobile application of CBM can be effective in changing attentional processes, and a preliminary swipe version of an approach/avoidance paradigm has already been developed (Kraus & Hofmann (2014). As such, the second pilot study presents a mobile version of an alcohol approach bias re-training (Wiers, Eberl, Rinck, Becker, & Lindenmeyer, 2011).

Part I - The Game Study

Goal of the study was twofold: First, we looked at whether adding (social) gaming elements increased the user experience and motivation to train. Second, we assessed the training effectiveness in terms of change in alcohol related memory bias and actual alcohol use after the training.

Study design and procedure

In the Game Study we compared four versions of the Alcohol/No-Go training (van Deursen et al., 2013): the original training (GNG-T), a neutral placebo training (GNG-P), a gamified version (GNG-G), where several game elements were added to the original task, and a social version of the game (GNG-SG) placed in a social network context (www.facebook.com). At baseline, participants were randomly allocated to one of the four conditions. Gaming experience and experience with other alcohol studies, alcohol problems and use, and the Go/No-Go task were assessed, and the first training was started immediately after. The training consisted of three sessions, at least one day and at most a week apart. At the end of the third training session the motivation and user experience questionnaires and the Go/No-Go assessment were done. A week later, participants received a follow-up questionnaire (TLFB) about alcohol use by email.

Participants

A sample of regularly drinking undergraduate students ($N = 77$, $M_{age} = 22.7$, $SD = 3.1$, range 18 to 29 years, 50.6 percent male) was recruited to participate in the training study in exchange for study credits. Participants were given the option to either train in the lab ($n = 35$) or from home ($n = 42$). At baseline, participants were informed about the study's training goal and procedure (but without mention of the game aspect, as this could potentially have a negative influence on those who did not receive game-training) and provided informed consent to participation. The study was approved by the Ethics Committee of the University of Amsterdam (Protocol Number: 2014-DP-3628).

Materials

User experience and motivation - User experience was measured with a set of questions based on the player enjoyment evaluation model (Sweetser & Wyeth, 2005). This model is structured on the theory on flow and consists of the following eight elements: concentration, challenge, player skills, control, clear goals, feedback, immersion, and social interaction. The resulting questionnaire included 21 questions that were rated on a five-point Likert scale going from "strongly disagree" to "strongly agree". Thirteen questions regarding motivation to train and playing the game were based on the attention, relevance, confidence, and satisfaction subscales by Liu and Chu (2010) and rated on the same Likert scale.

Alcohol use and problems - A shortened version of the Timeline Followback questionnaire (TLFB; Sobell, Sobell, Klajner, Pavan, & Basian, 1986; Wiers, Hoogeveen, Sergeant, & Gunning, 1997) was used to measure alcohol consumption per day over the past week. An additional question assessed whether participants drank more than 4 (female participants) or 5 glasses (male participants) of alcohol during one occasion in the past week to determine the number of binge drinking occasions. Alcohol-related problems were measured with the Alcohol Use Disorder Identification Test (AUDIT; Saunders, Aasland, Babor, de la Fuente, & Grant, 1993). This questionnaire included 10 multiple-choice questions regarding alcohol consumption and alcohol related problems. The overall AUDIT score ranges between 0 and 40, with a score of 8 or higher indicating an increased risk of alcohol-related problems in normal samples and 11 or higher in student samples (Fleming, Barry, & MacDonald, 1991).

Alcohol related memory bias - The Alcohol Go/No-Go assessment task looked similar to the version described by van Deursen et al. (2013). Here, it consisted of 14 practice trials with neutral images and 80 assessment trials, each showing either an image of an alcoholic or non-alcoholic beverage. On each picture, a cue (the letter P or F) was shown in one of the corners of the image, indicating whether the participant had to press a key (the spacebar) or not. For example, a participant could be instructed to press the response key as quickly as possible whenever the letter P (i.e., the Go cue) was visible, but not to press whenever the letter F (i.e., the No-Go cue) was shown. This cue-response matching was counterbalanced across participants. Each picture was paired with a Go cue equally often as with a No-Go cue. The bias score is calculated as the average reaction time on the alcohol-Go trials minus the average reaction time on the non-alcohol-Go trials.

Intervention

The Alcohol No-Go training, which was the basis for all of the four training conditions, was visually similar to the assessment version, except for the pairing of the image content and cues. In the training conditions (GNG-T, -G and -SG), images of alcohol were always paired with the No-Go cue and non-alcohol images with the Go cue. In the placebo training (GNG-P) there was no relation between the image content and the cues, similar to the assessment version. Each training session consisted of 200 training trials, showing images of beverages, mixed with 20 filler trials, showing neutral objects. The fillers were included to slightly mask the contingency between image content and Go/No-

Go cue (cf. Houben, Schoenmakers, & Wiers, 2010; Schoenmakers et al., 2010). The interface of the regular Go/No-Go task is displayed in Figure 1.

The game and social game versions of the Alcohol/No-Go training were called the Cheese Ninja Game. The aim of this game version was to include mechanics, backstory and aesthetics game aspects, as suggested by Schell (2014), while attempting to preserve as many of the key features of the paradigm. The main character in the game was a ninja mouse that walked through a hallway, passing by posters of the same beverages as used in the regular versions of the training. In front of each poster, a cue was presented in the form of something good or bad for the mouse, such as some cheese or a cat, respectively. The mouse had the ability to drop these objects, and the goal of the game was to collect as many good objects as possible while ignoring harmful objects. Similar to the regular training condition, the Go cues were consistently paired with non-alcoholic beverages on the posters, and the No-Go cues with alcoholic beverages.

In the social game condition, the game was registered as an app on the social network site Facebook (www.facebook.com). Participants were provided with pre-made test accounts, personalized with their first name and a neutral but unique profile picture. After each level the player could choose to post his or her level score and achievement to his Facebook timeline, which could be viewed, liked and commented on by the other GNG-SG participants. The interface of the social game within the Facebook environment is shown in Figure 1.

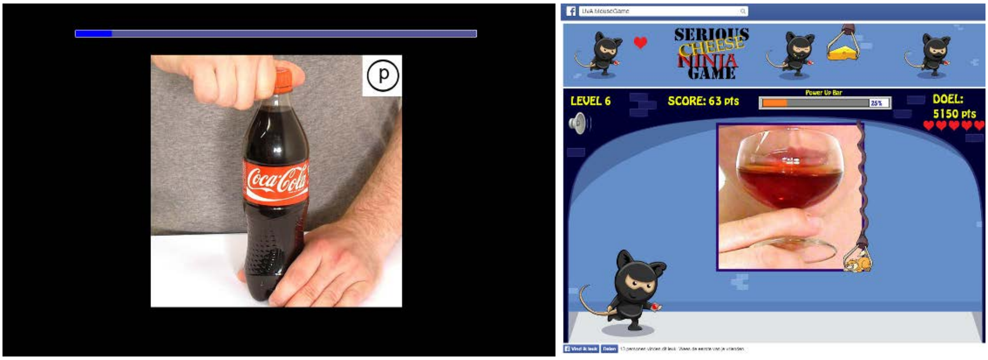

Fig. 1. The Alcohol/No-Go training: On the left is the original task; on the right is the social game version.

Game Study Results

Seven participants who missed the post-intervention assessment were excluded from the analyses. Next, all dependent variables were screened for univariate extreme outliers based on inspection of stem-and-leaf and box plots, which lead to the exclusion of one outlier in the Task Clarity factor of the user experience questionnaire.

Baseline characteristics

The average amount of alcohol consumed in the week before the pre-training assessment was 13.6 standard glasses (SD = 12.5). The mean AUDIT score was 10.2 (SD = 5.5), with 68.8 percent of the participants scoring ≥ 8 and 41.3 percent ≥ 11, indicating hazardous drinking in a large proportion of the sample (Fleming et al., 1991; Saunders et al., 1993). Nevertheless, and contrary to what was expected, they did not show a significant bias for alcohol, $t(76) = 1.053$, $p > 0.05$. Finally, participants differed with regard to their gaming habits. Participants in the social game condition appeared to play games slightly more often (see Table 1).

TABLE 1. BASELINE CHARACTERISTICS BY GROUP (GAME STUDY)

	Placebo (GNG-P)	*Original (GNG-T)*	*Game (GNG-G)*	*Social game (GNG-SG)*	*Total*	*p*
Total (male)	23 (9)	17 (9)	18 (8)	19 (13)	77 (39)	0.27
Age (years)	22.5 [3.1]	23.2 [3.3]	23.1 [3.3]	22.2 [3.1]	22.7 [3.1]	0.70
AUDIT	8.5 [5.3]	10.1 [5.0]	10.3 [4.5]	12.3 [6.5]	10.2 [5.5]	0.16
TLFB [a] (drinks)	10.7[10.4]	11.2 [9.0]	14.5 [7.9]	18.6 [19.0]	13.6[12.5]	0.19
GAME XP	3.4 [1.5]	3.3 [1.5]	3.3 [1.5]	2.2 [1.0]	3.1 [1.5]	0.03*
Alcohol Bias	1.2 [33.6]	-8.9[37.0]	-13.4[32.3]	3.0 [29.4]	-4.0[33.2]	0.37

Table shows means, with [standard deviations between brackets].

[a] One outlier removed.

* indicates $p < 0.05$.

AUDIT, Alcohol Use Disorder Identification Test; TLFB, Timeline Followback, shows the number of standardized drinks during the week prior to the pre-training assessment; GAME XP, frequency of playing games, where 1 = daily, 2 = weekly, 3 = monthly, 4 = yearly and 5 = never; Alcohol bias in milliseconds.

User experience and motivation to train

Exploratory factor analysis of the user experience questionnaire was carried out to test the factors composing the scale. A principal axis factor analysis was used with orthogonal Varimax rotation. Four factors with Eigenvalues > 1 (Kaiser, 1960) and a minimal 5 percent explained variance were identified. Table 2 shows these factors, which were labeled *Task Clarity*, *Ease of Use*, *Task Immersion* and *Task Demand,* with a total explained variance of 53.6 percent. Five items were discarded for not loading (minimal factor loading of 0.35, cf. Floyd & Widman, 1995) onto any factor. Similar analysis (without rotation) of the motivational questionnaire revealed only one factor explaining 52.94 percent of variance. This factor included all but two items, which did not load significantly (factor loading below 0.35). Cronbach's alpha was 0.89.

Using these factors, we examined whether adding (social) gaming elements increased the user experience and motivation to train. For this purpose we contrasted both game conditions with the non-game conditions, as well as the social game versus the regular game. ANOVA on the Task Clarity factor (after one outlier removed) revealed no significant difference between the groups, $F(3,64) = 1.978$, $p > 0.05$. A Kruskal-Wallis test on Ease of Use factor (due to violation of normality assumption) showed a significant group difference, $H(3) = 26.101$, $p < 0.001$. Follow-up Mann-Whitney pairwise tests showed that the game conditions together ($Mdn = 3.00$) were rated significantly less easy to use than the combined non-game conditions ($Mdn = 4.00$, $U = 319.5$, $z = -3.38$, $p = 0.001$, $r = -0.41$), but the game condition ($Mdn = 4.00$) was rated significantly easier to use than the social game condition ($Mdn = 2.33$; $U = 45.0$, $z = -3.44$, $p = 0.001$, $r = -0.59$). ANOVA on the Task Immersion factor also revealed a significant difference between groups, $F(3,65) = 4.520$, $p = 0.006$, with the social game ($M = 3.29$, $SD = 0.77$) being evaluated as more immersive than the normal game ($M = 2.34$, $SD = 0.62$), $t(65) = 3.646$, $p = 0.001$, $r = 0.41$. Finally, ANOVA on the Task Demand factor revealed a significant difference between groups, $F(3,65) = 5.154$, $p = 0.003$, where the game conditions together ($M = 3.13$, $SD = 0.77$) were rated as more demanding than the combined non-game conditions ($M = 2.47$, $SD = 0.64$), $t(65) = 3.769$, $p < 0.001$, $r = 0.42$. ANOVA on the motivational questionnaire revealed a

TABLE 2. ROTATED FACTOR SOLUTION FOR THE USER EXPERIENCE QUESTIONNAIRE (GAME STUDY): ITEM CONTENT, FACTOR LOADINGS (> .35) AND CRONBACH'S ALPHA FOR EACH FACTOR

Items	*Task clarity*	*Ease of use*	*Task immersion*	*Task demand*
- It was clear to me what was expected of me during the task	0.734*			
- It was clear to me how I performed during the exercise	0.705*			
- I thought all information that was shown on screen during the exercises was clear	0.567*	0.394		
- I had the idea that I kept improving at the exercises	0.555*			
- I thought that there were no redundant elements in the exercises	0.542*			
- I had the idea that I had complete control during the task		0.848*		
- The controls of the exercises were easy		0.751*		
- I thought the other visual elements on the screen were distracting		-0.730*		
- I forgot about the time while doing the exercises			0.753*	
- I was completely immersed in the task	0.383	-0.378	0.568*	
- While I was doing the exercises, I was unaware of my environment			0.702*	
- I got easily distracted from doing the exercises			-0.464*	
- The exercises were sufficiently challenging for me				0.421*
- The objects that I had to focus on were hard to discriminate		-0.397		0.607*
- The exercises required my full concentration				0.846*
- I didn't have to concentrate much to do the exercises	0.351			-0.450*
Cronbach's alpha	*0.78*	*0.82*	*0.73*	*0.63*

Items composing each factor are marked with an asterisk (*). Cross-loading items have been assigned to one factor based on the highest factor loading and content relevance. Items with a negative factor loading must be reverse scored.

significant difference between groups, $F(3,66) = 4.136$, $p = 0.010$, where the social game ($M = 3.23$, $SD = 0.51$) was rated more motivating than the normal game ($M = 2.64$, $SD = 0.71$), $t(66) = 2.806$, $p = 0.007$, $r = 0.33$.

Training effects

No bias change, $F(3,65) = 0.605$, $p > 0.05$, or change in drinking behavior, $F(3,65) = 1.111$, $p > 0.05$, were found after the training (see Table 3). We also examined the number of errors made on the Go/No-go assessment task. These were very low in all groups (mean error rates between 0.9 and 1.8 percent), indicating that in general participants were extremely accurate.

TABLE 3. TRAINING OUTCOMES (GAME STUDY)

	Placebo (GNG-P)	*Original (GNG-T)*	*Game (GNG-G)*	*Social game (GNG-SG)*	*p*
Alcohol bias change	-18.9 [47.3]	-20.2 [44.6]	-3.2 [29.2]	-6.6 [57.1]	0.614
TLFB change	2.3 [9.7]	-4.0 [7.8]	4.1 [20.5]	0.2 [10.2]	0.351

Table shows means, with [standard deviations between brackets].
TLFB, Timeline Followback, shows the number of standardized drinks during the week prior to the pre-training assessment.

Part II - The Mobile Study

Goal of this pilot study was to test reported user experience, motivational aspects and preliminary effectiveness of a mobile CBM intervention targeting automatic motivational approach tendencies towards alcohol, in comparison to the standard computerized version in a sample of young regular drinkers. This CBM intervention aims at training the participant to specifically avoid or approach a specific type of stimuli (Eberl et al., 2013; Wiers et al., 2011; Wiers, Rinck, Kordts, Houben, & Strack, 2010). Typically, it is a modified version of an assessment task, such as the Approach-Avoidance Task (AAT, Rinck & Becker, 2007; Wiers et al., 2010), with a built-in contingency that recasts it to re-training paradigm. However, performing such a task repeatedly on a computer may hinder motivation to train and training adherence (Beard et al., 2012), particularly with young adults (Gladwin et al., 2011). The deployment of CBM interventions on a mobile device could then maximize their effects and improve compliance.

User experience with the mobile condition was expected to be more positive than with the standard computerized condition, due to the perceived greater accessibility, convenience, and versatility of a mobile application (Hypothesis 1). Training technology was expected to predict number of completed training blocks after controlling for intrinsic motivation to train at baseline (Hypothesis

2). Alcohol use and approach bias were expected to decrease after the intervention in both conditions as a function of number of completed training blocks (Hypothesis 3).

Design and procedure

Participants were allocated to the mobile or computerized training intervention according to their smartphone operating system and gender. Participants with a non-Android™ (Google, Mountain View, CA) smartphone were assigned to the computerized training condition, whereas participants owning an Android smartphone were assigned to either condition stratified by gender.

At baseline, participants were screened for alcohol-related problems; whereas alcohol use and approach bias were assessed at baseline and after the intervention in the lab. Between the assessment sessions, participants could train for 14 days on their assigned device as much as they desired. Motivation to train was assessed at baseline and at the end of each training round. User experience was evaluated at post-intervention. Two weeks later, participants completed an online follow-up assessment of their alcohol use.

Participants

Recruited participants were 64 university students (M_{age} = 22.44, *SD* = 2.58; range 18 to 35 years, 60.94 percent female); 31 were assigned to the mobile training (all Android devices) and 32 to the computerized training (15.6 percent Android devices). One participant who did not fully complete the post-intervention session was excluded from the study. At baseline, participants were fully informed about the study goal and procedure and provided informed consent to participation. They were rewarded with two research credits or €20.00 upon completion of all assessment sessions. In addition, they received €0.80 for each completed training block. The study was approved by the Ethics Committee of the University of Amsterdam (Protocol Number: 2015-DP-4286).

Materials

Alcohol use and problems - Alcohol-related problems and total amount of standard units of alcohol consumption in the last two weeks was assessed with the AUDIT (Saunders et al., 1993) and TLFB (Sobell et al., 1986), respectively.

Approach bias - Besides a standard AAT, a second computerized reaction-time task was used to assess alcohol approach bias, namely the Stimulus Response Compatibility task (SRC, de Houwer, Crombez, Baeyens, & Hermans, 2001; Field, Kiernan, & Eastwood, 2008), in order to avoid practice effects due to the use of the modified AAT for training. Both tasks involved reacting as fast and as accurate as possible to alcoholic and non-alcoholic pictures by responding to the actual content of the picture (alcohol or soft drink) in the SRC and to an irrelevant feature of the picture (i.e., tilt direction) in the AAT.

In the SRC, participants were instructed to move a manikin away or towards a specific image by pressing two response keys (U or B) on the keyboard. The task was composed of a practice block (16 trials) and two test blocks (48 trials each) presenting 12 alcohol and 12 soft drink images repeated twice. In the 'alcohol approach' test block, participants had to move the manikin towards alcohol images and away from soft-drink images, whereas in the 'alcohol avoid' test block instructions were reversed. Block order was counterbalanced across participants. SRC score is computed by subtracting mean reaction time (RT) in the 'approach alcohol' block from mean RT in the 'avoid alcohol' block.

In the AAT, participants were instructed to push pictures tilted to the left away and pull pictures tilted to the right closer by pressing and holding two keys (up and down arrow keys) on the keyboard. Cue/response pairing was counterbalanced across participants. Upon response a zooming effect occurred, which increased picture size in the pulling closer response and decreased it in the pushing away, mimicking actual approach and avoidance (Wiers et al., 2009). The task was composed of a practice block (10 trials) with filler pictures (office supplies) and one test block, which consisted of 96 critical trials presenting six alcohol and six soft-drink stimuli repeated four times, mixed with 12 filler trials. AAT score is computed by subtracting the difference in mean RT between soft-drink/avoid and soft-drink/approach trials from the difference in mean RT between alcohol/avoid and alcohol/approach trials. All stimuli were taken from the Amsterdam Beverage Picture Set (ABPS, Pronk, van Deursen, Beraha, Larsen, & Wiers, 2015).

Motivation to train - At baseline, four questions asked about expectancies about the intervention and motivation to train. Each question was rated on a

5-point Likert scale from “totally disagree” to “totally agree”. Mean score across the four items was used for the analyses. As a proxy of motivation to continue training, participants indicated how much they enjoyed the training after each training round on a scale from 0 to 10.

User Experience - A questionnaire based on the User Experience Questionnaire (Rauschenberger, Schrepp, Olschner, Thomaschewski, & Cota, 2012) assessed user experience with both technologies by focusing on aspects such as ease of use, efficiency, stimulation and enjoyment of the training application. Questions were evaluated on the same 5-point Likert scale. Note that the user experience questionnaire used in this study differs from the one used in the Game Study. The reasons for this are twofold: First, there seems to be no golden standard in the literature when it comes to assess user’s experience, which prompted us to develop (and evaluate) a set of questions of our own, based on the relevant literature. Second, as the environments assessed in these studies were quite different (a (social) game setting versus a mobile environment), the questions were also different.

Intervention

The AAT was adapted for training by manipulating the stimulus-response contingency as to always avoid alcohol stimuli and approach soft-drink stimuli (Boffo, Pronk, Wiers, & Mannarini, 2015; Eberl et al., 2013; Wiers et al., 2011). Letters “P” and “F” were superimposed on the stimuli and used as response cues for push-away and pull-closer responses. Letter and response pairing was counterbalanced across participants. The training program was composed of 12 blocks of 72 trials presenting alcohol and soft-drink stimuli mixed with 8 filler trials. Six sets of six alcohol, soft drink and filler pictures were randomly presented throughout the 12 blocks. Stimuli were taken from the ABPS set (Pronk et al., 2015). After completion, the block sequence started over, allowing for unlimited training.

Both AAT training versions were visually similar to the assessment version of the task (see Figure 2). In the mobile version of the training, participants were instructed to respond with a swipe gesture on their touchscreen, with one hand holding the smartphone and the other swiping the stimuli away or towards themselves (Kraus & Hofmann, 2014). Similarly to the original AAT, a zooming effect was also implemented as well as an automated gliding motion animation to further emphasize an image moving towards or away from the user. Both training versions were programmed with Adobe ActionScript 3. Adobe Integrated Runtime (AIR) was then used to compile the application for Android by adjusting AIR settings to allow for maximum compatibility (from Android version 2.3 and above).

Fig. 2. The Alcohol/Avoid training: On the left is the PC version; on the right is the mobile version.

Mobile Study Results

Before running the analyses, all dependent variables were screened for univariate extreme outliers based on inspection of stem-and-leaf and box plots, which lead to the exclusion of two outliers in the number of completed training blocks, one outlier in the post-TLFB scores and one and six outliers in the baseline and post-test SRC scores, respectively.

Baseline characteristics

In the two weeks before the baseline assessment session, participants consumed on average 25.28 standard units of alcohol (SD = 20.10). Mean AUDIT scores was 11.27 (SD = 5.31) and 71.4 percent of participants reported an AUDIT score ≥ 8, and 58.7 percent ≥ 11, indicating hazardous drinking in a large proportion of the sample (Fleming et al., 1991; Saunders et al., 1993). Nonetheless, at baseline participants showed an alcohol avoidance bias score significantly different from 0 in the SRC task, $t(61) = -2.007$, $p = 0.049$, $r = 0.25$, and no alcohol bias in the AAT, $t(62) = 1.889$, $p = 0.064$. Group comparisons did not evidence any baseline difference for all relevant variables (see Table 4).

TABLE 4. BASELINE CHARACTERISTICS BY GROUP (MOBILE STUDY)

	Computer	*Mobile*	*Total*	*p*
Total (male)	32 (11)	31 (14)	63 (25)	0.38
Age (years)	22.0 [0.3]	23.0 [0.6]	22.5 [2.6]	0.12
AUDIT	11.3 [5.0]	11.3 [6.1]	11.3 [5.3]	0.99
TLFB	25.4 [19.1]	25.1 [21.4]	25.3 [20.1]	0.95
Motivation to train	2.9 [0.5]	3.1 [0.5]	9.0 [0.5]	0.09
Alcohol Approach Bias – SRC	-40.8 [76.2]	-1.8 [92.2]	-21.9 [85.9]	0.07
Alcohol Approach Bias – AAT	33.3 [74.8]	4.2 [83.7]	19.0 [80.0]	0.15

Table shows means, with [standard deviations between brackets].
AUDIT, Alcohol Use Disorder Identification Test; TLFB, Timeline Followback; SRC, Stimulus Response Compatibility task; AAT, Approach Avoidance task.

Hypotheses 1 and 2: User experience and motivation to train

Exploratory factor analysis was carried out on the user experience questions to identify the underlying factor structure. A principal axis factor extraction was used with Varimax orthogonal rotation. Four factors presenting Eigenvalues > 1.00 (Kaiser, 1960) and accounting for at least 5 percent of variance were identified. Six items did not significantly load onto any factor (factor loadings < 0.35, cf. Floyd & Widman, 1995) and were discarded. The final 4-factor solution explained 43.08 percent of the total variance. Table 5 shows the structure of the four factors, which were labeled *Ease of Use*, *Task Enjoyment*, *Player Involvement*, and *Task Compliance.*

A Mann-Whitney U test (due to violation of normality assumption) did not identify any significant difference between the two training conditions in Ease of Use and Task Compliance ($U = 439.5$ and $U = 382.5$, respectively, ps > 0.05). Player Enjoyment scores were also similar between technologies, $t(61) = -0.119$, $p > 0.05$; whereas Player Involvement showed a marginal difference,

Table 5. Rotated factor solution for the User Experience Questionnaire (Mobile Study): item content, factor loadings (> .35) and Cronbach's alpha for each factor

Items	*Ease of Use*	*Player Enjoyment*	*Player Involvement*	*Task Compliance*
- The task instructions were clear	0.708*			
- The task was easy to use	0.757*			
- Based on the instructions, the task functioned as I expected	0.738*			
- The task loaded quickly	0.754*			
- Swiping on the smartphone or pressing on the keyboard felt accurate	0.466*			
- I thought the task was fun to do		0.558*		0.431
- I quickly lost my interest in the task		-0.605*	0.505	
- I was very motivated to do the task		0.588*		
- I thought it was frustrating to do the task		-0.509*		
- Completing each task gave me a satisfied feeling		0.660*		
- I felt impatient while doing the task		-0.607*		
- I felt I had to concentrate hard during the task			-0.428*	
- I thought the task was very easy			0.371*	
- The task crashed often			-0.469*	
- I did the task while I was bored			0.598*	
- I didn't care whether I made a mistake				-0.608*
- I did the task because I felt the need to do so with regards to my drinking behavior				0.534*
- I thought the repeating of the swiping on the smartphone or pressing on the keyboard was exhausting				-0.488*
Cronbach's alpha	*0.80*	*0.77*	*0.51*	*0.57*

Items composing each factor are marked with an asterisk (*). Cross-loading items have been assigned to one factor based on the highest factor loading and content relevance. Items with a negative factor loading must be reverse scored.

$t(61) = -1.899$, $p = 0.06$, with a potentially greater involvement in the mobile group ($M = 11.23$, $SD = 2.12$) than the computerized group ($M = 10.13$, $SD = 2.46$). Mean scores of motivation to continue with training computed over all training rounds did not differ between the two conditions, $t(54) = 0.542$, $p > 0.05$.

A one-tailed Mann-Whitney U test showed that participants completed slightly more training blocks in the mobile training version (Mdn = 53, M = 72.38, SD = 65.73) than in the computerized one (Mdn = 17.00, M = 59.22, SD = 79.26), U = 348.00, z = -1.679, p = 0.046, r = -0.21. However, this effect disappeared after controlling for overall motivation to train at baseline with hierarchical regression analysis, R^2 = 0.091, F(1,59) = 5.816, p = 0.019; β = 0.293, p = 0.025.

Hypothesis 3: Training effects

Training effect on TLFB and SRC and AAT approach bias scores was examined with repeated-measures ANOVAs with number of completed training blocks as covariate and training condition as the between-subject factor.

The repeated-measures ANOVA on TLFB scores did not retrieve any main effect of time, F(2,88) = 1.804, $p > 0.05$, number of blocks, F(1,44) = 0.333, $p > 0.05$, or condition, F(1,44) = 0.030, $p > 0.05$. Post-hoc pairwise comparisons with Bonferroni correction found a significant difference between TLFB score at baseline (M = 25.54, SD = 18.57) and at follow-up (M = 20.12, SD = 16.92), t(49) = 3.321, p = 0.002. The repeated-measures ANOVA on SRC and AAT approach bias scores did not retrieve any main effect of time, F(1,53) = 0.365 and F(1,55) = 0.895, respectively, $ps > 0.05$, number of blocks, F(1,53) = 2.351 and F(1,55) = 1.849, respectively, $ps > 0.05$, or condition, F(1,53) = 3.647 and F(1,53) = 3.539, respectively, ps = 0.06. Post-hoc pairwise comparisons with Bonferroni correction found a marginal difference between AAT scores at baseline (M = 22.76, SD = 80.18) and at post-test (M = -0.18, SD = 42.40), p = 0.06, indicating a slight decrease in alcohol-approach bias over time.

General discussion

This article presented two pilot studies on motivational elements in CBM training. The Game Study showed that, while no effect of training was found on the bias score or drinking behavior, the games were seen as more demanding and harder to use than the regular training methods. Presenting the game in a social context also made it more immersive and more motivating compared to the stand-alone game. The Mobile Study showed a statistical trend indicating that players appeared to be more involved in the mobile group. The other measures of user experience, however, revealed no significant difference between both technologies. Initial analyses indicated the mobile trainers did complete more training blocks, but motivation to train prior to the training appeared of much greater influence, cancelling out the significance of this effect. Finally, while several marginal effects of training were found, these did not substantially differ between the conditions.

As these studies were pilot studies, several limitations have to be taken into account. For example, the Mobile Study did not include a placebo control condition, so the effects of the training cannot be solely attributed to the AAT. Also, the monetary reward appeared to be a very good motivator to some participants (with many earning well over €100). It remains unclear how many blocks would have been done if no money was involved. Another limitation is the samples used in these studies. While they did drink substantially, no significant bias scores were detected at baseline (in the Game Study) and only the SRC, but not the AAT, was significant in the Mobile Study. While this does not make it impossible to find effects of training, higher bias scores do allow for larger reductions as well (Eberl et al., 2013). Finally, some of the measures used were not optimal. For example, the self-developed user experience questionnaire used in the Mobile Study revealed two scales that were relatively unreliable (Cronbach's alpha < 0.60).

A more general point is that whenever significant adjustments are made to CBM paradigms (i.e., by transforming them into mobile or gamified applications), some elements will inevitably end up differently from their original, evidence-based counterparts. In these studies, for example, the introduction of movement of the stimuli, different controls and the added element of fun or mobility may make the training more challenging, but may also lead to more variability in reaction times, or in the extreme case may even render the training ineffective (Boendermaker, Prins, & Wiers, 2015; **Chapter 2** in this thesis). The difficulty of predicting which elements may have this effect, however, stresses the importance of performing these pilots.

Given that the main reason for training in the Mobile Study seemed to be having a strong motivation to train to begin with, perhaps future research should invest more in using elements of fun for maintaining any pre-existing motivation to train, rather than attempting to elicit it. If the next step in CBM research is to go mobile (e.g., Enock, Hofmann, & McNally, 2014; Kerst & Waters, 2014), in order to accommodate users' needs and technology preferences and take advantage of greater accessibility of mobile apps, it is necessary that training on a computer or on a mobile device does not result in different effects. While based on these results one could conclude that there is no real added value of going mobile with, or to introduce swiping gestures to, the alcohol avoidance training, it should be noted that the mobile training also did not perform worse than the regular training in this sample. As such, additional, placebo-controlled research to validate the efficacy of mobile training seems warranted, preferably in clinical samples. Moreover, as the (social) game elements did seem to increase motivation, future research could investigate their combination in mobile game versions of CBM.

Conclusions

The current studies indicate that adding (social) game elements can increase fun and motivation to train using CBM. Introducing ease of use elements, such as mobility, may increase motivation to train, but this seems less influential than initial motivation to train. Nevertheless, the mobile version did not underperform, which opens up new avenues for CBM training among younger participants. More research is needed to increase power and determine whether clinical effects can also be attained in the target populations.

Acknowledgements

The authors wish to thank the students involved in the data collection for these studies, Antonios Georgiadis, Ans de Nijs, and Wendelien Steltenpöhl (Part I - The Game Study), and Elisa van der Plas, Nienke van Bueren, Roberto Floris, and Vita Jongen (Part II - the Mobile Study).

Chapter 6
Attentional Bias Modification with Serious Game Elements: Evaluating the Shots Game

This chapter is based on:
Boendermaker, W. J., Sanchez Maceiras, S., Boffo, M., & Wiers, R W. (2016). Attentional bias modification with serious game elements: Evaluating the Shots Game. *Journal of Medical Internet Research: Serious Games, 4*(2): e20. DOI: 10.2196/games.6464

Abstract

Background: Young adults often experiment with heavy use of alcohol, which poses severe health risks and increases the chance of developing addiction problems. In clinical patients, cognitive re-training of automatic appetitive processes, such as selective attention towards alcohol (known as “Cognitive Bias Modification of Attention”, or CBM-A), has been shown to be a promising add-on to treatment, helping to prevent relapse. **Objective:** To prevent escalation of regular use into problematic use in youth, motivation appears to play a pivotal role. As CBM-A is often viewed as long and boring, this paper presents this training in the form of a serious game, as a novel approach aimed at enhancing motivation to train. **Methods:** 96 heavy drinking undergraduate students carried out either a regular CBM-A training, a gamified version (called ‘Shots’), or a placebo training version over four training sessions. Measures of motivation to change their behavior, motivation to train, drinking behavior and attentional bias for alcohol were included before and after training. **Results:** Alcohol attentional bias was reduced after training only in the regular training condition. Self-reported drinking behavior was not affected, but motivation to train decreased in all conditions, suggesting that the motivational features of the Shots Game were not enough to fully counteract the tiresome nature of the training. Moreover, some of the motivational aspects decreased slightly more in the Game condition, which may indicate potential detrimental effects of disappointing gamification. **Conclusions:** Gamification is not without its risks. When the motivational value of a serious game is less than expected by the adolescent, effects detrimental to their motivation may occur. We therefore advise caution when using gamification, as well as underscore the importance of careful scientific evaluation.

Introduction

Heavy alcohol use during adolescence and early adulthood has been related to health problems and academic underperformance (Singleton, 2007) and is an important predictor of addictive behaviors later in life (Kuntsche et al., 2013). Dual-process models of addiction (Deutsch & Strack, 2006; Wiers et al., 2007) suggest that prolonged use of alcohol, especially when initiated during adolescence (Gladwin, Figner, Crone, & Wiers, 2011), can lead to the development of strong automatically triggered reactions towards alcohol, which in turn facilitate the development of addictive behaviors. This is visible in heavy alcohol users' tendency to approach (Wiers, Rinck, Dictus, & van den Wildenberg, 2009) and selectively attend to (Field et al., 2007) alcohol-related stimuli more quickly, compared to non-alcohol related stimuli. Opposite to these strengthened automatic reactions are reflective cognitive processes, including control abilities (e.g., working memory, Grenard et al., 2008, and inhibition, Houben & Wiers, 2009) that can be too weak or too late to moderate the automatically triggered reactions (Gladwin et al., 2011). The resulting imbalance between automatically triggered appetitive processes and reflective control processes may contribute to escalation in problem drinking.

Research has shown that both in long-time heavy users and clinical patients, both types of processes can successfully be (re)trained, resulting in less craving and lower relapse rates (Bickel, Christensen, & Marsch, 2011; Eberl et al., 2013; Field et al., 2007; Houben, Wiers, & Jansen, 2011; Wiers, Gladwin, Hofmann, Salemink, & Ridderinkhof, 2013). Despite these promising results, application of cognitive training in younger populations has proven more difficult (Lindgren et al., 2015), for a number of reasons. First, youngsters tend to perceive stronger positive than negative effects of their alcohol use (National Institute on Alcohol Abuse and Alcoholism, 2005), perhaps because positive effects of alcohol tend to occur sooner than the negative effects Goldman (1999), making those positive associations stronger. This typically results in lower motivation to change their (drinking) behavior, compared with patient populations. Second, the fact that most training paradigms are long and often viewed as tedious and boring (Beard, Weisberg, & Primack, 2012) adds to the problem. To improve motivation to train, one potential solution could be to make the training methods more fun to do by adding game elements into the training paradigm. For example, Dovis, van der Oord, Wiers, and Prins (2012) offered children with Attentional Deficit/Hyperactivity Disorder (ADHD) a computer game version of a working memory task and observed that they normalized their persistence of performance to the level of children without ADHD. Dennis and O'Toole (2014) used a mobile game based on attentional bias modification training as an intervention for stress and anxiety in highly trait-anxious university students and showed a significant reduction in threat bias.

In the current paper, we apply similar gamification techniques to a typical CBM-A training task aimed at training attention away from pictures of alcohol-containing beverages: the Visual Probe Task (VPT; MacLeod, Rutherford, Campbell, Ebsworthy, & Holker, 2002; Schoenmakers et al., 2010). In the VPT participants are shown pairs of pictures, one of a relevant stimulus (e.g., a picture of an alcohol-containing beverage); the other a visually similar, neutral stimulus (e.g., a picture of a non-alcohol containing beverage). Next, a probe (e.g., an arrow pointing up or down) appears at the location of one of the stimuli, and the instruction is to quickly identify the probe (e.g., respond to the direction of the arrow). The contingency between the location of the probe and the stimulus it replaces can be manipulated. To assess attentional bias, the probe appears equally often at the location of both stimulus types; to train attention away from a certain set of stimuli (e.g., alcohol), the probe appears more often at the location of the other set of stimuli (e.g., non-alcohol). Schoenmakers and colleagues (2010) showed that CBM-A can indeed increase the ability to disengage from alcohol-related cues and found that alcohol dependent patients who had received the CBM-A training took significantly longer to relapse after training than patients who had not received CBM-A.

It should be noted that while adding game elements to a cognitive training task may help to increase participants' motivation to train, they usually also influence the specific task features and parameters and inevitably change to some degree the evidence-based nature of the task (Boendermaker, Prins, & Wiers, 2015; **Chapter 2** in this thesis). As such, we will compare the new gamified VPT training (VPT-G) both to a regular VPT training (VPT-R), to evaluate the added motivational effects of the game elements, as well as to a placebo version of the regular VPT training (VPT-P) to establish whether it has a training effect similar to the VPT-R. This results in the following hypotheses. First, we expect a significant reduction in attentional bias towards alcohol stimuli in both the VPT-R and VPT-G condition, compared to the

VPT-P condition (H1). This will be measured using both an assessment version of the VPT and another task that also measures attentional bias but is procedurally different: the Visual Search Task (VST; Dandenau & Baldwin, 2004). Next, we expect the same pattern of results between conditions with regard to decline of actual drinking behavior (H2). Finally, we expect to see that motivation to train is positively affected by the training in the VPT-G condition, but not in the VPT-R and VPT-P conditions (H3), while motivation to change is expected to remain unaffected, as it is not explicitly targeted by this training (H4).

Methods

Design and Procedure

The training consisted of four sessions, at least one day apart, over the course of two weeks. The first and last training sessions were combined with the assessment tasks in our lab; the two remaining training sessions were done at home. During the first session, participants gave digital informed consent and were randomly assigned to the VPT-P ($n = 33$), VPT-R ($n = 30$) or the VPT-G ($n = 33$) condition. They continued with digital versions of the Alcohol Use Disorders Identification Test (AUDIT), a short Readiness to Change Questionnaire (RCQ), the Timeline Followback questionnaire (TLFB), the Alcohol Use Questionnaire (AUQ), and a Motivation to Train Questionnaire (MTQ). After the questionnaires, they completed the VPT and VST baseline assessment and finished the session with the first VPT training. The following second and third session solely consisted of the VPT training task. The last session started with the fourth VPT training, after which they performed the VPT and VST post-training assessments, and the AUQ, TLFB, RCQ and MTQ questionnaires, supplemented by a brief set of questions about the training itself (EVAL). To evaluate drinking behavior after the training, a follow-up TLFB was filled in via email two weeks after session 4.

Participants

A sample of undergraduate students ($N = 96$, $M_{age} = 21.2$, $SD = 1.8$, range 18 to 28 years, 70.8 percent female) was recruited through the university lab's website, based on their drinking behavior (≥ 5 standard glasses of alcohol on average per week for males; ≥ 4 for females). Participants received study credits or €30 for taking part in the experiment. The study was approved by the Ethics Committee of the University of Amsterdam (Protocol Number: 2015-DP-4215).

Materials

Alcohol use and problems were measured with three questionnaires: A Timeline Followback procedure (TLFB; Sobell, Sobell, Klajner, Pavan, & Basian, 1986; Wiers, Hoogeveen, Sergeant, & Gunning, 1997) was used to measure alcohol consumption per day over the past week and also included a question about the number of binge drinking occasions during the past 30 days (>5 standard alcohol consumptions during one occasion for male participants; >4 for females). An adapted version of the AUQ (Mehrabian & Russell, 1978) was used to assess drink-specific alcohol consumption over the past six months. For analyses, Mehrabian and Russell's (1978) equation 2 was used to calculate the Habitual Alcohol Consumption (HAC), including those items regarding consumption of beer, wine and liquor, as well as our added items concerning alcohol pops. Alcohol-related problems were measured with the AUDIT (Saunders, Aasland, Babor, de la Fuente, & Grant, 1993), which includes 10 multiple-choice questions regarding alcohol consumption and alcohol related problems. The overall AUDIT score ranges between 0 and 40, with ⩾8 indicating an increased risk of alcohol-related problems in normal samples and ⩾11 in student samples (Fleming, Barry, MacDonald, 1991).

Motivation to train was assessed using a self-developed four-item questionnaire, each rated on a five-point Likert scale ranging from "strongly disagree" to "strongly agree". *Motivation to change* was assessed using a shortened version of the RCQ (Rollnick, Heather, Gold, Hall, 1992), consisting of three multiple-choice items. The EVAL questions concerned how they rated the training overall. See Supplementary Table 1 for an overview of translated questions.

Attentional bias was measured using assessment versions of the VST and VPT paradigms. In the VST (Dandenau & Baldwin, 2004; de Voogd, Wiers, Prins, & Salemink, 2014), participants were shown a grid of four by four pictures of beverages, where only one was of a different type: one alcohol-

SUPPLEMENTARY TABLE 1. MOTIVATIONAL QUESTIONS

Item	Dutch (original)	English (translation)
MTQ 1	Ik ben gemotiveerd om aan deze training mee te gaan doen.	I am motivated to participate in this training.
MTQ 2	Ik geloof dat deze training mij kan helpen om meer controle te krijgen over mijn drinkgedrag.	I believe that this training can help me to get more control over my drinking behavior.
MTQ 3	Ik wil graag mijn best doen op de training.	I want to do my best on the training.
MTQ 4	Ik vind de training zonde van mijn tijd.	I think the training is a waste of my time.
RCQSA 1	Denk je dat je meer drinkt dan goed voor je is?	Do you think that you drink more than is good for you?
RCQSA 2	Ben je van plan minder te gaan drinken?	Are you planning to drink less?
RCQSA 3	Ben je in de afgelopen drie maanden veranderd wat betreft je drinken?	Did you change your drinking behavior during the past 3 months?
EVAL 1	Ik denk dat deze training mensen in het algemeen kan helpen om meer controle te krijgen over hun drankgebruik.	I think this training can help people in general to get more control over their alcohol use.
EVAL 2	Ik vond het leuk om deel te nemen aan dit onderzoek.	I liked participating in this study.
EVAL 3	Ik zou deze training aanraden aan een vriend/vriendin met een soortgelijk drinkgedrag.	I would recommend this training to a friend who has a similar drinking behavior.
EVAL 4	Ik zou graag nog meer trainingssessies willen doen.	I would like to do more training sessions.
EVAL 5	Ik vond 4 trainingsessies te veel.	I thought 4 training sessions were too many.
EVAL 6	De vormgeving van de training motiveerde me om echt mijn best te doen.	The environment of the training motivated me to really do my best.

containing among 15 non-alcohol containing beverages, or vice versa. The instruction was to find and click the deviant type of beverage as quickly and accurately as possible. To focus visual attention to the center of the grid, each trial started with a fixation cross in the center of the screen the participants had to mouse over in order to start the trial. When an incorrect response was given, feedback was given and the trial had to be redone. The task consisted of six blocks of 18 trials, using active (person drinking) or passive (bottle/glass only) pictures of alcohol-containing and non-alcohol containing beverages, or neutral pictures of 5- or 7-petaled flowers, following the schedule in Table 1. The order of the blocks containing beverages was counterbalanced over participants. A progress bar indicated the number of trials left in each block. The attentional bias scores for active and passive stimuli were computed by subtracting the respective average reaction times (RTs) for selecting alcohol-containing beverages from the average RTs for selecting non-alcohol containing beverages. Given that faster RTs on alcohol trials suggest an attentional bias towards alcohol, a positive bias score thus indicated a bias towards alcohol.

Table 1. VST Block Distribution

Block	Target picture (1)	Non-target pictures (15)
1	Active alcohol	Active non-alcohol
2	Passive alcohol	Passive non-alcohol
3 / 4	5-petaled flower	7-petaled flowers
5	Active non-alcohol	Active alcohol
6	Passive non-alcohol	Passive alcohol

In the VPT, participants were shown pairs of alcohol and non-alcohol containing beverages, followed after 500 milliseconds by a small arrow probe in the location of one of the pictures, pointing up- or downwards. The instruction was to press the keyboard's arrow key corresponding to the arrow's direction as quickly and accurately as possible. The task consisted of 168 critical trials with pairs of beverages and 32 filler trials with pairs of neutral objects (office supplies), presented in random order over three blocks: a starting block of 10 neutral practice trials, then two test blocks of 100 trials (84 critical trials). In one of the two test blocks, the pictures disappeared as the arrow became visible ('GO' block); in the other block the pictures remained visible as the arrow probe was superimposed ('STAY' block). The STAY trials were included as they might better detect difficult disengagement from alcohol cues (Sportel, de Hullu, de Jong, & Nauta, 2013; Staugaard, 2009), while the standard go trials may be a better measure of rapid allocation and maintenance of attention on alcohol cues. The order of these blocks was counterbalanced over participants. Location of alcohol picture (left/right) and arrow (at alcohol or non-alcohol stimulus) were fully counterbalanced. The filler trials were included to slightly mask the contingency between arrow placement and the content of the pictures, as well as maintain participants' attention on the task and avoid anticipatory responses. When an incorrect response was given, the trial had to be redone after feedback was given, and the arrow direction was reapplied randomly. A progress bar indicated the number of trials left in each block. All picture pairs were matched by size and colors (see Figure 1). All stimuli originated from the Amsterdam Beverage Picture Set (ABPS; Pronk, van Deursen, Beraha, Larsen, & Wiers, 2015). Attentional bias scores for

STAY and GO trials were computed by respectively subtracting the average RT for trials with the arrow at the alcohol location from the average RT where the arrow was at the non-alcohol location. Given that faster responses on alcohol trials suggest an attentional bias towards alcohol, a positive bias score thus indicated a bias towards alcohol.

Fig. 1. The regular VPT (left) and the Shots Game implementation (right). The arrow size was matched between the tasks, but is enhanced slightly in the left pane (VPT) for visibility in print.

Training

The VPT training tasks were identical to the assessment version, except that after the practice block, there was only one training block of 156 critical and 16 filler trials, where STAY and GO trials were presented randomly. Additionally, in the VPT-R and VPT-G the arrow always appeared at the location of the non-alcohol containing beverage, thus training attention away from the alcohol-containing stimuli. In the VPT-G condition, participants trained using the Shots Game (van Schie & Boendermaker, 2014). The Shots Game was functionally identical to the VPT-R training, while looking like a slot machine game with two spinning wheels (see Figure 1; note that although the Shots Game looks like a slot machine, it has no gambling elements to it). The game elements used here constitute an *integrated gamification* of the VPT paradigm as defined by the CBM gamification model by Boendermaker and colleagues (2015; **Chapter 2** in this thesis). It mainly uses a coin-based reward system (cf. Step 1) and nicer looking graphics, animations and sound effects (cf. Step 3). The participant is rewarded for correct and fast responses (using time bonuses and special bonus trials), requiring a coin in order to spin the wheels (i.e., start a new trial) and eventually the possibility of reaching a new level (a new look for the machine). The game used picture stimuli similar to those in the ABPS, but slightly edited to fit the graphical style of the game.

Results

Two participants dropped out of the study after the first training session and were excluded from the analyses of training efficacy. Six additional participants failed to do the follow-up assessment and were excluded from the TLFB training effect analysis. Furthermore, baseline data from two participants on the RCQ question 3, and one participant on the VPT, were missing due to technical problems and therefore excluded from the relevant analyses.

Sample characteristics

At baseline, participants had consumed an average of 15.09 standard units of alcohol (SD = 11.46) during the previous 7 days, and binged on average on 6.65 occasions (SD = 3.48) during the previous 30 days. The AUQ average HAC score was 230.91 (SD = 17.17) and the mean AUDIT score was 13.58 (SD = 4.78), with 93.7 percent scoring $\geqslant 8$ and 71.8 percent $\geqslant 11$, indicating hazardous drinking in a large proportion of the sample (Fleming et al., 1991; Saunders et al., 1993). See Table 2 for an overview of baseline characteristics.

Training effects

All dependent variables were screened for univariate extreme outliers based on inspection of stem-and-leaf and box plots. When they were present, or one of the General Linear Model (GLM) assumptions were violated, a nonparametric method for factorial repeated-measures analysis of variance was used: The Aligned Rank Transform (ART) ANOVA (Wobbrock, Findlater, Gergle, & Higgins, 2011), in which data is aligned and then ranked as a preprocessing step, before applying GLM procedures (these results are marked with *).

Table 2. Baseline characteristics by group

	Placebo (VPT-P)	*Regular (VPT-R)*	*Game (VPT-G)*	*Total*	*P*
Total (male)	33 (10)	30 (8)	33 (10)	96 (28)	0.936
Age (years)	21.4 [2.1]	21.3 [1.4]	21.0 [2.0]	21.2 [1.8]	0.743
AUDIT	13.7 [5.5]	13.5 [3.9]	13.6 [4.9]	13.6 [4.8]	0.986
TLFB (drinks/7 days)	15.7 [14.2]	13.4 [7.6]	16.0 [11.4]	15.1 [11.5]	0.634
TLFB (binges/30 days)	6.9 [4.3]	6.3 [2.8]	6.7 [3.3]	6.7 [3.5]	0.809
AUQ (HAC)	229.7[17.8]	233.9 [20.5]	229.2 [12.4]	230.9 [17.2]	0.516
RCQ-1	2.2 [0.8]	2.3 [0.7]	2.5 [0.9]	2.4 [0.8]	0.454
RCQ-2	3.8 [1.9]	4.0 [1.9]	3.7 [1.9]	3.8 [1.9]	0.682
RCQ-3	1.8 [0.7]	2.0 [0.6]	1.9 [0.7]	1.9 [0.7]	0.499
MTQ	15.8 [2.4]	16.7 [1.8]	16.4 [2.0]	16.3 [2.1]	0.249
VPT-GO Alcohol Bias	-3.7 [20.8]	10.2 [25.1]	0.5 [22.9]	2.1 [23.4]	0.053
VPT-STAY Alcohol Bias	-0.2 [37.2]	1.9 [26.1]	3.2 [36.0]	1.6 [33.3]	0.920
VST-ACT Alcohol Bias	64.8[511.8]	250.9[596.4]	18.9[495.8]	107.2[538.0]	0.200
VST-PAS Alcohol Bias	116.8[515.2]	118.3[536.1]	232.7[594.4]	157.1[547.0]	0.624

Table shows means, with [standard deviations between brackets].
AUDIT, Alcohol Use Disorder Identification Test; TLFB, Timeline Followback; AUQ, Alcohol Use Questionnaire; HAC, Habitual Alcohol Consumption; RCQ-1/2/3, Readiness to Change Questionnaire, item 1/2/3; MTQ, Motivation to Train Questionnaire; VPT-GO/STAY, Visual Probe Task trials where the stimulus picture disappeared or remained visible when the probe appeared, respectively (in milliseconds); VST-ACT/PAS, Visual Search Task with active or passive beverage-related stimuli, respectively (in milliseconds).

6

Attentional bias change (H1) - There was a significant reduction of alcohol attentional bias over time on the VPT-GO trials, $F^*(1,90) = 9.407$, $p = 0.003$, $\eta_p^2 = 0.095$, as well as a significant interaction with condition, $F^*(2,90) = 8.685$, $p < 0.001$, $\eta_p^2 = 0.162$. Tukey-adjusted contrasts indicated this was due to a significant decrease of bias in the VPT-R condition over time, $t(90) = 3.094$, $p = 0.031$, $r = 0.310$, and also confirmed the result presented in Table 2 that the VPT-GO bias score at baseline was significantly higher in the VPT-R condition, compared to the VPT-P condition, $t(179.05) = 3.055$, $p = 0.031$, $r = 0.223$. The VPT-STAY trials also showed a significant reduction of bias over time, $F^*(1,90) = 10.894$, $p = 0.001$, $\eta_p^2 = 0.108$, without an interaction with condition. In contrast, no significant changes over time were found on the VST. A significant overall difference was found between the conditions on the VST-ACT trials: $F(2,91) = 5.480$, $p = 0.006$, $\eta_p^2 = 0.107$, but post-hoc analyses revealed no significant contrasts.

Behavioral change (H2) - There was no significant reduction in TLFB scores for both binges and total use (*P*s > 0.05).

Motivation to train (H3) - The MTQ demonstrated sufficient internal consistency, Cronbach's alpha = 0.69. Exploratory principal axis factor analysis indicated a single factor. Therefore the sum score was analyzed. There was a significant decrease in motivation to train over time, $F^*(1,91) = 54.377$, $p < 0.001$ $\eta_p^2 = 0.374$, with no interaction with condition, indicating motivation decreased similarly in all conditions. Participants' responses on the EVAL questions only differed between conditions on the question whether they would like to do more training sessions (EVAL-4), $H(2) = 9.987$, $p = 0.007$. Contrasts indicated that the VPT-G conditions scored significantly lower than both the VPT-P condition, $U = 356.0$, $p = 0.011$, $r = -0.313$, and the VPT-R condition, $U = 273.0$, $p = 0.004$, $r = -0.364$.

Motivation to change (H4) - The RCQ showed an overall increase of the degree to which participants planned to drink less after the training (a *lower* score on RCQ-2), $F^*(1,91) = 5.863$, $p = 0.017$, $\eta_p^2 = 0.061$. However, there also was a significant interaction between time and condition, $F^*(2,91) = 3.865$, $p = 0.024$, $\eta_p^2 = 0.078$. Tukey-adjusted contrasts indicated a lower motivation to drink less over time for the VPT-G condition, $t(91) = -2.985$, $p = 0.041$, $r = 0.299$. The other RCQ items did not show significant effects. See Table 3 for an overview of estimated marginal and interaction means.

TABLE 3. TRAINING EFFECTS – ESTIMATED MEANS

	Placebo (VPT-P)	Regular (VPT-R)	Game (VPT-G)	Time
TLFB (drinks/7 days)[a]				
Baseline	124.4 [14.4]	157.5 [14.5]	137.7 [13.7]	134.3 [8.2]
post-training	130.7 [14.4]	129.2 [14.5]	120.3 [13.7]	133.2 [8.2]
2 week follow-up	127.9 [14.4]	127.1 [14.5]	137.8 [13.7]	130.0 [8.2]
Condition	133.9 [11.5]	125.0 [11.6]	138.7 [11.2]	
TLFB (binges/30 days)[a]				
Baseline	128.9 [14.4]	139.3 [14.6]	140.5 [13.9]	136.4 [8.2]
post-training	122.2 [14.4]	138.2 [14.6]	133.4 [13.9]	135.6 [8.2]
2 week follow-up	129.1 [14.4]	121.3 [14.6]	139.5 [13.9]	125.5 [8.2]
Condition	134.5 [12.7]	122.9 [12.7]	140.1 [12.3]	
MTQ[a]				
baseline	92.9 [9.7]	94.8 [10.1]	100.7 [9.7]	113.6 [5.3]
post-training	96.0 [9.7]	90.7 [10.1]	91.9 [9.7]	75.4 [5.3]
Condition	85.0 [8.5]	102.9 [8.75]	95.6 [8.5]	
RCQ-1[a]				
Baseline	106.3 [9.5]	76.0 [9.9]	91.6 [9.5]	99.0 [5.5]
post-training	110.4 [9.5]	89.3 [9.9]	93.5 [9.5]	90.0 [5.5]
Condition	86.5 [8.4]	99.9 [8.7]	97.2 [8.4]	
RCQ-2[a]				
Baseline	108.2 [9.5]	76.1 [9.8]	86.5 [9.5]	100.4 [5.6]
post-training	100.8 [9.5]	85.4 [9.8]	110.1 [9.5]	88.6 [5.6]
Condition	84.3 [8.4]	101.4 [8.7]	97.7 [8.4]	
RCQ-3[a]				
Baseline	91.6 [9.2]	69.8 [9.6]	88.6 [9.3]	97.5 [5.4]
post-training	116.9 [9.2]	83.4 [9.6]	104.7 [9.3]	87.5 [5.4]
Condition	84.1 [7.4]	103.1 [7.6]	90.2 [7.5]	

Table shows means, with [standard errors between brackets].

[a] Using Aligned Rank Transform procedure

[b] Using regular ANOVA procedure

TLFB, Timeline Followback, shows the number of standardized drinks during the week prior to the pre-training assessment; MTQ, Motivation to Train Questionnaire; RCQ-1/2/3, Readiness to Change Questionnaire, questions 1/2/3, respectively.

6

TABLE 3. TRAINING EFFECTS – ESTIMATED MEANS (CONTINUED)

	Placebo (VPT-P)	Regular (VPT-R)	Game (VPT-G)	Time
VPT-GO Alcohol Bias[a]				
Baseline	77.1 [9.2]	118.1 [9.7]	80.9 [9.3]	104.8 [5.5]
post-training	105.0 [9.2]	76.5 [9.7]	103.3 [9.3]	82.2 [5.5]
Condition	91.7 [7.0]	87.8 [7.2]	101.0 [7.0]	
VPT- STAY Alcohol Bias[a]				
Baseline	79.8 [9.5]	103.3 [10.0]	92.6 [9.6]	106.9 [5.5]
post-training	105.6 [9.5]	88.6 [10.0]	91.0 [9.6]	80.1 [5.5]
Condition	99.6 [6.5]	88.4 [6.7]	92.4 [6.6]	
VST-ACT Alcohol Bias[b]				
baseline	64.8 [93.5]	226.6 [101.5]	18.9 [93.5]	103.4 [55.6]
post-training	50.2 [104.7]	535.1 [113.7]	28.0 [104.7]	204.4 [62.2]
Condition	57.5 [79.7]	380.8 [86.5]	23.5 [79.7]	
VST-PAS Alcohol Bias[b]				
baseline	116.8 [95.8]	153.6 [104.0]	232.7 [95.8]	167.7 [56.9]
post-training	29.6 [96.8]	178.9 [105.1]	235.5 [96.8]	147.9 [57.5]
Condition	73.2 [75.7]	166.1 [82.2]	234.1 [75.7]	

Table shows means, with [standard errors between brackets].

[a] Using Aligned Rank Transform procedure

[b] Using regular ANOVA procedure

VPT-GO/STAY, Visual Probe Task trials where the stimulus picture disappeared or remained visible when the probe appeared, respectively (in milliseconds); VST-ACT/PAS, Visual Search Task with active or passive beverage-related stimuli, respectively (in milliseconds).

Discussion

This study aimed to decrease attentional bias towards alcohol and hazardous drinking behavior in young adults by using a CBM-A training with game elements. After training, there was an overall decline of attentional bias on the VPT task, but this effect was primarily driven by the regular VPT training condition, where a stronger bias was observed at baseline. The training effect did not generalize to the VST task, nor was there a decline of alcohol use after the training. Motivation to train decreased equally in all conditions, indicating that the training indeed became boring over time, but also that the motivational elements of the Shots Game could not sufficiently counteract this effect. Interestingly, motivation to change, with respect to planning to drink less in the future, increased in the regular and placebo training, but decreased in the game training condition. Moreover, participants in the game condition indicated a lower motivation to continue training, compared to the other conditions.

These findings regarding the motivational effects of the training may have important implications on the potential risks involved with using certain types of game elements. While the gamified training arguably looked fancier than the regular training, the game elements in this study merely consisted of upgraded visuals and a coin-based reward system. There was no storyline and only limited progression and personalization options were available in the game. As such, it is likely that these rather minimal game elements alone were insufficient to increase motivation to train in our student sample. Moreover, most participants will likely compare any gamified training to what they believe a game experience should look and feel like (Boendermaker et al., 2015; **Chapter 2** in this thesis). If the game training experience then disappoints, motivation could indeed take a dive, even going below the level observed in the regular training conditions, because expectations were higher to begin with. This was perhaps reflected in our finding that participants in the game condition, specifically, were less motivated to continue training, as well as start drinking less, after the training. Finally, the current results may also be related to the visual and auditory game elements used, which might have distracted participants from the training elements, rendering it less effective. Indeed, the standard training condition with no game elements did show a small change in attentional bias. Moreover, motivation to change increased only in the non-game conditions. It could be that the exposure to alcohol cues gave participants a push towards a readiness to change, but only when there were no distracting game elements surrounding those cues. Although this last point is speculative, it is clear that some game elements may be detrimental not only to motivation to train but also to the training mechanisms themselves.

Limitations

Some limitations apply to this study. First, despite hazardous drinking in a substantial part of the sample, the modest training effects in this sample may be partially due to a relatively small alcohol attentional bias at baseline, which is a known moderator of training effects (Boettcher et al., 2013; Eberl et al., 2013; Kuckert et al., 2014). Furthermore, this study included a total of 624 critical training trials divided over four sessions. Although this number is similar to that used in other research (e.g., Dennis & O'Toole, 2014), other ABM studies have used markedly larger numbers (e.g., Schoenmakers et al., 2010, where participants completed 2640 training trials over five sessions). Given the very likely dose-response relationship between use and effectiveness of cognitive training paradigms, the amount of training practice may have prevented the training from efficiently changing attentional bias. Finally, a recent meta-analysis (Mogoaşe, David, & Koster, 2014) concluded that online CBM-A studies usually show smaller effect sizes than lab-based studies. While the assessments took place in the lab, it is possible that the option to train at home had a negative effect on participants' motivation, for example by making participants take the training less seriously.

Conclusions

In sum, the novel game-like approach used in this study proved insufficient to motivate young adults to train, in comparison with a regular CBM-A training. In fact, some aspects of motivation appeared to deteriorate, rather than improve, suggesting that gamification can have drawbacks if not done optimally. It could be concluded from this study that a point-based reward system in combination with fancy graphics do not satisfy participants' expectations of what constitutes a game. Because one expects a game to be fun, this may have detrimental effects on motivation. Moreover, when those game elements distract participants from the training elements, they may actually impair performance. A second notion that can be taken from this study is that the observed attentional biases towards alcohol as measured with the VPT in this heavy drinking student sample were remarkably low. Whether this has implications for the presence of attentional bias in adolescent samples in general or merely pertains to the VPT paradigm as a valid assessment measure of attentional bias remains to be determined by future research. If nothing else, however, these results underscore the importance of careful scientific evaluation, before serious games are used as interventions.

Acknowledgements

The authors wish to thank the students involved in the data collection for these studies: Aida Alai, Guillaume Rensink, and especially Sylvia van Schie, who also provided the game's graphics.

Chapter 7
Summary, General Discussion, and Recommendations

Parts of this chapter are based on:
Boendermaker, W. J., Peeters, M., Prins, P. J. M., & Wiers, R. W. (2017). Using serious games to (re)train cognition in adolescents. In M. Ma, & A. Oikonomou (Eds.), *Serious Games and Edutainment Applications, Volume II*. UK: Springer-Verlag. ISBN: 978-3-319-51643-1

Introduction

In this final part of this thesis the results from the studies presented in this thesis will be summarized and discussed following the framework set forth in **Chapter 2**. Limitations of the research, as well as general implications and suggestions for future research will be discussed. Finally, a critical reflection is made on several of the more practical aspects of the projects, with lessons to be learned for future design, development and evaluation of serious games.

Summary of main findings

The general aim of the projects presented in this thesis was to investigate if and how serious gaming techniques could be used to motivate adolescents and young adults to do cognitive training that is aimed at decreasing their alcohol intake. First, we developed a framework on which to base these serious games. The framework distinguishes between a number of steps in the development of serious games, based on cognitive training (cognitive bias modification, CBM, in particular), starting with an original, evidence based training paradigm and adding game-elements in various degrees. Based on this framework, we developed and investigated a number of serious games, which were the basis of the other chapters in this thesis. There were three goals in each of these serious game training methods. The first goal was to restore balance to the cognitive system, by either strengthening the adolescents' cognitive control functions (e.g., working memory, inhibition), or modifying their automatic, appetitive processes (e.g., attention or approach tendencies towards alcohol) that typically develop through prolonged alcohol use. The second goal, as a result of the cognitive change, was to decrease adolescents' problematic alcohol use. The third and final goal was to evaluate whether the game variant of the training indeed increased motivation to train amongst the sample beyond the level of that in the regular, non-gamified training.

The first two empirical studies described in this thesis (**Chapters 3** and **4**) made use of a shell-type game called the CityBuilder Game, which relates to Step 4 in the model. **Chapter 3** showed that serious game elements can help to increase motivation to increase working memory capacity in high school adolescents. Although the training turned out to be successful at increasing working memory capacity, it did not do so beyond the level found in the control condition in which participants only trained with relatively easy exercises (sequences of three squares). Importantly, motivation to do the post-training working memory assessment was found to be significantly lower in the gamified WMC-training condition, compared to the non-game and control training conditions. This may be due to the sudden lack of motivating rewards during the post-training assessment after multiple sessions of game-training in this condition, specifically. It could therefore be that the training gain of participants in the game condition was underestimated. Interestingly, the self-reported, explicit measure of motivation to train did not show participants had a preference for the game version of the training, but participants in the gamified condition did train significantly longer. The game was thus able to increase motivation to train, but the effect did fade over time until it was no longer different from the non-game conditions. As there was no differential effect on the (hypothesized) mediator (WMC), the training did not lead to significantly lower drinking (cf. MacLeod & Clarke, 2015). This lack of effect may have also been due to the fact that drinking levels in this school-based sample were very low at baseline.

The second study with the CityBuilder Game (**Chapter 4**) was specifically aimed at identifying the specific role of point-rewards (cf. Step 1) and point-rewards with value in a shell game (cf. Step 4) as alcohol-specific inhibition reinforcers. Alcohol-specific inhibition was found to increase, and alcohol intake was found to be lower, in all conditions after the Go/No-Go training. However, the effect of the training on alcohol-specific inhibition was more pronounced when the training itself was also alcohol-specific. Rewarding of correct behavior during training with arbitrary points or with points that held value in the shell-game was found not to significantly reinforce this training effect beyond giving minimal feedback only.

The next two chapters (**Chapters 5** and **6**) focused on exploring integrating elements of fun into the training paradigms (cf. Steps 2 and 3) by adding a social gaming context, swiping gestures in a mobile environment and extensive visuals. The first study in **Chapter 5** describes the Cheese Ninja Game, where game elements were integrated into an evidence-based CBM paradigm. The game was also embedded within a social media environment (www.facebook.com) to measure the added effects of social feedback. While no cognitive training effects were found, adding (social) game elements did increase appreciation for the training as well as participants' motivation to train. The second study presented in this chapter concerned a mobile CBM training, which appeared to increase motivation to train in terms how often participants trained, but this effect disappeared after controlling for baseline motivation to train. Importantly, despite changing several key aspects of the normal training paradigms, both variants did not underperform compared to their regular training counterparts in this student sample. Finally, in **Chapter 6** the Shots Game was evaluated. This game, based on the visual probe attentional bias retraining paradigm, introduced mainly an elaborate point-rewards system (cf. Step 1), as well as some fancy graphics and sounds, integrated well into the task (cf. Step 2). However, contrary to the other games presented in this thesis, the Shots Game had no additional gameplay elements included. The results showed that the attentional bias for alcohol was reduced only by the regular visual probe training, and not by playing the game. Nevertheless, self-reported drinking behavior was not affected by any of the training variants. As expected, motivation to train was shown to decrease over time, but this happened in all conditions, suggesting that the motivational features of the Shots Game were not enough to counteract the tiresome nature of the training. Moreover, motivation to starting drinking less after the training actually decreased after playing the Game, which may indicate potential detrimental effects of a disappointing gamification. Thus we showed here, inadvertently, that gamification is not without its risks, which has important implications for future serious game development (discussed further below).

In sum, this thesis has provided a first, evidence-based framework for the development and evaluation of serious games based on cognitive training paradigms, accompanied by a number of studies investigating the various potential benefits of such modifications. There are, however, some limitations to this work, as well as a number of important implications and suggestions for future research, which will be discussed in the sections below.

General Discussion

First I will discuss the limitations of the current research and how they affect the implications on the effectiveness of serious games for cognitive training. Then I will briefly reflect upon the general project flow over the past five years, and I will elaborate on how the use of Adobe Flash has played a major role in the project, with some important technical lessons to be learned. Finally, I will look to the future and make several more general suggestions aimed at other researchers who want to design, develop or evaluate serious games based on cognitive training paradigms.

Critical Reflections I - On the effectiveness of gamified cognitive training

The field of serious games research is ever growing. When this project started in 2011, not much was yet known with regard to CBM gamification, specifically, but now that the project has been completed, much progress has been made. How do our results fit into that field, what lessons can be learned and what conclusions should (not!) be drawn? First of all, there are several limitations to the research we have done. For example, as the gamified training interventions developed in these studies were all novel to some degree, we mainly included student samples for evaluation. However, despite high levels of drinking in those samples, we often found limited levels of bias at baseline (e.g., **Chapter 5**). This could indicate that the measures we used to detect cognitive biases were not optimal (for example, the visual probe paradigm is known to have particularly poor psychometric properties; van Bockstaele, Salemink, Bögels, & Wiers, 2015). But it could also support the notion that a longer period of heavy drinking is necessary in order to develop the biases that characterize heavy drinking adult samples. There are some studies (e.g., Amir, Taylor, & Donohue, 2011; Salemink & Wiers, 2012) that found that a stronger bias before the training predicted stronger CBM effects in anxiety, alcohol use (Eberl et al., 2013) and smoking (Elfeddali, de Vries, Bolman, Pronk, & Wiers, 2016). Indeed, several studies have indicated that the chances of finding significant biases in relatively light drinkers are typically low (approach bias: Wiers et al., 2009; attentional bias: Field & Cox, 2008). As such, the fact that biases at baseline were sometimes low made the evaluation of the games somewhat harder in terms of establishing training efficacy. But evaluation in a school setting is also more practical in terms of accessibility of the large numbers of participants necessary for all the experimental conditions. Nevertheless, the inferences of the current research are limited in the sense that it remains unclear how adolescents with much higher levels of use and/or bias would fare on these games. A second limitation concerns the development of the games. As all games presented in this thesis were developed by our research group (more details on this are in the next section), their level of fanciness is of course also limited. That means that although some of the games did indeed turn out to motivate adolescents beyond the level of the regular, non-gamified training (e.g., **Chapters 3** and **5**), any conclusions drawn from these data should take into account that they were by no means as much fun as commercial games. Although it can be argued (see **Chapter 2**) that serious games may never reach that level of fun (if only because of the serious aim of changing problem behaviors), serious games more motivating than the ones used in these studies are certainly possible to develop. Such games could potentially increase the effects found in these studies, or be better able to detect those that were not (yet) found. What we can learn from these studies, however, is what aspects of gamification, and of cognitive training itself, work better than others. For example, when only points, fancy visuals, and sounds are used, but no storyline or character development (**Chapter 6**), the training may look nicer but it may certainly not make the training more fun to do. A third limitation also relates to the student samples we used, in that despite their heavy drinking habits, and sometimes reported significant problems with their alcohol use, *motivation to actually change* their behavior was typically low (e.g., **Chapter 6**). Although serious games based on cognitive training may be able to increase participants' *motivation to train*, it seems unlikely that behavioral change can be achieved without some level of explicit motivation to make those changes. As a case in point, Kerst and Waters (2014) also found that adults who smoke but were not motivated to quit smoking did show a bias change after attentional retraining but no behavioral changes. As such, lab- or school-based studies such as these, with participants who are not necessarily motivated to make real behavioral changes, should perhaps not be used as the core evidence as to whether CBM in general does (not) work. Although this issue could suggest that the evaluation and application of serious games should rather be done in a more seriously drinking population, it should be noted that it can also be unethical to experiment with novel techniques on severe alcohol users if some of them will have to be allocated to a placebo control condition. Therefore, the logical first step remains the evaluation of these games in a more accessible, less heavy drinking sample, such as students, to rule out any inadvertent iatrogenic effects. Fourth, any experimental evaluation of a serious game training, compared to a non-game variant, will run the risk of suffering from a methodological fallacy in that motivation may be lowered during the post-training assessment of the training effect (**Chapter 3**). A sudden lack of motivating rewards during the post-training assessment after multiple sessions of game-training may have negatively consequences on participants' motivation, potentially leading to an underestimation of the actual training effect. This is a troublesome issue that seems difficult to counter. One option could be to gamify the assessment tasks as well, but the distraction coming from game elements may be detrimental to the sensitivity of the assessment itself (a problem that training tasks may suffer from as well, but to a much lesser degree), and prohibits the inclusion of a non-game control condition. Another option could be to incorporate mini-assessments during the training sessions (cf. van Deursen, Salemink, Smit, Kramer, & Wiers, 2013), but this may also intensify the entire training program by prolonging its overall duration. Fifth, as we have seen in **Chapter 2**, it is important to consider the delicate nature of evidence-based cognitive training paradigms in order to be able to use them as a basis for developing serious games. Most paradigms are structured as repeated stimulus-response exercises and tend to be very sensitive to slight structural changes (e.g., changing the display duration of a cue from 500 to 2000 milliseconds may give very different results; Field, Mogg, Mann, Bennett, & Bradley, 2013). As such, there is always a risk involved in adding game elements to such paradigms, as these may render the task less effective. As it is difficult to be completely sure how every single game element affects the training's efficacy, rigorous evaluation is of the utmost importance, especially if clinical groups are a potential (future) target of the game training. Although a lack of training effects (beyond placebo, or beyond the regular, non-game training, e.g., **Chapter 5**) seems a disappointing finding, it also shows that the introduction of various game elements to these training paradigms did not have undesired effects (with the notable exception of the Shots Game in **Chapter 6**). Both findings, however, stress the importance of careful evaluation in terms of both motivational value and (transfer of) cognitive effects.

Critical Reflections II - On Practical Issues during the Project

An important aspect of the project that affected the final form of this thesis has to do with a number of practical issues. These issues affected the main part of the project (the shell-game RCTs) that eventually made it impossible to include them in this thesis. However, as they took up the bulk of the project, I want to briefly describe them here.

We decided early on in the project, based on several pilot studies and additional literature research, to exclude two paradigms from the shell-game based RCTs. The alcohol-specific task switch training was dropped due to a lack of a readily available, reliable, evidence based training paradigm; the working memory training was excluded because it turned out that ten sessions of training were not enough to see a training effect (see **Chapter 2**) meaning that this paradigm would not match well with the other paradigms in terms of training duration. The remaining training paradigms (attentional and approach bias, and inhibition) were tested in a number of two (alcohol x cannabis) by five (placebo, attentional bias, approach bias, inhibition, and combined training) by four (training sessions) factorial repeated measures study designs. These studies included 160 seventeen to twenty year old university students and 260 sixteen to nineteen year old students at a school for mid-level vocational education. Both of these large-scale studies were plagued by practical issues. The first one (at university) was well underway when a bug in the Flash player plugin (used for the game) was detected by one of the research assistants. This bug caused the game to randomly crash for some participants in a way that also made it impossible for the system to (systematically) detect the crash. After several weeks of rigorous attempts at finding the source of the bug (more on this in the next section), it was eventually fixed, but there remained two problems with the data set. First, despite additional interviews held after the training, it was unknown *exactly* which participants had encountered this problem at which points during the training. Second, it could not be ascertained what influence these crashed may have had on participants' motivation to train. Despite our asking about this in the interview, the data was deemed too unreliable to publish and was therefore excluded from the thesis. The second study (at the MBO school) did not have this issue, but suffered from another problem: despite elaborate meetings with the school's coordinator before the study, who repeatedly assured us everything was in order, the very day before the first assessments were to start it was communicated to us that the teachers 'suddenly' did not fully support cooperation with the study. It was decided nevertheless to go forward with the study, but in practice, this meant that some of the teachers flat-out denied our research assistants access to the classroom, openly criticizing the study and researchers, and some walked out of the classroom at random, leaving the assistants alone to manage groups of unmotivated, sometimes rebellious students. In short, despite our best efforts these aspects led to a huge percentage of drop-out, with only one student actually finishing all sessions, and less than half of the students showing up for the post-training assessment. This is not to say that only the school was to blame, but again, the entire data set was deemed insufficient to use for publication. The remaining data from the other participating schools eventually turned out to provide insufficient power to present in this thesis. This is also the main reason for why only alcohol-related data is presented here. In all of these studies, participants were given the option at the start of the training to either train in the context of gaining more control over their alcohol use, or their cannabis use. As is to be expected, the number of students who chose alcohol was much higher than for those who chose cannabis. As such, to give the thesis more focus it was decided to only include the alcohol data.

Critical Reflections III - On the Use of Adobe Flash

The serious games, non-game training tasks, and assessment tasks presented in this thesis were developed using Adobe Flash, which turned out to be a source of several problems. As such, I want to spend a few words to describe some of the technical details of the bugs we have encountered, so that others who have used, or are planning to use, Adobe Flash for research purposes in general can learn from these issues.

First of all, despite the unusually large number of issues with Flash in this project, software bugs affecting research outcomes are not as uncommon as one might think. For example, it was recently discovered that some of the most widely used software for fMRI data analysis contains erroneous code that tends to inflate false-positive rates (Eklund, Nichols, & Knutsson, 2016). As such it is of the utmost importance that researchers be open about these issues and share their findings with others.

When speaking about Adobe Flash, I effectively refer to its programming environment that is based on the ActionScript 3 language. Many members of the ADAPT research team at the University of Amsterdam (and beyond) have been using ActionScript-based tasks for online assessment, training and serious games since 2009. Our choice for the use of ActionScript was made after careful consideration of several programming languages (e.g., Java, JavaScript, C++). Flash provided the ability to present the training programs over the internet (which was a requirement for this project), offered many high-end features and had a very clear syntax. Flash also had its own environment (a browser plug-in), which allowed for more control over the user experience than, for instance, JavaScript, which was more browser-sensitive. This all changed a few years into the project when Adobe (Flash and ActionScript's sole developer) decided to re-focus, away from browser-based applications (for which the Flash player plug-in is used) and towards mobile applications (for which the, still very usable Adobe AIR, also based on ActionScript, is used). This effectively meant that more bugs appeared, and support for Flash on many devices (e.g., Apple's iPhone and iPad devices) and browsers was discontinued. Despite this decline of quality over time, the fact that all training and assessment tasks, and most importantly the games, were already developed in Flash (running well into the tens of thousands of lines of code) made the decision to switch to another platform difficult. With regard to the project presented in this thesis, changing platforms midway was simply impossible.

To give an idea of the specific problems we have encountered, in the next section I will briefly explain their technical aspects, as well as the scientific consequences they may have had. The first issue concerned 'PepperFlash': when Flash was run inside certain browsers (mainly Google Chrome, using a specific type of plug-in system called Pepper Plugin Application Programming Interface; PPAPI, instead of the older Netscape Plugin Application Programming Interface; NPAPI), tasks ran notably sluggish. For example, stimuli that were presented relatively briefly (200 ms) sometimes completely failed to appear on screen. While it remains unknown whether and how this affected the recorded reaction times, this clearly was an issue to be avoided. As such, after finding this problem we implemented a browser-, and Flash plug-in version check, so that certain browsers could be avoided. As this specific issue lies in the connection between the plug-in itself and the browser plug-in system, there really wasn't much we could do (on the coding side). This issue has been reported to Adobe but no solution has been offered to date.

The second issue concerned the aforementioned Flash player plug-in crash: the CityBuilder Game appeared to occasionally crash the browser plug-in, which also prevented the logging of the problem. It appeared that a certain piece of code of the game (actually, a single 1 that should have been a 0) sometimes caused a loop-problem that should be handled without problem by the Flash player plug-in, but in some situations it crashed the Flash player. While this was again an issue to be fixed by Adobe (no piece of code should ever lead to a crash), no solution has been offered to this date.

The third issue occurred in some (but not all) of our tasks (e.g., the IAT in **Chapter 4**) where an unusually large number of the recorded reaction times were in fact multiples of 10 (that is, they ended with a -0, e.g., 340, 560, 770; instead of e.g., 345, 561, 778). While the number of RTs ending with a -0 should be around ten percent of all RTs, in some data sets it was around 25 percent and sometimes even as high as 85 percent. Although we couldn't find a source of, or solution to this problem, upon closer inspection, we came to the following conclusions: First, the problem seemed to occur in different tasks, and different versions of tasks that used the exact same code structures when it came to anything related to the measurement of reaction times. Second, the problem also *didn't* occur in many other participants, and sometimes it occurred only in certain blocks of a task, within the same participant, while this was not the case with other participants. This strongly suggested that this issue did not concern a bug in the tasks' code, as that would very likely cause the problem to happen all the time, but rather in a sensitivity of the Flash Player plug-in to external sources of distress that occur only at certain times, in some participants. The fact that the applicable code has also been used in many other tasks that do not have this problem also supported this notion (but it cannot be fully ruled out at this point). We have looked at and ruled out all kinds of potentially related factors, such as when was the task generating the problematic data executed, on which PC, which task version was used, and so forth. However, so far nothing has explained the phenomenon. While the data suggested that the reaction times were (probably) rounded to the nearest ten, which would mean that the reported reaction times were slightly less accurate (to the level of $1/100^{th}$ of a second instead of $1/1000^{th}$ of a second), importantly, we ran a number of simulations on the data which seemed to suggest that this did not affect block mean scores very much. That is, we calculated block mean scores for participants with and without this problem (so, those with normal RTs and those with RTs ending in -0 only), and compared them to the same RTs which were rounded to the nearest 10, and subsequently added a random integer between -5 and +5, creating randomized 1000^{th} digits. The resulting block averages deviated from the

7

original block averages by less than 1 millisecond, suggesting that any derived bias scores would be affected only very slightly, and effects on the group level would be negligible. This simulation was done based on IAT blocks of 24 and 48 trials, with similar results. Nevertheless, to be on the safe side, these tasks were not included in the chapters presented in this thesis.

Finally, on a general note, a disturbing issue is the fact that while we currently know about these culprits, the question arises what issues we don't know about. While this logic holds true for any task, in any programming language, admittedly Flash has by now acquired a bad reputation. Nevertheless, until there is evidence of the contrary there should be no need to doubt that the many Flash tasks used that did not show one of the problems mentioned above provided unreliable data and as such I believe their data can be trusted just as well as data from tasks built in other languages.

In sum, although Adobe Flash at the time of this writing has obviously lost its appeal to game designers and researchers alike, the effects it has had on previously collected data remain. It should also be noted that Flash' counterpart, Adobe AIR, so far has not had these problems, which means that ActionScript 3 as a programming language, and effectively most of the code developed for the tasks and games can still be used. Just not inside a browser.

Recommendations for future research, development and evaluation of serious games based on cognitive training

Based on the research presented in this thesis, what can we recommend about the use of serious games for the purposes of cognitive training? In general, serious games can be a promising new way to reach at risk youth, through prevention as well as intervention, and cognitive training can be a firm scientific basis for the design of those serious games. As this field is relatively young, more research is needed to determine for whom these cognitive training programs work best, and which game elements should be used or avoided. For example, the different game types described have thus far not been compared directly to see which one works best. Similarly, certain game elements, such as loud sound effects, flashing visual distractors, and real-time scoring, can also distract the participant during their performance and lead to reduced task performance (e.g., Katz, Jaeggi, Buschkuehl, Stegman, & Shah, 2014). This underscores the importance of validation of the new gamified measure, with regard to the degree to which the game-elements add to the cognitive load during task performance. Other interesting questions concern, for example, whether the combination of explicit and implicit motivational techniques, targeting motivation to change, for example, through motivational interviewing or cognitive behavioral therapy, as well as motivation to train through the use of game elements works best. Perhaps a serious game can be designed that incorporates elements of implicit cognitive training as well as more explicit cognitive behavioral therapeutic elements. And if game elements can improve cognitive training efficacy, can they also improve the quality of cognitive assessment data (Hawkins, Rae, Nesbitt, & Brown, 2013)? The fact that despite a lack of an explicit preference, participants may still implicitly prefer a game version of a cognitive training task (e.g., **Chapter 3**) has consequences for the way we evaluate their motivation to train. It is therefore recommended that any evaluation of a new serious game training in terms of motivation includes not only explicit, but also implicit, behavioral measures of motivation, as they may paint a different, perhaps also more reliable picture.

While many of these questions are currently being studied (e.g., Lumsden, Edwards, Lawrence, Coyle, & Munafò, 2016), several critical notions also apply. For example, as a typical cognitive (re)training program can take up to 600-2000 trials over multiple sessions (e.g., van Deursen et al., 2013), serious games that incorporate such large numbers of trials will need to keep participants' motivation high over a longer period of time in order for them to be able to sustain a sufficient level of performance. Interestingly, the use of game elements may therefore also introduce a new risk: when the training is presented as a game, participants' expectations of the level of fun during training will be raised, as the word "game" undoubtedly creates certain expectations based on previous experience (**Chapter 2**). As such, they now expect to be entertained. If this does not happen, however, their disappointment may also be greater than when they did not have these expectations in the first place (as is presumably the case when participating in a regular, non-game training intervention). Because of this, it is important to use the word 'game' with caution when presenting a serious game to participants. It would also be interesting to study participant expectations about the (gamified) training, and their effects on motivation and treatment outcome.

Despite the evidence that game-elements can increase participants' motivation for doing cognitive training, the level of fun may never reach that of commercial games. However, the question is, can or should we expect them to be? Buday, Baranowski, and Thompson (2012) suggest that a direct comparison with commercial games should perhaps be avoided altogether. Given that even expensive commercial games sometimes fail to keep critical players interested for long, there is indeed a challenge for typically low-budget serious games to keep motivation reasonably high, while keeping expectations relatively low. As such, serious game research, as well as training outcomes, could benefit from keeping expectations modest.

Appendix A
References

Adams, E. (2013). *Fundamentals of game design. Third Edition.* San Francisco, CA: New Riders. ISBN: 978-0321929679

Akaike, H. (1973). Information theory and an extension of the maximum likelihood principle. In B. N. Petrov, & F. Csaki (Eds.), *Proceedings of the Second International Symposium on Information Theory* (pp. 267–281). Budapest, Hungary: Akademia Kiado.
DOI: 10.1007/978-1-4612-1694-0_15 (reprint)

Allom, V., Mullan, B., & Hagger, M. (2016). Does inhibitory control training improve health behaviour? A meta-analysis. *Health Psychology Review, 10*(2), 168–186. DOI: 10.1080/17437199.2015.1051078

Amir, N., Beard, C., Burns, M., & Bomyea, J. (2009). Attention modification program in individuals with generalized anxiety disorder. *Journal of Abnormal Psychology, 118*, 28–33. DOI: 10.1037/a0012589

Amir, N., Taylor, C. T., & Donohue, M. C. (2011). Predictors of response to an attention modification program in generalized social phobia. *Journal of Consulting and Clinical Psychology, 79*(4), 533–541.
DOI: 10.1037/a0023808.

Anguera, J. A., Bernard, J. A., Jaeggi, S.M., Buschkuehl, M., Benson, B. L., Jennett, S., & Seidler, R. D. (2012). The effects of working memory resource depletion and training on sensorimotor adaptation. *Behavioural Brain Research, 228*, 107–115. DOI: 10.1016/j.bbr.2011.11.040

Baddeley, A. (1992). Working memory. *Science, 255*, 556–559. DOI: 10.1126/science.1736359

Bar-Haim, Y., Lamy, D., Pergamin, L., Bakermans-Kranenburg, M. J., & van IJzendoorn, M. H. (2007). Threat-related attentional bias in anxious and non-anxious individuals: A meta-analytic study. *Psychological Bulletin, 133*, 1–24. DOI: 10.1037/0033-2909.133.1.1

Barnett, E., Sussman, S., Smith, C., Rohrbach, L., & Spruijt-Metz, D. (2012). Motivational Interviewing for adolescent substance use: A review of the literature. *Addictive Behaviors, 37*(12), 1325–1334. DOI: 10.1016/j.addbeh.2012.07.001

Baumeister, R. F. (1991). *Meanings of life.* New York: Guilford. ISBN: 978-0898627633

Beard, C., Weisberg, R. B., & Primack, J. (2012). Socially anxious primary care patients' attitudes toward cognitive bias modification (CBM): A qualitative study. *Behavioural and Cognitive Psychotherapy, 40*, 618–633. DOI: 10.1017/S1352465811000671

Bedwell, W. L., Pavlas, D., Heyne, K., Lazzara, E. H., & Salas, E. (2012). Toward a taxonomy linking game attributes to learning: An empirical study. *Simulation & Gaming, 43*(6), 729–760. DOI: 10.1177/1046878112439444

Bickel, W. K., Christensen, D. R., & Marsch, L. A. (2011). A review of computer-based interventions used in the assessment, treatment, and research of drug addiction. *Substance Use and Misuse, 46*, 4–9. DOI: 10.3109/10826084.2011.521066

Blakemore S. J., & Choudhury S. (2006). Development of the adolescent brain: Implications for executive function and social cognition. *Journal of Child Psychology and Psychiatry, 47*, 296–312. DOI: 10.1111/j.1469-7610.2006.01611.x

Boendermaker, W. J., Boffo, M., & Wiers, R. W. (2015). Exploring elements of fun to motivate youth to do cognitive bias modification. *Games for Health Journal, 4*(6), 434–443. DOI: 10.1089/g4h.2015.0053

Boendermaker, W. J., Prins, P. J. M., & Wiers, R. W. (2013). *Documentation of the CityBuilder Game. Theoretical background and parameters.* Amsterdam, the Netherlands: University of Amsterdam.

Boendermaker, W. J., Prins, P. J. M., & Wiers, R. W. (2015). Cognitive bias modification for adolescents with substance use problems – Can serious games help? *Journal of Behavior Therapy and Experimental Psychiatry, 49*, 13–20. DOI: 10.1016/j.jbtep.2015.03.008

Boendermaker, W. J., Sanchez Maceiras, S., Boffo, M., & Wiers, R. W. (2016). *Attentional bias modification with serious game elements: Evaluating the Shots Game.* Manuscript submitted for publication.

Boettcher, J., Leek, L., Matson, L., Holmes, E. A., Browning, M., MacLeod, C., Andersson, G., & Carlbring, P. (2013). Internet-based attention bias modification for social anxiety: A randomized controlled comparison of training towards negative and training towards positive cues. *PLoS One, 8*, e71760. DOI: 10.1371/journal.pone.0071760

Boffo, M., Pronk, T., Wiers, R. W., & Mannarini, S. (2015). Combining cognitive bias modification training with motivational support in alcohol dependent outpatients: Study protocol for a randomised controlled trial. *BioMed Central Trials, 16*, 63. DOI: 10.1186/s13063-015-0576-6

Bonomo, Y. A., Bowes, G., Coffey, C., Carlin, J. B., & Patton, G. C. (2004). Teenage drinking and the onset of alcohol dependence: A cohort study over seven years. *Addiction, 99*, 1520–1528.
DOI: 10.1111/j.1360-0443.2004.00846.x

Boot, W. R., Kramer, A. F., Simons, D. J., Fabiani, M., & Gratton, G. (2008). The effects of video game playing on attention, memory, and executive control. *Acta Psychologica, 129*(3), 387–98.
DOI: 10.1016/j.actpsy.2008.09.005

Buday, R., Baranowski, T., & Thompson, D. (2012). Fun and games and boredom. *Games for Health Journal, 1*(4), 257–261.
DOI: 10.1089/g4h.2012.0026

Carver, C. S., & White, T. L. (1994). Behavioral inhibition, behavioral activation and affective responses to impending reward and punishment: The BIS/BAS scales. *Journal of Personality and Social Psychology, 67*(2), 319–333. DOI: 10.1037/0022-3514.67.2.319

Clarke, P. J. F., Notebaert, L., & Macleod, C. (2014). Absence of evidence or evidence of absence: Reflecting on therapeutic implementations of attentional bias modification. *BMC Psychiatry, 14*, 8.
DOI: 10.1186/1471-244X-14-8

A

Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences* (2nd ed.). New Jersey: Lawrence Erlbaum Associates. ISBN: 978-0805802832

Cooper, M. L. (1994). Motivations for alcohol use among adolescents: Development and validation of a four-factor model. *Psychological Assessment, 6 (2)*, 117–128. DOI: 10.1037/1040-3590.6.2.117

Cousijn, J., Goudriaan, A. E., Ridderinkhof, K. R., van den Brink, W., Veltman, D. J., & Wiers, R. W. (2012). Approach-bias predicts development of cannabis problem severity in heavy cannabis users: Results from a prospective FMRI study. *PLoS One, 7*, e42394. DOI: 10.1371/journal.pone.0042394

Cousijn, J., Goudriaan, A.E., Ridderinkhof, K. R., van den Brink, W., Veltman, D., & Wiers, R. W. (2013). Neural responses associated with cue-reactivity in frequent cannabis users. *Addiction Biology, 18*(3), 570–580. DOI: 10.1111/j.1369-1600.2011.00417.x

Cousijn, J., Goudriaan, A. E., & Wiers, R. W. (2011). Reaching out towards cannabis: Approach-bias in heavy cannabis users predicts changes in cannabis use. *Addiction, 106*, 1667–1674. DOI: 10.1111/j.1360-0443.2011.03475.x

Cristea, I. A., Mogoașe, C., David, D., & Cuijpers, P. (2015). Practitioner review: Cognitive bias modification for mental health problems in children and adolescents: A meta-analysis. *Journal of Child Psychology and Psychiatry, 56*(7), 723–734. DOI: 10.1111/jcpp.12383

Crone, E. A., & Dahl, R. E. (2012). Understanding adolescence as a period of social-affective engagement and goal flexibility. *Nature Reviews Neuroscience, 13*(9), 636–650. DOI: 10.1038/nrn3313

Dandeneau, S. D., & Baldwin, M. W. (2004). The inhibition of socially rejecting information among people with high versus low self-esteem: The role of attentional bias and the effects of bias reduction training. *Journal of Social and Clinical Psychology, 23*, 584–602. DOI: 10.1521/jscp.23.4.584.40306.

Danielsson, A.-K., Wennberg, P., Hibell, B., & Romelsjö, A. (2011). Alcohol use, heavy episodic drinking and subsequent problems among adolescents in 23 European countries: Does the prevention paradox apply? *Addiction, 107*, 71–80. DOI: 10.1111/j.1360-0443.2011.03537.x de Boer, A., Peeters, M., & Koning, H. M. (2016). An experimental study of risk taking behavior among adolescents: A closer look at peer and sex influences. *Journal of Early Adolescence.* Advance online publication. DOI: 10.1177/0272431616648453

de Houwer, J., Crombez, G., Baeyens, F., & Hermans, D. (2001). On the generality of the affective Simon effect. *Cognition and Emotion, 15*, 189–206. DOI: 10.1080/02699930125883

de Voogd, E. L., Wiers, R. W., Prins, P. J. M., & Salemink, E. (2014). Visual search attentional bias modification reduced social phobia in adolescents. *Journal of Behavior Therapy and Experimental Psychiatry, 45*, 252–259. DOI: 10.1016/j.jbtep.2013.11.006

de Wit, H. (2009). Impulsivity as a determinant and consequence of drug use: A review of underlying processes. *Addiction Biology, 14*(1), 22–31. DOI: 10.1111/j.1369-1600.2008.00129.x

Deci, E. L., Koestner, R., & Ryan, R. M. (1999). A meta-analytic review of experiments examining the effects of extrinsic rewards on intrinsic motivation. *Psychological Bulletin, 125*(6), 627–668. DOI: 10.1037/0033-2909.125.6.627

Deci, E. L., & Ryan, R. M. (1985). *Intrinsic motivation and self-determination in human behavior.* New York: Plenum. ISBN: 978-1489922731

Dennis, T. A., & O'Toole, L. J. (2014). Mental health on the go: Effects of a gamified attention bias modification mobile application in trait anxious adults. *Clinical Psychological Science, 2*(5), 576–590. DOI: 10.1177/2167702614522228

Derryberry, D., & Reed, M. (2002). Anxiety-related attentional biases and their regulation by attentional control. *Journal of Abnormal Psychology, 111*, 225–236. DOI: 10.1037/0021-843X.111.2.225

Deutsch, R., & Strack, F. (2006). Reflective and impulsive determinants of addictive behaviors. In R. W. Wiers & A. W. Stacy (Eds.), *Handbook of implicit cognition and addiction* (pp. 45–57). Thousand Oaks, CA: Sage. ISBN: 978-1412909747

Diamond, A. (2013). Executive functions. *Annual Review of Psychology, 64*, 135–168. DOI: 10.1146/annurev-psych-113011-143750

Dickman, S. J. (1990). Functional and dysfunctional impulsivity: Personality and cognitive correlates. *Personality Processes and Individual Differences, 58*(1), 95–102. DOI: 10.1037/0022-3514.58.1.95

A

Dovis, S., van der Oord, S., Wiers, R. W., & Prins, P. J. M. (2012). Can motivation normalize working memory and task persistence in children with attention-deficit/hyperactivity disorder? The effects of money and computer-gaming. *Journal of Abnormal Child Psychology, 40*, 669–681. DOI: 10.1007/s10802-011-9601-8

Dovis, S., van der Oord, S., Wiers, R. W., & Prins, P. J. M. (2013). What part of working memory is not working in ADHD? Short-term memory, the central executive and effects of reinforcement. *Journal of abnormal child psychology, 41*(6), 901–917. DOI: 10.1007/s10802-013-9729-9

Duka, T., Gentry, J., Malcolm, R., Ripley, T. L., Borlikova, G., Stephens, D. N., Veatch, L. M., Becker, H. C., & Crews, F. T. (2004). Consequences of multiple withdrawals from alcohol. *Alcohol Clinical Experimental Research, 28*, 233–246. DOI: 10.1097/01.ALC.0000113780.41701.81

Eberl, C., Wiers, R. W., Pawelczack, S., Rinck, M., Becker, E. S., & Lindenmeyer, J. (2013). Approach bias modification in alcohol dependence: Do clinical effects replicate and for whom does it work best? *Developmental Cognitive Neuroscience, 4*, 38–51. DOI: 10.1016/j.dcn.2012.11.002

Eklund, A., Nichols, T. E., & Knutsson, H. (2016). Cluster failure: Why fMRI inferences for spatial extent have inflated false-positive rates. *Proceedings of the National Academy of Science, 113*(28), 7900–7905. DOI: 10.1073/pnas.1602413113

Elfeddali, I., de Vries, H., Bolman, C., Pronk, T., & Wiers, R.W. (in press). Preventing smoking relapse among smokers via a web-based multiple-session attentional bias modification training: a randomized controlled trial. *Health Psychology.*

Emmelkamp, P. M. G. (2012). Attention bias modification: The emperor's new suit? *BioMed Central Medicine, 10*, 63. DOI: 10.1186/1741-7015-10-63

Enock, P. M., Stefan, G. H., & McNally, R. J. (2014). Attention bias modification training via smartphone to reduce social anxiety: A randomized controlled multi-session experiment. *Cognitive Therapy and Research, 28*, 200–216. DOI: 10.1007/s10608-014-9606-z

Entertainment Software Association (2014). Essential facts about the computer and video game industry. Retrieved from http://www.theesa.com/wp-content/uploads/2014/10/ESA_EF_2014.pdf on April 2, 2016.

Fadardi, J. S., & Cox, W. M. (2009). Reversing the sequence: Reducing alcohol consumption by overcoming alcohol attentional bias. *Drug and Alcohol Dependence, 101*, 137–145. DOI: 10.1016/j.drugalcdep.2008.11.015

Fagerström, K.O. (1978). Measuring degree of physical dependency to tobacco smoking with reference to individualization of treatment. *Addictive Behaviors, 3*, 235–241. DOI: 10.1016/0306-4603(78)90024-2

Festinger, L. (1957). *A Theory of Cognitive Dissonance.* Stanford, CA: Stanford University Press. ISBN: 978-0804709118

Field, M., & Cox, W. M. (2008). Attentional bias in addictive behaviors: A review of its development, causes, and consequences. *Drug and Alcohol Dependence, 97*, 1–20. DOI: 10.1016/j.drugalcdep.2008.03.030

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*, 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*, 350–357. DOI: 10.1007/s00213-005-0202-5

Field, M., Eastwood, B., Bradley, B. P., & Mogg, K. (2006). Selective processing of cannabis cues in regular cannabis users. *Drug and Alcohol Dependence, 85*, 75–82. DOI: 10.1016/j.drugalcdep.2008.03.030

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*, 209–218. DOI: 10.1016/j.jbtep.2007.06.001

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

Fleming, M. F., Barry, K. L., & MacDonald, R. (1991). The alcohol use disorders identification test (AUDIT) in a college sample. *International Journal of the Addictions, 26*, 1173–1185. DOI: 10.3109/10826089109062153

Floyd, F. J., & Widman, K. F. (1995). Factor analysis in the development and refinement of clinical assessment instruments. *Psychological Assessment, 7*, 286–299. DOI: 10.1037/1040-3590.7.3.286

Forbes, E. E., & Dahl, R. E. (2010). Pubertal development and behavior: Hormonal activation of social and motivational tendencies. *Brain and Cognition, 72*(1), 66–72. DOI: 10.1016/j.bandc.2009.10.007

Foxcroft, D. R., & Tsertsvadze, A. (2012). Universal school-based prevention programs for alcohol misuse in young people. *Evidence-based Child Health: A Cochrane Review Journal, 7*, 450–575. DOI: 10.1002/ebch.1829

Franken, I. H. A., Muris, P., & Rassin, E. (2005). Psychometric properties of the Dutch BIS/BAS scales. *Journal of Psychopathology and Behavioral Assessment, 27*(1), 25–30. DOI: 10.1007/s10862-005-3262-2

Geurts, H. M., Verte, S., Oosterlaan, J., Roeyers, H., & Sergeant, J. A. (2004). How specific are executive functioning deficits in ADHD and autism? *Journal of Child Psychology and Psychiatry, 45*(4), 836–854. DOI: 10.1111/j.1469-7610.2004.00276.x

Gladwin, T. E., Figner, B., Crone, E. A., & Wiers, R. W. (2011). Addiction, adolescence, and the integration of control and motivation. *Developmental Cognitive Neuroscience, 1*(4), 364–376. DOI: 10.1016/j.dcn.2011.06.008

Gladwin, T. E., & Wiers, R. W. (2011). How do alcohol cues affect working memory? Persistent slowing due to alcohol-related distracters in an alcohol version of the Sternberg task. *Addiction Research & Theory 20*(4), 284–290. DOI: 10.3109/16066359.2011.614977.

Goldman, M. S. (1999). Risk for substance abuse: Memory as final common etiological pathway. *Psychological Science, 10*, 196–198. DOI: 10.1111/1467-9280.00133

Goldstein, R. Z., & Volkow, N. D. (2002). Drug addiction and its underlying neurobiological basis: Neuroimaging evidence for the involvement of the frontal cortex. *The American Journal of Psychiatry, 159*, 1642–1652. DOI: 10.1176/appi.ajp.159.10.1642

Granic, I., Lobel, A., & Engels, R. C. (2014). The benefits of playing video games. *American Psychologist, 69*(1), 66–78. DOI: 10.1037/a0034857

Grant, V. V., Stewart, S. H., O'Connor, R. M., Blackwell, E., & Conrod, P. J. (2007). Psychometric evaluation of the five-factor modified drinking motives questionnaire-revised in undergraduates. *Addictive Behaviors, 32*, 2611–2632. DOI: 10.1016/j.addbeh.2007.07.004

Green, C. S., & Bavelier, D. (2003). Action video game modifies visual selective attention. *Nature, 423*(6939), 534–537. DOI: 10.1038/nature01647

Greenwald, A. G., McGhee, D. E., & Schwartz, J. L. K. (1998). Measuring individual differences in implicit cognition: The implicit association test. *Journal of Personality and Social Psychology, 74,* 1464–1479. DOI: 10.1037/0022-3514.74.6.1464

Grenard, J. L., Ames, S. L., Wiers, R. W., Thush, C., Sussman, S., & Stacy, A. W. (2008). Working memory capacity moderates the predictive effects of drug-related associations on substance use. *Psychology of Addictive Behaviors, 22*, 426–432. DOI: 10.1037/0893-164X.22.3.426

Habgood, M. J., & Ainsworth, S. E. (2011). Motivating children to learn effectively: Exploring the value of intrinsic integration in educational games. *Journal of the Learning Sciences, 20*, 169–206. DOI: 10.1080/10508406.2010.508029

Hadwin, J. A., & Richards, H. J. (2016). Working memory training and CBT reduces anxiety symptoms and attentional biases to threat: A preliminary study. *Frontiers in Psychology, 7*, 47. DOI: 10.3389/fpsyg.2016.00047

Hawkins, G. E., Rae, B., Nesbitt, K. V., & Brown, S. D. (2013). Gamelike features might not improve data. *Behavior Research Methods, 45*(2), 301–318. DOI: 10.3758/s13428-012-0264-3.

Hibell, B., Guttormsson, U., Ahlström, S., Balakireva, O., Bjarnason, T., Kokkevi, A., & Kraus, L. (2012). *The 2011 ESPAD report: Substance use among students in 36 European countries.* The Swedish Council for Information on Alcohol and Other Drugs, Stockholm, Sweden.

Holmes, J., Gathercole, S. E., & Dunning, D. L. (2009). Adaptive training leads to sustained enhancement of poor working memory in children. *Developmental Science, 12*, F9–F15. DOI: 10.1111/j.1467-7687.2009.00848.x

Hosenbocus, S., & Chahal, R. (2012). A review of executive function deficits and pharmacological management in children and adolescents. *Journal of the Canadian Academy of Child and Adolescent Psychiatry, 21*(3), 223–229. PMID: 22876270

Houben, K., Havermans, R. C., Nederkoorn, C., & Jansen, A. (2012). Beer a no-go: Learning to stop responding to alcohol cues reduces alcohol intake via reduced affective associations rather than increased response inhibition. *Addiction, 107*, 1280–1287. DOI: 10.1111/j.1360-0443.2012.03827.x

Houben, K., Nederkoorn, C., Wiers, R. W., & Jansen, A. (2011). Resisting temptation: Decreasing alcohol-related affect and drinking behavior by training response inhibition. *Drug and Alcohol Dependence, 116*, 132–136. DOI: 10.1016/j.drugalcdep.2010.12.011

Houben, K., Schoenmakers, T. M., & Wiers, R. W. (2010). I didn't feel like drinking but I don't know why: The effects of evaluative conditioning on alcohol-related attitudes, craving and behavior. *Addictive Behaviors, 35*, 1161–1163. DOI: 10.1016/j.addbeh.2010.08.012

Houben, K., & Wiers, R. W. (2009). Response inhibition moderates the relationship between implicit associations and drinking behavior. *Alcoholism: Clinical and Experimental Research, 33*, 626–633. DOI: 10.1111/j.1530-0277.2008.00877.x

Houben, K., Wiers, R. W., & Jansen, A. (2011). Getting a grip on drinking behavior: Training working memory to reduce alcohol abuse. *Psychological Science, 22*, 968–975. DOI: 10.1177/0956797611412392

Jaeggi, S. M., Buschkuehl, M., Jonides, J., & Shah, P. (2011). Short- and long-term benefits of cognitive training. *Proceedings of the National Academy of Sciences, 108*, 10081–10086. DOI: 10.1073/pnas.1103228108

Jaeggi, S. M., Buschkuehl, M., Shah, P., & Jonides, J. (2014). The role of individual differences in cognitive training and transfer. *Memory & Cognition, 42*, 464–480. DOI: 10.3758/s13421-013-0364-z

Johnston, L. D., O'malley, P. M., Bachman, J. G., & Schulenberg, J. E. (2012). *Monitoring the Future: National Survey Results on Drug Use, 1975-2012. Volume I: Secondary School Students.* NIH Publication No. 10-7584. National Institute on Drug Abuse (NIDA).

Kaiser, H. F. (1960). The application of electronic computers to factor analysis. *Educational and Psychological Measurement, 20*, 141–151. DOI: 10.1177/001316446002000116

Kane, M. J., & Engle, R. W. (2002). The role of prefrontal cortex in working-memory capacity, executive attention, and general fluid intelligence: An individual-differences perspective. *Psychonomic Bulletin & Review, 9*(4), 637–671. DOI: 10.3758/BF03196323

Kato, P. M. (2010). Video games in health care: Closing the gap. *Review of General Psychology, 14*(2), 113–121. DOI: 10.1037/a0019441

Katz, B., Jaeggi, S., Buschkuehl, M., Stegman, A., & Shah, P. (2014). Differential effect of motivational features on training improvements in school-based cognitive training. *Frontiers in Human Neuroscience, 8,* 242. DOI: 10.3389/fnhum.2014.00242

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

Kharrazi, H., Lu, A. S., Gharghabi, F., & Coleman, W. (2012). A scoping review of health game research: Past, present, and future. *Games for Health Journal, 1*(2), 153–164. DOI: 10.1089/g4h.2012.0011

Klingberg, T. (2010). Training and plasticity of working memory. *Trends in Cognitive Sciences, 14*, 317–324. DOI: 10.1016/j.tics.2010.05.002

Klingberg, T., Fernell, E., Olesen, P. J., Johnson, M., Gustafsson, P., Dahlström, K., Gillberg, C. G., Forssberg, H., & Westerberg, H. (2005). Computerized training of working memory in children with ADHD: A randomized, controlled trial. *Journal of the American Academy of Child & Adolescent Psychiatry, 44*, 177–86. DOI: 10.1097/00004583-200502000-00010

Kopetz, C., & Orehek, E. (2015). When the end justifies the means: Self-defeating behaviors as "rational" and "successful" self-regulation. *Current Directions in Psychological Science, 24*(5), 386–391. DOI: 10.1177/0963721415589329

Koster, E. H. W., Fox, E., & MacLeod, C. (2009). Introduction to the special section on cognitive bias modification in emotional disorders. *Journal of Abnormal Psychology, 118*(1), 1–4. DOI: 10.1037/a0014379

Kramer, J. R., Chan, G., Hesselbrock, V. M., Kuperman, S., Bucholz, K. K., Edenberg, H. J., Schuckit, M. A., Nurnberger, J. I., Foroud, T., Dick, D. M., Bierut, L. J., & Porjesz, B. (2010). A principal components analysis of the abbreviated Desires for Alcohol Questionnaire (DAQ). *Journal of Studies on Alcohol and Drugs, 71,* 150–155. PMID: 20105425

Kraus, A. A., & Hofmann, W. (2014). Getting in touch with motivation: The Swipe Approach–Avoidance Procedure (SwAAP). In: Kraus, A. *Indirect procedures for the measurement of approach-avoidance motivation,* pp. 176–198 (Unpublished doctoral thesis). Aarhus University: Aarhus, Denmark.

Krueger, R. F., Caspi, A., Moffitt, T. E., White, J., & Stouthamer-Loeber, M. (1996). Delay of gratification, psychopathology, and personality: Is low self-control specific to externalizing problems? *Journal of Personality, 64*(1), 107–129. DOI: 10.1111/j.1467-6494.1996.tb00816.x

Kuckertz, J. M., & Amir, N. (2015). Attention bias modification for anxiety and phobias: Current status and future directions. *Current Psychiatry Reports, 17,* 9. DOI: 10.1007/s11920-014-0545-x

Kuckertz, J. M., Gildebrant, E., Liliequist, B., Karlström, P., Väppling, C., Bodlund, O., Stenlund, T., Hofmann, S. G., Andersson, G., Amir, N., & Carlbring, P. (2014). Moderation and mediation of the effect of attention training in social anxiety disorder. *Behaviour Research and Therapy, 53*, 30–40. DOI: 10.1016/j.brat.2013.12.003

Kuntsche, E., Rossow, I., Simons-Morton, B., ter Bogt, T., Kokkevi, A., & Godeau, E. (2013). Not early drinking but early drunkenness is a risk factor for problem behaviors among adolescents from 38 European and North American countries. *Alcoholism: Clinical and Experimental Research, 37*(2), 308–314. DOI: 10.1111/j.1530-0277.2012.01895.x

Larimer, M. E., & Cronce, J. M. (2007). Identification, prevention and treatment revisited: Individual-focused college drinking prevention strategies 1999–2006. *Addictive Behaviors, 32*, 2439–2468. DOI: 10.1016/j.addbeh.2007.05.006

Lindgren, K. P., Wiers, R. W., Teachman, B., Gasser, M. L., Westgate, E. C., Cousijn, J., Enkema, M., & Neighbors, C. (2015). Attempted training of alcohol approach and drinking identity associations in US undergraduate drinkers: Null results from two studies. *Plos One, 10*(8), e0134642. DOI: 10.1371/journal.pone.0134642

Liu, T.-Y., & Chu, Y.-L. (2010). Using ubiquitous games in an English listening and speaking course: Impact on learning outcomes and motivation. *Computers & Education, 55*, 630–643. DOI: 10.1016/j.compedu.2010.02.023

Logan, G. D., & Cowan, W. B. (1984). On the ability to inhibit thought and action: A theory of an act of control. *Psychological Review, 91,* 295–327. DOI: 10.1037/0033-295X.91.3.295

Love, A., James, D., & Willner, P. (1998). A comparison of two alcohol craving questionnaires. *Addiction, 93,* 1091–1102. DOI: 10.1046/j.1360-0443.1998.937109113.x

Lumsden, J., Edwards, E., Lawrence, N., Coyle, D., & Munafò, M. (2016). Gamification of cognitive assessment and cognitive training: A systematic review of applications and efficacy. *JMIR Serious Games, 4*(2), e11. DOI: 10.2196/games.5888

Lumsden, J., Skinner, A., Woods, A. T., Lawrence, N. S., & Munafò, M. (2016). The effects of gamelike features and test location on cognitive test performance and participant enjoyment. *PeerJ, 4*, e2184. DOI: 10.7717/peerj.2184

Luna, B., Garver, K. E., Urban, T. A., Lazar, N. A., & Sweeney, J. A. (2004). Maturation of cognitive processes from late childhood to adulthood. *Child Development, 75*(5), 1357–1372. DOI: 10.1111/j.1467-8624.2004.00745.x

MacLeod, C., & Clarke, P. J. F. (2015). The attentional bias modification approach to anxiety intervention. *Clinical Psychological Science, 3*(1), 58–78. DOI: 10.1177/2167702614560749

MacLeod, C., Mathews, A., & Tata, P. (1986). Attentional bias in emotional disorders. *Journal of Abnormal Psychology, 95*, 15–20. DOI: 10.1037/0021-843X.95.1.15

MacLeod, C., Rutherford, E., Campbell, L., Ebsworthy, G., & Holker, L. (2002). Selective attention and emotional vulnerability: Assessing the causal basis of their association through the experimental manipulation of attentional bias. *Journal of Abnormal Psychology, 111*(1), 107–123. DOI: 10.1037//0021-843X.111.1.107

Martinussen, R., Hayden, J., Hogg-Johnson, S., & Tannock, R. (2005). A meta-analysis of working memory impairments in children with attention-deficit/hyperactivity disorder. *Journal of the American Academy of Child and Adolescent Psychiatry, 44*(4), 377–384. DOI: 10.1097/01.chi.0000153228.72591.73

Maurage, P., Joassin, F., Speth, A, Modave, J., Philippot, P., & Campanella, S. (2012). Cerebral effects of binge drinking: Respective influences of global alcohol intake and consumption pattern. *Clinical Neurophysiology, 123*(5), 892–901. DOI: 10.1016/j.clinph.2011.09.018

Mehrabian, A., & Russell, J. A. (1978). A questionnaire measure of habitual alcohol use. *Psychological Reports, 43,* 803–806. DOI: 10.2466/pr0.1978.43.3.803

Merry, S. N., Stasiak, K., Shepherd, M., Frampton, C., Fleming, T., & Lucassen, M. F. (2012). The effectiveness of SPARX, a computerised self help intervention for adolescents seeking help for depression: Randomised controlled non-inferiority trial. *British Medical Journal, 344*, e2598. DOI: 10.1136/bmj.e2598

Miller, W. R., & Rollnick, S. (2002). *Motivational interviewing: Preparing people to change addictive behaviors (2nd ed.)*. New York: Guilford. ISBN: 978-1572305632

Mogoaşe, C., David, D., & Koster, E. H. W. (2014). Clinical efficacy of attentional bias modification procedures: An updated meta-analysis. *Journal of Clinical Psychology, 70*, 1133–1157. DOI: 10.1002/jclp.22081

Moran, T. P. (2016). Anxiety and working memory capacity: A meta-analysis and narrative review. *Psychological Bulletin, 142*(8), 831–864. DOI: 10.1037/bul0000051

Morrison, A. B., & Chein, J. M. (2011). Does working memory training work? The promise and challenges of enhancing cognition by training working memory. *Psychonomic Bulletin Review, 18*, 46–60. DOI: 10.3758/s13423-010-0034-0

National Institute on Alcohol Abuse and Alcoholism (NIAAA) (2005). The effects of alcohol on physiological processes and biological development. *Alcohol Research and Health, 28*, 125–131. PMID: 19093686

Nikolaou, K., Critchley, H., & Duka, T. (2013). Alcohol affects neuronal substrates of response inhibition but not of perceptual processing of stimuli signalling a stop response. *PLoS One, 8, e*76649. DOI: 10.1371/journal.pone.0076649

Noble, A., Best, D., Sidwell, C., & Strang, J. (2000). Is an arcade-style computer game an effective medium for providing drug education to schoolchildren? *Education for Health, 13*(3), 404–406. DOI: 10.1080/135762800750059525

Notebaert, L., Clarke, P. J., Grafton, B., & MacLeod, C. (2015). Validation of a novel attentional bias modification task: The future may be in the cards. *Behaviour Research and Therapy, 65*, 93–100. DOI: 10.1016/j.brat.2014.12.007

Patton, J. H., Stanford, M. S., & Barratt, E. S. (1995). Factor structure of the Barratt impulsiveness scale. *Journal of clinical psychology, 51,* 768–774. DOI: 10.1002/1097-4679(199511)51:6<768::AID-JCLP2270510607>3.0.CO;2-1

Paus, T., Keshavan, M., & Giedd, J. N. (2008). Why do many psychiatric disorders emerge during adolescence? *Nature Reviews Neuroscience, 9*(12), 947–957. DOI: 10.1038/nrn2513

Peckham, A. D., McHugh, R. K., & Otto, M. W. (2010). A meta-analysis of the magnitude of biased attention in depression. *Depression and Anxiety, 27*, 1135–1142. DOI: 10.1002/da.20755

Peeters, M., Janssen, T., Monshouwer, K., Boendermaker, W., Pronk, T., Wiers, R. W., & Vollebergh, W. (2015). Weaknesses in executive functioning predict the initiating of adolescents' alcohol use. *Developmental Cognitive Neuroscience, 16*, 139–146. DOI: 10.1016/j.dcn.2015.04.003

Peeters, M., Wiers, R. W., Monshouwer, K., van de Schoot, R., Janssen, T., & Vollebergh, W.A. (2012). Automatic processes in at-risk adolescents: The role of alcohol-approach tendencies and response inhibition in drinking behavior. *Addiction, 107*, 1939–1946. DOI: 10.1111/j.1360-0443.2012.03948.x

Petrides, M., & Milner, B. (1982). Deficits on subject-ordered tasks after frontal- and temporal-lobe lesions in man. *Neuropsychologia, 20*(3), 249–262. DOI: 10.1016/0028-3932(82)90100-2

Presley, C., Meilman, P., Cashin, J., & Lyerla, R. (1996). *Alcohol and Drugs on American College Campuses: Use, Consequences, and Perceptions of the Campus Environment, Volume III: 1991-1993.* Carbondale, IL: Core Institute, Southern Illinois University.

Prins, P. J. M., ten Brink, E., Dovis, S., Ponsioen, A., Geurts, H. M., de Vries, M., & van der Oord, S. (2013). "Braingame Brian": Toward an executive function training program with game elements for children with ADHD and cognitive control problems. *Games for Health Journal, 2*(1), 44–49. DOI: 10.1089/g4h.2013.0004

Prins, P. J. M., Dovis, S., Ponsioen, A., ten Brink, E., & van der Oord, S. (2011). Does computerized working memory training with game elements enhance motivation and training efficacy in children with ADHD? *Cyberpsychology, Behavior, and Social Networking, 14*(3), 115–122. DOI: 10.1089/cyber.2009.0206

Prokhorov, A. V., DeMoor, C., Pallonen, U. E., Hudmon, K. S., Koehly, L., & Hu, S. (2000). Validation of the modified Fagerstrom tolerance questionnaire with salivary cotinine among adolescents. *Addictive Behaviors, 25*(3), 429–433. DOI: 10.1016/S0306-4603(98)00132-4

Pronk, T., van Deursen, D., Beraha, E., Larsen, H., & Wiers, R. W. (2015). Validation of the Amsterdam beverage picture set: A controlled picture set for cognitive bias measurement and modification paradigms. *Alcoholism: Clinical and Experimental Research, 39*(10), 2047–2055. DOI: 10.1111/acer.12853

Psychology Software Tools, Inc. [E-Prime]. Retrieved from http://www.pstnet.com.

Rauschenberger, M., Schrepp, M., Olschner, S., Thomaschewski, J., & Cota, M.P. (2012). Measurement of user experience: A Spanish language version of the User Experience Questionnaire (UEQ). In A. Rocha, J. A. Calvo-Manzano, L. P. Reis, & M. P. Cota (Eds.), *Proceedings of the 7th Iberian Conference on Information Systems and Technologies (CISTI)* (pp. 471–476. Madrid, Spain. ISBN 978-9899624764

Raven, J., Raven, J. C., & Court, J. H. (2004). *Manual for Raven's Progressive Matrices and Vocabulary Scales. Section 3: The Standard Progressive Matrices.* San Antonio, TX: Harcourt Assessment.

Rinck, M., & Becker, E. S. (2007). Approach and avoidance in fear of spiders. *Journal of Behavior Therapy and Experimental Psychiatry, 38*(2), 105–120. DOI: 10.1016/j.jbtep.2006.10.001

Robinson, S., Goddard, L., Dritschel, B., Wisley, M., & Howlin, P. (2009). Executive functions in children with autism spectrum disorders. *Brain and Cognition, 71*(3), 362–368. DOI: 10.1016/j.bandc.2009.06.007

Rollnick, S., Heather, N., Gold, R., & Hall, W. (1992). Development of a short 'readiness to change' questionnaire for use in brief, opportunistic interventions among excessive drinkers. *British Journal of Addiction, 87*(5), 743–754. DOI: 10.1111/j.1360-0443.1992.tb02720.x

Ross, T. P., Hanouskova, E., Giarla, K., Calhoun, E., & Tucker, M. (2007). The reliability and validity of the self-ordered pointing task. *Archives of Clinical Neuropsychology, 22*(4), 449–458. DOI: 10.1016/j.acn.2007.01.023

Ryan, R. M., & Deci, E. L. (2000). Intrinsic and extrinsic motivations: Classic definitions and new directions. *Contemporary Educational Psychology, 25*, 54–67. DOI: 10.1006/ceps.1999.1020

Salemink, E., & Wiers, R.W. (2012). Adolescent threat-related interpretive bias and its modification: The moderating role of regulatory control. *Behaviour Research and Therapy, 50*(1), 40–46. DOI: 10.1016/j.brat.2011.10.006

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*, 791–804. DOI: 10.1111/j.1360-0443.1993.tb02093.x

Schell, J. (2014). *The Art of Game Design: A Book of Lenses. Second Edition.* Boca Raton, FL: CRC Press. ISBN: 978-1466598645

Schmidt, N. B., Richey, J. A., Buckner, J. D., & Timpano, K. R. (2009). Attention training for generalized social anxiety disorder. *Journal of Abnormal Psychology, 118*, 5–14. DOI: 10.1037/a0013643

Schoemaker, K., Mulder, H., Dekovic, M., & Matthys, W. (2013). Executive functions in pre-school children with externalizing behavior problems: A meta-analysis. *Journal of Abnormal Child Psychology, 41*, 457–471. DOI: 10.1007/s10802-012-9684-x

Schoenmakers, T., de Bruin, M., Lux, I. F., Goertz, A. G., van Kerkhof, D. H., & Wiers, R. W. (2010). Clinical effectiveness of attentional bias modification training in abstinent alcoholic patients. *Drug and Alcohol Dependence, 109*, 30–36. DOI: 10.1016/j.drugalcdep.2009.11.022

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*, 399–405. DOI: 10.1111/j.1360-0443.2006.01718.x

Schwarz, G. (1978). Estimating the dimension of a model. *Annals of Statistics, 6*(2), 461–464. Stable URL: http://www.jstor.org/stable/2958889

Seppa, K., Lepisto, J., & Sillanaukee, P. (1998). Five shot questionnaire on heavy drinking. *Alcoholism: Clinical and Experimental Research, 22*, 1788–1791. DOI: 10.1111/j.1530-0277.1998.tb03981.x

Shipstead, Z., Redick, T. S., & Engle, R. W. (2012). Is working memory training effective? *Psychological Bulletin, 138*, 628–654. DOI: 10.1037/a0027473

Singleton, R. A. (2007). Collegiate alcohol consumption and academic performance. *Journal of Studies on Alcohol and Drugs, 68*(4), 548–555. DOI: 10.15288/jsad.2007.68.548

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*, 149–161. DOI: 10.1016/0306-4603(86)90040-7

Somerville, L. H. (2013). The teenage brain: Sensitivity to social evaluation. *Current Directions in Psychological Science, 22*(2), 129–135. DOI: 10.1177/0963721413476512

Sonuga-Barke, E. J., Brandeis, D., Cortese, S., Daley, D., Ferrin, M., Holtmann, M., . . . Sergeant, J., European ADHD Guidelines Group (2013). Nonpharmacological interventions for ADHD: Systematic review and meta-analyses of randomized controlled trials of dietary and psychological treatments. *American Journal of Psychiatry, 170*, 275–289. DOI: 10.1176/appi.ajp.2012.12070991

Sportel, B. E., de Hullu, E., de Jong, P. J., & Nauta, M. H. (2013). Cognitive bias modification versus CBT in reducing adolescent social anxiety: A randomized controlled trial. *PLoS One, 8*(5), e64355. DOI: 10.1371/journal.pone.0064355

Stacy, A. W. (1997). Memory activation and expectancy as prospective predictors of alcohol and marijuana use. *Journal of Abnormal Psychology, 106*, 61–73. DOI: 10.1037/0021-843X.106.1.61

Staugaard, S. R. (2009). Reliability of two versions of the dot-probe task using photographic faces. *Psychology Science Quarterly, 51*, 339–350.

Steinberg, L. (2007). Risk taking in adolescence. New perspectives from brain and behavioral science. *Current Directions in Psychological Science, 16*(2), 55–59. DOI: 10.1111/j.1467-8721.2007.00475.x

Sweetser, P., & Wyeth, P. (2005). GameFlow: A model for evaluating player enjoyment in games. *Computers in Entertainment, 3*, 3–24. DOI: 10.1145/1077246.1077253

Tapert, S. F., Cheung, E. H., Brown, G. G., Frank, L. R., Paulus, M. P., Schweinsburg, A. D., Meloy, M. J., & Brown, S. A. (2003). Neural response to alcohol stimuli in adolescents with alcohol use disorder. *Archives of General Psychiatry, 60*, 727–735. DOI: 10.1001/archpsyc.60.7.727

Thatcher, D. L., & Clark, D. B. (2008). Adolescents at risk for substance use disorders. Role of psychological dysregulation, endophenotypes, and environmental influences. *Alcohol Research & Health, 31*(2), 168–176. PMID: 23584818

Thush, C., Wiers, R. W., Ames, S. L., Grenard, J. L., Sussman, S., & Stacy, A. W. (2008). Interactions between implicit cognition and working memory in the prediction of alcohol use in at-risk adolescents. *Drug and Alcohol Dependence, 94*, 116–124. DOI: 10.1016/j.drugalcdep.2007.10.019

Thush, C., Wiers, R. W., Moerbeek, M., Ames, S. L., Grenard, J. L., Sussman, S., & Stacy, A. W. (2009). Influence of motivational interviewing on explicit and implicit alcohol-related cognition and alcohol use in at-risk adolescents. *Psychology of Addictive Behaviors, 23* (1), 146–151. DOI: 10.1037/a0013789

Townshend, J. M., & Duka, T. (2002). Patterns of alcohol drinking in a population of young social drinkers: A comparison of questionnaire and diary measures. *Alcohol & Alcoholism, 37,* 187–92.
DOI: 10.1093/alcalc/37.2.187

van Bockstaele, B., Salemink, E., Bögels, S. M., & Wiers, R. W. (2015). Limited generalisation of changes in attentional bias following attentional bias modification with the visual probe task. *Cognition and Emotion, 23*, 1–8. DOI: 10.1080/02699931.2015.1092418

van der Oord, S., Ponsioen, A. J. G. B., Geurts, H. M., ten Brink, E. L., & Prins, P. J. M. (2012). A pilot study of the efficacy of a computerized executive functioning remediation training with game elements for children with ADHD in an outpatient setting: Outcome on parent and teacher-rated executive functioning and ADHD behavior. *Journal of Attention Disorders, 18*(8), 699–712. DOI: 10.1177/1087054712453167

van Deursen, D. S., Salemink, E., Boendermaker, W. J., Pronk, T., Hofmann, W., & Wiers, R. W. (2015). Executive functions and motivation as moderators of the relationship between automatic associations and alcohol use in problem drinkers seeking online help. *Alcoholism: Clinical and Experimental Research, 39*(9), 1788–1796. DOI: 10.1111/acer.12822 van Deursen, D. S., Salemink, E., Smit, F., Kramer, J., & Wiers, R. W. (2013). Web-based cognitive bias modification for problem drinkers: Protocol of a randomized controlled trial with a 2x2x2 factorial design. *BMC Public Health, 13*, 674. DOI: 10.1186/1471-2458-13-674

van Schie, S., & Boendermaker, W. J. (2014). *Measuring attentional bias towards alcohol in adolescents using motivating game elements* (Unpublished master's thesis). University of Amsterdam: Amsterdam, the Netherlands.

Verbeken, S., Braet, C., Goossens, L., & van der Oord, S. (2013). Executive function training with game elements for obese children: A novel treatment to enhance self-regulatory abilities for weight-control. *Behaviour Research and Therapy*, *51* (6), 290–299. DOI: 10.1016/j.brat.2013.02.006.

Verdejo-Garcia, A., Lawrence, A. J., & Clark, L. (2008). Impulsivity as a vulnerability marker for substance-use disorders: Review of findings from high-risk research, problem gamblers and genetic association studies. *Neuroscience and Biobehavioral Reviews, 32*, 777–810. DOI: 10.1016/j.neubiorev.2007.11.003

Wechsler, D. (1981). *Wechsler Adult Intelligence Scale-Revised*. Psychological Corporation, San Antonio, TX. DOI: 10.1177/073428298300100310

Werch, C. E., & Owen, D. M. (2002). Iatrogenic effects of alcohol and drug prevention programs. *Journal of Studies on Alcohol, 63*, 581–590. DOI: 10.15288/jsa.2002.63.581

White, H. R., & Labouvie, E. W. (2000). Longitudinal trends in problem drinking as measured by the Rutgers alcohol problem index. *Alcoholism: Clinical and Experimental Research, 24*, 76A.

Wiers, R. W., Bartholow, B. D., van den Wildenberg, E., Thush, C., Engels, R. C. M. E., Sher, K. J., Grenard, J., Ames, S. L., & Stacy, A. W. (2007). Automatic and controlled processes and the development of addictive behaviors in adolescents: A review and a model. *Pharmacology, Biochemistry, and Behavior, 86*(2), 263–283. DOI: 10.1016/j.pbb.2006.09.021

Wiers, R. W., Boelema, S., Nikolaou, K., & Gladwin, T. E. (2015). On the development of implicit and control processes in relation to substance use in adolescence. *Current Addiction Reports, 2*(2), 141–155. DOI: 10.1007/s40429-015-0053-z

Wiers, R. W., Eberl, C., Rinck, M., Becker, E., & Lindenmeyer, J. (2011). Retraining automatic action tendencies changes alcoholic patients' approach bias for alcohol and improves treatment outcome. *Psychological Science, 22*, 490–497. DOI: 10.1177/0956797611400615

Wiers, R. W., Gladwin, T. E., Hofmann,W., Salemink, E., & Ridderinkhof, K. R. (2013). Cognitive bias modification and cognitive control training in addiction and related psychopathology: Mechanisms, clinical perspectives, and ways forward. *Clinical Psychological Science, 1*(2), 192–212. DOI: 10.1177/2167702612466547

Wiers, R. W., Hoogeveen, K. J., Sergeant, J. A., & Gunning, B. W. (1997). High- and low-dose alcohol-related expectancies and the differential associations with drinking in male and female adolescents and young adults. *Addiction, 92,* 871–888. DOI: 10.1111/j.1360-0443.1997.tb02956.x

Wiers, R. W., Houben, K., Fadardi, J. S., van Beek, P., Rhemtulla, M. T., & Cox, W. M. (2015). Alcohol cognitive bias modification training for problem drinkers over the web. *Addictive Behaviors, 40*, 21–26. DOI: 10.1016/j.addbeh.2014.08.010

Wiers, R. W., Rinck, M., Dictus, M., & van den Wildenberg, E. (2009). Relatively strong automatic appetitive action-tendencies in male carriers of the OPRM1 G-allele. *Genes Brain & Behavior, 8*, 101–106. DOI: 10.1111/j.1601-183X.2008.00454.x

Wiers, R. W., Rinck, M., Kordts, R., Houben, K., & Strack, F. (2010). Retraining automatic action-tendencies to approach alcohol in hazardous drinkers. *Addiction, 105*, 279–287. DOI: 10.1111/j.1360-0443.2009.02775.x

Wiers, R. W., van de Luitgaarden, J., van den Wildenberg, E., & Smulders, F. T. Y. (2005). Challenging implicit and explicit alcohol-related cognitions in young heavy drinkers. *Addiction, 100*, 806–819. DOI: 10.1111/j.1360-0443.2005.01064.x

Wilcutt, E. G., Doyle, A. E., Nigg, J. T., Faraone, S. V., & Pennington, B. F. (2005). Validity of the executive function theory of ADHD: A meta-analytic review. *Biological Psychiatry, 57*, 1336–1346. DOI: 10.1016/j.biopsych.2005.02.006

Wobbrock, J. O., Findlater, L., Gergle, D., & Higgins, J. J. (2011). The aligned rank transform for nonparametric factorial analyses using only ANOVA procedures. In D. S. Tan, S. Amershi, B. Begole, W. A. Kellogg, & M. Tungare (Eds.), *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 143–146). New York, NY: ACM Press. ISBN: 978-1450302289

Wood, M. D., Sher, K. J., & McGowan, A. K. (2000). Collegiate alcohol involvement and role attainment in early adulthood: Findings from a prospective high-risk study. *Journal of Studies on Alcohol, 61* (2), 278–289. DOI: 10.15288/jsa.2000.61.278

A

Appendix B
Summary in English

Serious Gamification - Motivating Adolescents to do Cognitive Training

Adolescence is a developmental period characterized by a considerable increase in the prevalence externalizing behavior, such as experimenting with alcohol. Although this behavior does not necessarily lead to mental problems, excessive use at this age can lead to health problems, school dropout, and addiction problems later in life. As such, early intervention is important to prevent escalation. Training or re-training of cognitive processes, such as working memory capacity and automatic attentional processes towards alcohol, can be effective in helping adolescents to get a grip on and decrease their alcohol use. However, these training procedures are usually perceived as long and boring. The use of serious games may provide a solution by aiding in motivating adolescents to do cognitive training. The research behind the application of gaming techniques to evidence-based cognitive training paradigms, and cognitive bias modification (CBM) in particular, is still in its infancy. As such, the general aim of this projects was to investigate if and how serious gaming techniques can be used to motivate adolescents and young adults to do cognitive training that is aimed at decreasing their alcohol intake. To do so we first developed a framework on which to base these serious games. The framework distinguishes between a number of steps in the development of serious games, based on cognitive training (CBM in particular), starting with an original, evidence based training paradigm and adding game-elements in various degrees.

Based on this framework, we developed and investigated a number of serious games, which were the basis of the other chapters in this thesis. There were three goals in each of these serious game training studies. The first goal was to restore balance to the cognitive system, by either strengthening the adolescents' cognitive control functions (e.g., working memory, inhibition), or modifying their automatic, appetitive processes (e.g., attention or approach tendencies towards alcohol) that typically develop through prolonged alcohol use. The second goal, as a result of the cognitive change, was to decrease adolescents' problematic alcohol use. The third and final goal was to evaluate whether the game variant of the training indeed increased motivation to train amongst the sample beyond the level of that in the regular, non-gamified training. The first two empiric studies described in this thesis (**Chapters 3** and **4**) made use of a shell-type game called the CityBuilder Game, which relates to Step 4 in the model (**Chapter 2**). **Chapter 3** shows that serious game elements can help to increase motivation to increase working memory capacity in high school adolescents. Although the training turned out to be successful at increasing working memory capacity, it did not do so beyond the level of increase found in the control condition in which participants only trained with relatively easy exercises (sequences of three squares). Importantly, motivation to do the post-training working memory assessment was found to have significantly lowered in the gamified WMC-training condition, relative to the non-game and control training conditions. This may be due to the sudden lack of motivating rewards during the post-training assessment after multiple sessions of game-training in this condition, specifically. It could therefore be that the training gain of participants in the game condition was underestimated. Interestingly, the self-reported, explicit measure of motivation to train did not show participants had a preference for the game version of the training, but participants in the gamified condition did train significantly longer. The game was thus able to increase motivation to train, but the effect did fade over time until it was no longer different from the non-game conditions. The training did not lead to significantly lower drinking, which may have been due to the fact that drinking levels in the school-based sample were very low at baseline.

The second study with the CityBuilder Game (**Chapter 4**) was specifically aimed at identifying the specific role of point-rewards (cf. Step 1) and point-rewards with value in a shell game (cf. Step 4) as alcohol-specific inhibition reinforcers. Alcohol-specific inhibition was found to increase, and self-reported alcohol intake was found to be lower, in all conditions after the Go/No-Go training. However, the effect of the training on alcohol-specific inhibition was more pronounced when the training itself was also alcohol-specific. Rewarding of correct behavior during training with arbitrary points or with points that held value in the shell-game was found not to significantly reinforce this training effect beyond giving minimal feedback only.

The next two chapters (**Chapters 5** and **6**) focused on exploring integrating elements of fun into the training paradigms (cf. Steps 2 and 3) by adding a social gaming context, swiping gestures in a mobile environment, and extensive visuals. The first study in **Chapter 5** described the Cheese Ninja Game, where game elements were integrated into an evidence-based CBM paradigm. The game was also embedded inside a social media environment (www.facebook.com) to measure the added effects of social feedback. While no cognitive training effects were found, adding (social) game elements did increase appreciation for the training as well as participants' motivation to train. The second study presented in this chapter concerned a mobile CBM training, which appeared to increase motivation to train in terms how often participants trained, but this effect disappeared after controlling for baseline motivation to train. Importantly, despite changing several key aspects of the normal training paradigms, both variants did not underperform compared to their regular training counterparts in this student sample. Finally, in **Chapter 6** the Shots Game was evaluated. This game, based on the visual probe attentional bias retraining paradigm, introduced mainly an elaborate point-rewards system (cf. Step 1), as well as some fancy graphics and sounds, integrated well into the task (cf. Step 2). However, contrary to the other games presented in this thesis, the Shots Game had no additional gameplay elements included. The results showed that the attentional bias for alcohol was reduced only by the regular visual probe training, and not by playing the game. Nevertheless, self-reported drinking behavior was not affected by any of the training variants. As expected, motivation to train was shown to decrease over time, but this happened in all conditions, suggesting that the motivational features of the Shots Game were not enough to counteract the tiresome nature of the training. Moreover, motivation to start drinking less after the training actually decreased after playing the game, which may indicate potential detrimental effects of a disappointing gamification. Thus we showed here, inadvertently, that gamification is not without its risks, which has important implications for future serious game development and research.

Conclusions

In sum, this thesis has provided a new, evidence-based framework for the development and evaluation of serious games based on cognitive training paradigms, accompanied by a number of studies investigating the various potential benefits of such modifications. It can be concluded that serious games can be a promising new way to reach at risk youth, through prevention as well as intervention, and cognitive training can be a firm scientific basis for the design of those serious games. However, as this field is still young, more research is needed to determine for whom these cognitive training games work best, and which game elements should be used or avoided. As such, serious game research, as well as training outcomes, could benefit from keeping expectations modest.

For a broader discussion of the results presented in this thesis, the reader is referred to **Chapter 7**.

Appendix C
Samenvatting in het Nederlands

Serious Gamification - Adolescenten Motiveren voor Cognitieve Training

De adolescentie is een periode in de ontwikkeling van een kind die wordt gekenmerkt door een opvallende toename van externaliserend gedrag, waaronder het experimenteren met alcohol. Hoewel dit soort gedrag niet per definitie tot problemen hoeft te leiden, kan overmatig drankgebruik op deze leeftijd wel leiden tot gezondheidsproblemen, schooluitval en op latere leeftijd verslavingsgerelateerde problemen. Daarom is het belangrijk dat er vroeg wordt ingegrepen, om escalatie te voorkomen. Het trainen, of her-trainen, van cognitieve processen, zoals de werkgeheugencapaciteit en automatische aandachtsprocessen naar alcohol toe, kunnen effectief zijn om adolescenten meer grip te laten krijgen op hun alcoholgebruik. Deze trainingen worden echter vaak ervaren als lang en saai. Het gebruiken van serious games kan hier wellicht uitkomst bieden, door adolescenten extra te motiveren om de cognitieve trainingen te doen. Het onderzoek achter de toepassing van speltechnieken bij op wetenschappelijk onderzoek gebaseerde cognitieve trainingparadigma's, en *cognitive bias modification* (het hertrainen van cognitieve biassen, ook wel bekend als CBM) in het bijzonder, bevindt zich nog in de kinderschoenen. Daarom was het algemene doel van dit project om te onderzoeken of en hoe *serious gaming* technieken kunnen worden toegepast om adolescenten en jongvolwassenen te motiveren om cognitieve trainingen te volgen die erop gericht zijn om hun alcoholinname te verminderen. Om dit te onderzoeken hebben we eerst een model ontwikkeld waarop dit soort *serious* *games* kunnen worden gebaseerd. Het model maakt onderscheid tussen een aantal stappen in de ontwikkeling van *serious games*, die gebaseerd worden op cognitieve trainingen (CBM in het bijzonder), beginnende bij het originele, op wetenschappelijk onderzoek gebaseerde trainingsparadigma, waar vervolgens in verschillende mate game elementen aan worden toegevoegd. Op basis van dit model hebben we een aantal *serious games* ontwikkeld en onderzocht, wat de basis vormt van de hierna volgende hoofdstukken in dit proefschrift. Elk van deze *serious game* training studies had drie doelen. Het eerste doel was om de balans in het cognitieve systeem te herstellen, door enerzijds de cognitieve controlefuncties van adolescenten te versterken (bijvoorbeeld werkgeheugen, inhibitievermogen), of anderzijds door het hertrainen (modificeren) van hun automatische, aan de alcohol gerelateerde processen (zoals aandacht en toenaderingneigingen naar de alcohol toe), die zich vaak ontwikkelen bij langdurig alcoholgebruik. Het tweede doel, als gevolg van het eerste doel, was om het daadwerkelijke alcoholgebruik van de jongeren te verminderen. Het derde en laatste doel was om te evalueren of de spelvariant van de training inderdaad de motivatie om te trainen in de steekproefpopulatie kon verhogen tot voorbij het niveau van de reguliere training zonder spelelementen. De eerste twee empirische studies die worden beschreven in dit proefschrift (**Hoofdstuk 3** en **4**) maakten daarbij gebruik van een schil-type spel genaamd de CityBuilder Game (zie Stap 4 in het model in **Hoofdstuk 2**). **Hoofdstuk 3** laat zien dat *serious game* elementen inderdaad kunnen helpen bij het verhogen van de motivatie van middelbare scholieren om de werkgeheugencapaciteit te verhogen. Hoewel de training succesvol bleek in het vergroten van de werkgeheugencapaciteit, was dit effect niet groter dan in de controleconditie waar de deelnemers met relatief makkelijke oefeningen trainden (sequenties van drie blokjes). Een andere belangrijke bevinding van het onderzoek was dat de motivatie om de werkgeheugen meting te doen significant verlaagd was na de training in de spelconditie, vergeleken met de andere condities. Dit zou kunnen komen door het plotseling ontbreken van motiverende beloningen tijdens de nameting, na meerdere sessies van belonende speltraining, wat alleen in deze conditie gebeurde. Het zou daardoor kunnen zijn dat de vooruitgang op de training van deelnemers in de spelconditie is onderschat. Een interessante bevinding is ook dat de zelf-gerapporteerde, expliciete maat van motivatie om te trainen niet liet zien dat deelnemers een voorkeur hadden voor de spelconditie, maar dat deelnemers in de spelconditie wel significant langer doortrainden. Het spel was dus wel degelijk in staat om de motivatie om te trainen effectief te verhogen, maar dit effect doofde wel uit naar mate de training voortduurde, tot op het punt waar het niet meer verschilde van de andere condities. De training leidde niet tot een vermindering in drinken, maar dit heeft wellicht te maken met het feit dat er aan het begin van de studie al erg weinig alcohol werd gedrronken in deze middelbare school steekproef. De tweede studie met de CityBuilder Game (**Hoofdstuk 4**) was er specifiek op gericht om de precieze rol te identificeren van beloningspunten (zie Stap 1) en beloningspunten met waarde in een game-schil (zie Stap 4) als bekrachtigers van alcohol-specifieke inhibitie. Alcohol-specifieke inhibitie bleek vergroot, en de zelf-gerapporteerde inname van alcohol was lager, in alle condities na de training. Echter, het effect van de training op alcohol-specifieke inhibitie was duidelijker te zien wanneer de training zelf ook alcohol-specifiek was. Het belonen van correct gedrag tijdens de training door middel van arbitraire punten of met punten die waarde hadden in de spel-schil bleek dit trainingseffect niet significant meer te versterken dan wanneer er alleen minimale feedback werd gegeven. In de volgende twee hoofdstukken (**Hoofdstuk 5** en **6**) lag de focus op het verkennen van het integreren van motiverende elementen in de trainingsparadigma's zelf (zie Stap 2 en 3), door het toevoegen van een *social gaming* context, veegbewegingen op een mobieltje, en met uitgebreide visuele elementen. De eerste studie in **Hoofdstuk 5** beschrijft de Cheese Ninja Game, waar spelelementen geïntegreerd werden in een op wetenschappelijk onderzoek gebaseerd CBM paradigma. Het spel werd tevens ingebed in een *social media* omgeving (www.facebook.com) om te kijken naar het toegevoegde effect van sociale feedback. Hoewel er geen cognitief trainingseffect werd gevonden bleek dat het toevoegen van (sociale) spelelementen wel zorgde voor een hogere waardering van de training, alsmede een grotere motivatie om te trainen. De tweede studie die wordt gepresenteerd in dit hoofdstuk betrof een mobiele CBM training, die de motivatie om te trainen in termen van hoeveel de deelnemers trainden leek te verhogen, maar nadat werd gecontroleerd voor motivatie aan de start van de training verdween dit effect. Een belangrijk punt is dat hoewel diverse kernelementen van de normale trainingsparadigma's werden aangepast, bleek dat beide varianten het ook niet slechter deden dan de reguliere varianten in deze studentensteekproef. Tot slot werd in **Hoofdstuk 6** de Shots Game geëvalueerd. Dit spel was gebaseerd op het *visual probe* aandachts bias hertrainingsparadigma en introduceerde voornamelijk een uitgebreid punten beloningssysteem (zie Stap 1) en een mooie visuele omgeving, aangevuld met diverse geluidseffecten, allemaal goed geïntegreerd in de taak zelf (zie Stap 2). In tegenstelling tot de andere spellen die in dit proefschrift werden beschreven, bevatte de Shots Game echter geen aanvullende *gameplay* elementen. De resultaten lieten zien dat de aandachtsbias voor alcohol alleen verminderde in de reguliere *visual probe* training, en niet door het spelen van het spel. Desalniettemin bleek het zelf-gerapporteerde drinkgedrag door geen van de trainingsvarianten te worden beïnvloed. Zoals verwacht bleek dat motivatie om te trainen afnam naarmate de training voortduurde, maar dit gebeurde in alle condities, wat suggereerde dat de motivationele aspecten van de Shots Game niet genoeg waren om de saaie aard van de training tegen te gaan. Bovendien bleek dat de motivatie van deelnemers om minder te gaan drinken na de training zelfs verminderd was na het spelen van het spel, wat kan wijzen op mogelijke schadelijke effecten van een teleurstellende vergaming. We hebben hier dus, onbedoeld, aangetoond dat gamificatie niet geheel zonder risico is, wat belangrijke implicaties heeft voor toekomstige ontwikkeling van en onderzoek naar *serious games*.

Conclusies

Dit proefschrift biedt een nieuw, op wetenschappelijk onderzoek gebaseerd model voor de ontwikkeling en evaluatie van *serious games* die gebaseerd worden op cognitieve trainingsparadigma's, vergezeld door een aantal studies die de verschillende potentiële voordelen van dergelijke aanpassingen bestuderen. Er kan worden geconcludeerd dat *serious games* een veelbelovende, nieuwe manier kunnen zijn om jongere risicogroepen te bereiken, zowel via preventie als interventie, en dat cognitieve training een solide, wetenschappelijke basis kan vormen voor het ontwerpen van die *serious games*. Aangezien dit onderzoeksveld echter nog vrij jong is, zal er meer onderzoek moeten worden gedaan om nader te bepalen voor welke mensen dit soort cognitieve trainingsspellen nu het beste werken, en welke spelelementen beter wel of niet kunnen worden gebruikt. Daarom doet men er vooralsnog verstandig aan om bij serious game onderzoek, alsmede bij het interpreteren van de uitkomsten van de game trainingen, de verwachtingen bescheiden te houden.

Voor een uitgebreidere discussie omtrent de bevindingen die voortkomen uit dit proefschrift wordt de lezer verwezen naar **Hoofdstuk 7**.

Appendix D
Contributions of co-authors

All researchers were supported by the National Initiative Brain & Cognition Grant 433-11-1510, unless noted otherwise.

Parts of **Chapter 1** are based on the following publication:
Boendermaker, W. J., Peeters, M., Prins, P. J. M., & Wiers, R. W. (2017). Using serious games to (re)train cognition in adolescents. In M. Ma & A. Oikonomou (Eds.), *Serious Games and Edutainment Applications, Volume II.* UK: Springer-Verlag. Chapter (#14) in press. ISBN: 978-3-319-51643-1

Wouter Boendermaker wrote the manuscript. Margot Peeters contributed to the parts that were in the publication (*Training Cognitive Processes* and *Motivation(s)*); Pier Prins and Reinout Wiers reviewed and approved the entire manuscript.

Chapter 2 is based on the following publication:
Boendermaker, W. J., Prins, P. J. M., & Wiers, R. W. (2015). Cognitive bias modification for adolescents with substance use problems – Can serious games help? *Journal of Behavior Therapy and Experimental Psychiatry, 49*, 13-20. DOI: 10.1016/j.jbtep.2015.03.008

Wouter Boendermaker designed the theoretic framework with feedback from Pier Prins and Reinout Wiers; Wouter Boendermaker wrote the manuscript; All co-authors reviewed and approved the manuscript.

Chapter 3 has been submitted for publication as:
Boendermaker, W. J., Gladwin, T. E., Peeters, M., Prins, P. J. M., & Wiers, R. W. (2016). Evaluating working memory training with serious game elements in adolescents: A randomized controlled trial.

Wouter Boendermaker designed the study with feedback from Thomas Gladwin, Pier Prins, and Reinout Wiers. Wouter Boendermaker programmed the computerized assessment and training tasks, including the game, and set up the experimental procedure using the UvA SurveyTool. Elisabeth Enthoven provided the game's graphics. Wouter Boendermaker supervised the participant recruitment and data collection process, aided by students Nathan Bleijenberg, Bouwien Westerhuis, Emmily Harwig, Marcia Dominicus, and Wendy Kuijn. Wouter Boendermaker and Margot Peeters conducted the statistical analyses. Wouter Boendermaker wrote the manuscript. All co-authors reviewed and approved the manuscript. Thomas Gladwin and Reinout Wiers were also supported by the VICI grant 453-08-001, funded by the Dutch National Science Foundation (N.W.O.). Thomas Gladwin was also supported by the European Foundation for Alcohol Research (E.R.A.B.) grant EA 12 39.

Chapter 4 is in preparation for submission as:
Boendermaker, W. J., Nikolaou, K., & Wiers, R. W. (2016). Cognitive training in adolescents: A comparison of training reinforcements.

Kyriaki Nikolaou and Wouter Boendermaker designed the study with feedback from Reinout Wiers. Wouter Boendermaker programmed the computerized assessment and training tasks, including the game, with the exception of the cue-reactivity task and the stop-signal task, which were provided by Kyriaki Nikolaou. Bruno Boutin and Tim de Jong assisted in setting up the experimental procedure in LOTUS. Elisabeth Enthoven provided the game's graphics. Kyriaki Nikolaou supervised the participant recruitment and fMRI data collection process, aided by students Annelieke Hagen and Sam Prinssen. Wouter Boendermaker conducted the statistical analyses with feedback from Kyriaki Nikolaou. Wouter Boendermaker and Kyriaki Nikolaou wrote the manuscript. All co-authors reviewed and approved the manuscript.

Chapter 5 is based on the following publication:
Boendermaker, W. J.*, Boffo, M.*, Wiers, R. W. (2015). Exploring elements of fun to motivate youth to do cognitive bias modification. *Games for Health Journal, 4*(6), 434-443. DOI: 10.1089/g4h.2015.0053

* Wouter Boendermaker and Marilisa Boffo share first authorship.

Wouter Boendermaker designed the Game Study. Wouter Boendermaker and Marilisa Boffo designed the Mobile Study. Wouter Boendermaker programmed the computerized assessment and training tasks, including the game and the mobile application, and set up the experimental procedure in LOTUS. Wouter Boendermaker supervised the participant recruitment and data collection process for the Game Study, aided by students Antonios Georgiadis, Ans de Nijs, and Wendelien Steltenpöhl. Wouter Boendermaker and Marilisa Boffo supervised the participant recruitment and data collection process for the Mobile Study, aided by students Elisa van der Plas, Nienke van Bueren, Roberto Floris, and Vita Jongen. Wouter Boendermaker and Marilisa Boffo conducted the statistical analyses. Wouter Boendermaker wrote the manuscript. All co-authors reviewed and approved the manuscript. Marilisa Boffo was also supported by the NextLevel project grant 314-99-102, and Reinout Wiers by the VICI grant 453-08-001, both funded by the Dutch National Science Foundation (N.W.O.).

Chapter 6 is based on the following publication:
Boendermaker, W. J., Sanchez Maceiras, S., Boffo, M., & Wiers, R W. (2016). Attentional bias modification with serious game elements: Evaluating the Shots Game. *Journal of Medical Internet Research: Serious Games, 4*(2): e20. DOI: 10.2196/games.6464

Wouter Boendermaker and Soraya Sanchez Maceiras designed the study. Wouter Boendermaker programmed the computerized assessment and training tasks, including the game, and set up the experimental procedure in LOTUS. Sylvia van Schie provided the game's graphics. Wouter Boendermaker supervised the participant recruitment and data collection process, aided by students Soraya Sanchez Maceiras, Aida Alai, Guillaume Rensink, and Sylvia van Schie. Wouter Boendermaker conducted statistical analyses with feedback from Marilisa Boffo. Wouter Boendermaker wrote the manuscript. All co-authors reviewed and approved the manuscript. Marilisa Boffo was also supported by the NextLevel project grant 314-99-102, and Reinout Wiers by the VICI grant 453-08-001, both funded by the Dutch National Science Foundation (N.W.O.).

Parts of **Chapter 7** are based on the following publication:
Boendermaker, W. J., Peeters, M., Prins, P. J. M., & Wiers, R. W. (2017). Using serious games to (re)train cognition in adolescents. In M. Ma & A. Oikonomou (Eds.), *Serious Games and Edutainment Applications, Volume II*. UK: Springer-Verlag. Chapter (#14) in press. ISBN: 978-3-319-51643-1

Wouter Boendermaker wrote the manuscript; Margot Peeters contributed to the parts of the manuscript that were in the publication (parts of *Critical Reflections I - on the effectiveness of serious games for cognitive training* and *Recommendations for future research, development and evaluation of serious games based on cognitive training*); Pier Prins and Reinout Wiers reviewed and approved the entire manuscript.

D

Appendix E
Acknowledgements (dankwoord)

Het schrijven van een proefschrift is een omvangrijke onderneming die niet had kunnen plaatsvinden zonder de nodige input van anderen, zowel op professioneel als op persoonlijk vlak. Allereerst wil ik daarom mijn begeleiders bedanken voor hun bijdragen gedurende de afgelopen jaren.

Reinout, ik leerde je kennen toen ik bij de TOP kwam te werken op je VICI project. Sindsdien bood je me steeds de gelegenheid om door te groeien, zowel binnen het ADAPT team als daarbuiten: van coördinator van de CSCA summer school en zij-projecten op de VU en met Nijmegen, via onze subsidieaanvraag naar de promotieplek. Ik zal ook niet snel vergeten hoe je me het blijde nieuws vertelde, dat de subsidie (eindelijk) was toegekend: Tim belde in lichte paniek dat ik zo snel mogelijk naar de afdeling moest komen omdat 'het helemaal misging met het Loes-systeem van Thomas'; jij zou woedend zijn en ik meen zelfs dat ik Janna hoorde gillen op de achtergrond. Eenmaal aangekomen bleek dat de champagne klaarstond en, ik geef het eerlijk toe, ik was er goed ingestonken! Dat tekende ook de sfeer binnen het team: altijd een vrolijke noot. Bovendien genoot ik tijdens het project veel vrijheid en vertrouwen, maar was je altijd beschikbaar voor een brainstormsessie, en dat heb ik altijd zeer gewaardeerd. **Pier**, jou kende ik al uit mijn studietijd en als begeleider van mijn masterthese. Hoewel je tijdens het project met emeritaat ging was je altijd bereikbaar voor feedback en, belangrijker nog, wijze raad met betrekking tot het omgaan met een aantal van de tegenslagen die we in het project hebben gehad. Dat heb ik altijd als zeer waardevol ervaren. Tot slot ook nadrukkelijk dank aan mijn co-promotor en dagelijks begeleider tijdens de eerste twee jaar van het project, **Thomas Gladwin**. Naast de talloze, en soms

eindeloze, maar altijd boeiende discussies over werk- en andersgerelateerde onderwerpen, die soms tot in de late uurtjes konden duren was er bij jou altijd tijd voor gezelligheid. Bedankt voor al je wijze inzichten!

Ik wil hier ook graag mijn dank uitspreken aan de leden van de beoordelingscommissie, de hoogleraren **Hilde Geurts**, **Jaap Murre** en **Geertjan Overbeek** van de Universiteit van Amsterdam, **Isabela Granic** van de Radboud Universiteit Nijmegen en **Lydia Krabbendam** van de Vrije Universiteit Amsterdam, alsook dr. **Brian Ostafin** van de Rijksuniversiteit Groningen, voor het nemen van de tijd en moeite om mijn proefschrift van voor naar achter door te spitten.

Wat is een promovendus zonder zijn paranimfen? **Mieke**, we hebben nogal wat lief en leed met elkaar gedeeld de afgelopen jaren, zowel op werk- als persoonlijk vlak. Je was altijd bereid om mijn klaagzang aan te horen (soms had je zelf ook een mooi couplet); je was een echte kameraad. **Marilisa**, words cannot express how much admiration I have for you. You are without a doubt one of the strongest, most caring, stubborn, loving creatures I have ever met, and I'll be forever thankful that I have. Not only did you contribute to several of my studies, you were always there for some friendly advice, whether it be about statistics or life in general. Your relentless optimism has undoubtedly got me through several of the more difficult periods of the project, and I thank you for that, my lady.

Door de jaren heen heb ik met verschillende kamergenoten lief en leed mogen delen. Natuurlijk wil ik iedereen bedanken voor de gezelligheid. In het bijzonder **Bas**, jou kende ik al vanuit de masterthese bij Pier en het BrainGame Brian project. We konden soms flink botsen in onze discussies, maar je bent ook met recht een rots in de branding geweest, en je oprechtheid en eerlijkheid was doorgaans een ware verfrissing. Bovendien wisten we elkaar voor allerlei zaken te vinden, of het nu om een symposium, een gastcollege of om technisch werk ging, dus ik heb altijd fijn met je samengewerkt. **Leone**, wij zaten vlak bij elkaar en dat leidde vaak tot grappige verhalen en ook nog eens diverse samenwerkingen. Dankzij die samenwerking kon ik bovendien de grafische ondersteuning voor game uitbesteden aan **Elisabeth Enthoven** (wat overigens echt een verbetering was!). Ook wij hebben lief en leed gedeeld en ik

wil je dan ook bedanken voor een heel gezellige tijd. Tot slot **Bianca**, jij vulde ons kamerkwartet aan. Samen hebben wij op onze kamer denk ik meer dan honderd studenten begeleid en jij was daarin beslist onze topper! Vaak heb ik je dan ook om advies gevraagd als er weer eens een lastige situatie was ontstaan. Petje af voor wat jij allemaal hebt gedaan!

Naast kamergenoot, TOP-buddy en part-time nemesis natuurlijk ook een knipoog naar **Thomas Pronk**. Wat ben jij toch een veelzijdig figuur: net als je denkt dat je voorspelbaar begint te worden verbaas je me weer met een van je prachtige verhalen. Toch verdien je ook mijn welgemeende dank voor al je tips en trucs over programmeren, en natuurlijk je allereerste uitnodiging om bij de TOP/ADAPT te komen werken. We hebben onze nukken gehad, maar al met al zou ik mijn UvA tijd niet zonder jou kunnen voorstellen. Samen met **Tim** en **Denise** hebben we in de eerste jaren van onze projecten veel ups en downs doorgemaakt, heel veel geleerd en een onvergetelijke tijd gehad. **Tim**, ook jouw rol was een heel bijzondere: niet alleen als goede vriend en gewaardeerde collega in Amsterdam, maar ook toen ik van Amsterdam naar Eindhoven verhuisde om met Sanne te gaan samenwonen sleepte je me gelijk mee naar de zaterdag(lamzakken)voetballerij en de vele spellenavonden. En natuurlijk zal ik ook de ‘sexy pancakes and horror movies’ tijdens mijn vrijgezellenavond nooit vergeten. Kortom, ook jij hebt een belangrijke bijdrage geleverd aan een onvergetelijke tijd.

Als lid van het ADAPT onderzoeksteam heb ik natuurlijk met nog veel meer mensen samengewerkt, om adviezen gevraagd en gezelligheid mee gehad. **Elske**, **Helle, Bram**, **Nicole**, **Kiki**, **Janna**, **Si**, **Yang**, **Xinyu**, **Iman**, **Tess**, **Panote**, **Eva**, **Özlem**, **Anke**, **Dana**, **Joeri**, **David**, **Suzanne**, **Hans**, **Andrea**, **Esther**, **Poppy** en **Mae**, heel erg bedankt! In het bijzonder wil ik hierbij **Helle**, **Kiki**, **Margot** en **Denise** nog eens noemen, met wie ik op diverse projecten heb samengewerkt en van wie ik ook veel heb kunnen leren. Bovendien heb ik dankzij **Margot** een nieuwe aanstelling en een geweldig team gevonden aan de Universiteit Utrecht.

Buiten het ADAPT onderzoeksteam ook dank voor de overige leden van de afdeling ontwikkelingspsychologie. In het bijzonder natuurlijk onze (voormalige) helden van het afdelingssecretariaat, **Ellen** en **Eveline** (en in het begin **Borre**). Jullie verrichten, vaak onzichtbare, bergen werk en waren talloze malen te vinden voor allerlei praktische zaken, super! Daarnaast de activiteitencommissie (accie): **Tjitske**, **Maaike**, **Laura** en **Leone**: enorm veel lol mee gehad bij (het voorbereiden van) alle afdelingsuitjes. Hier mogen verder ook de TOPpers hier niet onvermeld blijven: **Marco**, **Bruno**, **Jasper**, **GJ**, **Marcus** en de rest. Bedankt voor de fijne samenwerking en de technische support door de jaren heen!

Voor mijn tijd als promovendus heb ik veel mogelijkheden mogen krijgen om onderzoekservaring op te doen, bij onder anderen **Sylvan Smulders**, **Marijke van de Laar**, **Mariëtte Huizinga**, en in het bijzonder **Ute-Christine Klehe**. Ik wil jullie hiervoor heel erg bedanken. Ten slotte **Eveline Zandvliet**, bedankt voor je tomeloze inzet voor de research master en haar studenten.

Buiten de UvA heb ik tevens samengewerkt met diverse onderzoekers, bedrijven, zorginstellingen en scholen, die ik hier ook graag wil bedanken: Met name **Frank Nack** voor alle samenspraak bij de begeleiding van informatiekundestudenten, **Elly Konijn**, **Johan Hoorn** en **Jolanda Veldhuis** van de Game Cella' an de VU, **Caroline Braet** en **Sandra Verbeken** van de Universiteit Gent, **Harold de Groot** van Shosho, **Michaël Bas** van Ranj en **Valentijn Visch** aan de TU Delft, de vele medewerkers bij het Trimbos Instituut, GGD Hollands Midden, Arkin/Jellinek, Mondriaan GGZ en Tactus Verslavingszorg, en tot slot de teams op het Wellant College Rijnsburg, het ROC van Twente en het ROC op Maat West in Amsterdam.

Dit onderzoek was zeker niet mogelijk geweest zonder de vele honderden scholieren en studenten die als **proefpersoon** hebben meegedaan aan het onderzoek. Ook onmisbaar was de eindeloze inzet van alle studenten die hielpen bij de dataverzameling. **Peter**, **Jade**, **Wieger**, **Christin**, **Florien**, **Nathan**, **Bouwien** en **Emmily**, **Αντώνιος** en **Roseline**, **Johanna**, **Ans** en **Sylvia**, **Tim** en **Annelieke** (waar een Lotus workshop al toe kan leiden ☺), **Martine** en **Maxime**, **Sahira**, **Nynke**, **Michael** en **Marlijn**, **Roberto**, **Rosanne**, **Nienke**, **Vita** en **Elisa**, **Soraya**, **Guillaume** en **Aïda**, superbedankt voor al

jullie inzet en vooral ook de gezelligheid die jullie meebrachten. Bovenal dank aan **Tim**, **Johanna** en **Soraya**, die zo veel extra's hebben gedaan en van wie ik minstens net zo veel heb geleerd als andersom (hoop ik)!

Ten slotte alle mensen buiten het werk om die mij in deze periode hebben gesteund, afgeleid of allebei: **Bouke**, **Bart**, **Sanne**, **Hylke** en **Kaya**, en **Pim**, **Bart**, **Frank**, **Elise**, **Joost**, **Maurice** en **Tim**. Belangrijk is ook de continue steun geweest van mijn (schoon)familie: mijn ouders, **Rien** en **Flien**, jullie hebben me altijd voor de volle honderd procent gesteund, altijd daar voor wijze adviezen als het eens wat minder ging, ontzettend bedankt! Mijn grote kleine zusje **Annemieke**, die mij altijd weer laat zien dat je nog veel harder kunt werken. Mijn schoonouders, **Frits** en **Mariet**, ook altijd open voor een goed gesprek en filosofische discussies over het wel en wee in de wetenschap. Maar bovenal mijn lieve echtgenote **Sanne**, zonder wie ik hier misschien wel niet aan was begonnen, en mijn allerliefste schatten **Jasper** en **Jade**; jullie maken het alle moeite dubbel en dwars waard!

Om dan toch in game-termen af te sluiten rest mij nog slechts één game-quote:

Job's done!

–Hearthstone; peasant in Warcraft II
(Blizzard Entertainment)

E

Appendix F
About the author

Short biography and curriculum vitae

Short biography

Wouter Boendermaker was born in Drachten, the Netherlands, on August 16th, 1982. After graduating high school at the Murmellius Gymnasium in Alkmaar, he studied information science at the Free University (VU) in Amsterdam for two years, as well as psychology at the University of Amsterdam (UvA), where he obtained his BSc in developmental psychology in 2006. In 2009 he obtained his MSc in clinical and developmental psychology, which included a clinical internship working with children with learning disabilities, as well as a Dutch psychodiagnostic assessment certificate. In 2009 he also obtained a second MSc from an international two-year research master of psychology (cum laude), with clinical psychology as major and developmental psychology as minor. After graduating he started working as a scientific programmer for prof. Reinout Wiers at the ADAPT laboratory. During this period, he assisted Reinout in writing the PhD project proposal that lead to this dissertation, which received funding in 2011. Wouter currently works as a post-doctoral researcher at the Interdisciplinary Social Science department at Utrecht University.

For more information, please visit my personal website at www.wouboe.nl.

Publications (published or in press)

Boendermaker, W. J., Veltkamp, R., Beun, R. J., van de Schoot, R., & Peeters, M. (2016). Introducing The Fling – An innovative serious game to train behavioral control in adolescents: Protocol of a randomized controlled trial. In R. Bottino, J. Jeuring, & R. C. Veltkamp (Eds.), *Proceedings of the 2016 Games and Learning Alliance (GALA) Conference* (pp. 120-129). Springer Lecture Notes in Computer Science (LNCS 10056). DOI: 10.1007/978-3-319-50182-6_11

Boendermaker, W. J., Sanchez Maceiras, S., Boffo, M., & Wiers, R. W. (2016). Attentional bias modification with serious game elements: Evaluating the Shots Game. *Journal of Medical Internet Research: Serious Games, 4*(2): e20. DOI: 10.2196/games.6464

Boendermaker, W. J., Peeters, M., Prins, P. J. M., & Wiers, R. W. (2017). Using serious games to (re)train cognition in adolescents. In M. Ma & A. Oikonomou (Eds.), *Serious Games and Edutainment Applications, Volume II*. UK: Springer-Verlag. Chapter (#14) in press. ISBN: 978-3-319-51643-1

de Voogd., E. L., Wiers, R. W., Prins, P. J. M., de Jong, P. J., **Boendermaker, W. J.**, Zwitser, R.J., & Salemink, E. (2016). Online attentional bias modification training targeting anxiety and depression in unselected adolescents: short- and long-term effects of a randomized controlled trial. *Behaviour Research and Therapy, 87*, 11-22. DOI: 10.1016/j.brat.2016.08.018

Boendermaker, W. J., Boffo, M., & Wiers, R. W. (2015). Exploring elements of fun to motivate youth to do cognitive bias modification. *Games for Health Journal, 4*(6), 434-443. DOI: 10.1089/g4h.2015.0053

van Deursen, D. S., Salemink, E., **Boendermaker, W. J.**, Pronk, T., Hofmann, W., & Wiers, R. W. (2015). Executive functions and motivation as moderators of the relationship between automatic associations and alcohol use in problem drinkers seeking online help. *Alcoholism: Clinical and Experimental Research, 39*(9), 1788-1796. DOI: 10.1111/acer.12822

Peeters, M., Janssen, T., Monshouwer, K., **Boendermaker, W.**, Pronk, T., Wiers, R., & Vollebergh, W. (2015). Weaknesses in executive functioning predict the onset of adolescents' alcohol use. *Developmental Cognitive Neuroscience, 16*, 139-146. DOI: 10.1016/j.dcn.2015.04.003

Janssen, T., Larsen, H., Peeters, M., **Boendermaker, W. J.**, Vollebergh, W. A. M., & Wiers, R. W. (2015). Do online assessed self-report and behavioral measures of impulsivity predict onset of substance use in adolescents? *Addictive Behaviors Reports, 1*, 12-18. DOI: 10.1016/j.abrep.2015.01.002

Boendermaker, W. J., Prins, P. J. M., & Wiers, R. W. (2015). Prevention in addiction: Using serious games to (re)train cognition in Dutch adolescents. In S. Göbel, M. Ma, J. Baalsrud Hauge, M. F. Oliveira, J. Wiemeyer, & V. Wendel (Eds.), *Serious Games: First Joint International Conference, JCSG 2015, Huddersfield, UK, June 3-4, 2015, Proceedings* (pp. 173-178). Switzerland: Springer International Publishing. DOI: 10.1007/978-3-319-19126-3_15.

Boendermaker, W. J., Prins, P. J. M., & Wiers, R. W. (2015). Cognitive bias modification for adolescents with substance use problems – Can serious games help? *Journal of Behavioral Therapy and Experimental Psychiatry, 49*, 13-20. DOI: 10.1016/j.jbtep.2015.03.008

Larsen, H., Kong, G., Becker, D., Cousijn, J., **Boendermaker, W.**, Cavallo, D., Krishnan-Sarin, S., & Wiers, R. (2014). Implicit motivational processes underlying smoking in American and Dutch adolescents. *Frontiers in Psychiatry, 5*, 51. DOI: 10.3389/fpsyt.2014.00051.

Drabbe, J. P., Drost, R., Klehe, U.-C., van Vianen, A.E.M., & **Boendermaker, W.** (2008). Personeelsselectie in Nederland. Aanbevelingen voor selectie van personeel in tijden van krapte (External report). Diemen: Yacht M&C. http://dare.uva.nl/record/1/294039

Manuscripts in submission or preparation

Braet, C., Verbeken, S., Houben, K., & **Boendermaker, W. J.** (2017). Short report: Computer training of attention and inhibition for youngsters with obesity: A pilot-study. Manuscript in preparation for publication.

Boendermaker, W. J., Peeters, M., & Wiers, R. W. (2017). Serious games inzetten bij drinkende jongeren. *Tijdschrift Verslaving.* Manuscript under review.

Dovis, S., & **Boendermaker, W. J.** (2017). The Stop-Range Task: A faster, more reliable alternative for the Stop task. Manuscript in preparation for publication.

Dovis, S., van der Oord, S., **Boendermaker, W. J.**, & Prins, P. J. M. (2017) Assessing executive functioning in ADHD: Why tasks and rating scales don't get along. Manuscript in preparation for publication.

Koning, I., **Boendermaker, W. J.**, Turrisi, R., & Vollebergh, W. (2017) Alcohol-specific parenting, adolescents' self-control and their implicit associations with alcohol. Manuscript in preparation for publication.

Gladwin, T. E., **Boendermaker, W. J.**, & Wiers, R. W., (2017). Directional alcohol associations and hazardous drinking. Manuscript in preparation for publication.

Boendermaker, W. J., Nikolaou, K., & Wiers, R. W. (2016). Cognitive training in adolescents: A comparison of training reinforcements. Manuscript in preparation for publication.

Boendermaker, W. J., Gladwin, T. E., Peeters, M., & Wiers, R. W. (2016). Evaluating working memory training with serious game elements in adolescents: A randomized controlled trial. Manuscript submitted for publication.

Verbeken, S.*, **Boendermaker, W. J.***, Loeys, T., Moens, E., Goossens, L., & Braet, C. (2016). Effectiveness of adding an approach avoidance training with game elements to a residential childhood obesity treatment. Manuscript submitted for publication. * Contributed equally.

National and international conference presentations

Boendermaker, W. J., Veltkamp, R., Beun, R. J., van de Schoot, R., & Peeters, M. (2016, December). *Introducing The Fling: An Innovative Serious Game to Train Behavioral Control in Adolescents: protocol of a randomized controlled trial.* **Presentation** at the Games and Learning Alliance (GALA) International Conference in Utrecht, NL.

Boendermaker, W.J., Sanchez Maceiras, S., Boffo, M., & Wiers, R. (2016, December). *Evaluating the Shots Game: Attentional Bias Modification with Serious Game Elements.* **Poster** presented at Always look on the bright side of life? Symposium in Carré, Amsterdam, NL

Boendermaker, W. J. (2016, November). *The Fling: Een Demonstratie.* **Demonstration** at Vereniging voor Gedragstherapie en Cognitieve Therapie (VGCt) Najaarscongres in Veldhoven, NL.

Boendermaker, W. J., Veltkamp, R., Beun, R. J., van de Schoot, R., & Peeters, M. (2016, November). *Introducing The Fling: A Serious Game to Train Behavioral Control in Adolescents.* **Presentation** at the 6th Games for Health Europe Conference in Utrecht, NL.

Boendermaker, W.J., Sanchez Maceiras, S., Boffo, M., & Wiers, R. (2016, June). *Evaluating the Shots Game: Attentional Bias Modification with Serious Game Elements.* **Poster** presented at The 39th Annual Scientific Meeting of The Research Society on Alcoholism, New Orleans, LA (USA).

Boendermaker, W.J., Boffo, M., Prins, P., & Wiers, R. (2016, February). *Exploring Elements of Fun to Motivate Heavy Drinking Youth to (re)train Cognition.* **Poster** presented at The Second Behaviour Change Conference: Digital Health and Wellbeingon, University College London (UCL), UK.

Boendermaker, W.J. (2015, November). *Motiveren van Jongeren bij Preventie van Verslaving: Kunnen Serious Games het (her)trainen van Cognities Ondersteunen?* **Presentation** at the 25th Annual Forum Alcohol en Drugs Onderzoek (FADO) Conference in Utrecht, NL.

Boendermaker, W.J. (2015, November). *Prevention in Addiction: Using serious games to motivate adolescents to (re)train their brain.* **Presentation** at the Experimental Psychopathology conference on Emotional Memory, Heeze, NL.

Boendermaker, W.J. (2015, November). *Prevention in Addiction: Using serious games to motivate adolescents to (re)train their brain.* **Presentation** at the 5th Games for Health Europe Conference in Utrecht, NL.

Boendermaker, W.J., Prins, P.J.M., & Wiers, R.W. (2015, June). *Prevention in addiction: Using serious games to (re)train cognition in Dutch adolescents.* **Presentation** and **poster** at The 1st Joint International Conference on Serious Games (JCSG), Huddersfield, UK.

Boendermaker, W.J., Prins, P.J.M., & Wiers, R.W. (2014, December). *Preventie van Verslaving bij Jongeren: Cognitief Trainen met een Serious Game.* **Poster** presented at NWO Brain Products Day, organized by HCMI, Utrecht, NL.

Boendermaker, W.J. (2014, September). *Prevention of Addiction. Using serious games to motivate adolescents to (re)train their brain.*

Presentation at the Annual Conference of the Developmental Section of the British Psychological Society, Amsterdam, NL.

Boendermaker, W.J. (2014, March). *Preventie van Verslaving: Het gebruik van Serious Games om jongeren te motiveren hun brein te (her)trainen.* **Presentation** at Tactus Prevention in Deventer, NL.

Boendermaker, W.J. (2014, January). *Preventie van Verslaving: Het gebruik van Serious Games om jongeren te motiveren hun brein te (her)trainen.* **Presentation** at Arkin/Jellinek, Amsterdam, NL.

Boendermaker, W.J. (2013, December). *Prevention of Addiction. Using serious games to motivate adolescents to (re)train their brain.* **Presentation** at NHL University of Applied Sciences, Leeuwarden, NL.

Boendermaker, W.J. (2013, November). *Preventie van Verslaving. Hoe een serious game het trainen van executieve functies bij adolescenten kan ondersteunen.* **Presentation** and **organized symposium** at VGCt Najaarscongres, Veldhoven, NL.

Boendermaker, W.J. (2013, November). *Preventie van Verslaving bij Jongeren Met Serious Games.* **Presentation** at landelijk hoofdenoverleg Verslavingspreventie Nederland, Amersfoort, NL.

Boendermaker, W.J., Gladwin, T.E., & Wiers, R.W. (2013, May). *Prevention in addiction - Using a serious game to train working memory in adolescents.* **Poster** presented at KNAW Conference on Brain and Cognition: Societal Innovation: How cognitive neuroscience can guide therapy, Amsterdam, NL.

Boendermaker, W.J., Gladwin, T.E., & Wiers, R.W. (2013, March). *Prevention in addiction - Using a serious game to train working memory in adolescents.* **Poster** presented at the Entertainment Software and Cognitive Neurotherapeutics Society (ESCONS) Conference, Los Angeles, CA (USA).

Boendermaker, W.J. (2012, February). *Impliciete Trainingen & Serious Games.* **Presentation** during the 'Dag van de Psychoanalyse' Conference at the Free University (VU), Amsterdam, NL.

Boendermaker, W.J., Prins, P.J.M., & Wiers, R.W. (2011, December). *Play hard, learn hard! Serious gaming om alcohol- en drugsproblemen te voorkomen bij adolescenten.* **Poster** presented at HCMI Conference Cognitie en Psychopathologie: Voorspellen en Veranderen, Utrecht, NL.

GAMIFICATION

SERIOUS GAMIFICATION
MOTIVATING ADOLESCENTS TO DO COGNITIVE TRAINING

WOUTER J. BOENDERMAKER

Adolescence is a developmental period in which an increase in risk taking behavior, such as experimenting with large amounts of alcohol, is common. Although this kind of behavior does not always lead to problematic behavior, excessive use of alcohol at this age can cause significant health problems and school drop-out, and it increases the chance of developing **addiction**-related problems later in life. It is therefore important to take action at an early stage, to prevent escalation. Previous research has shown that **training** certain cognitive control processes, such as working memory and inhibition, can be effective means to help adolescents get more grip on their alcohol use. Additionally, a number of relatively automatic processes, such as selective attention for alcohol, can be re-trained, away from the alcohol. Although both of these forms of computerized training have been shown to be effective, they are also often seen as long and boring. **Serious games** may provide a way to motivate adolescents to complete these cognitive training tasks. This PhD thesis describes a framework for the development of such serious games, with several examples of game-training tasks that were developed. It can be concluded that serious games can indeed play an important, motivating role in cognitive training of adolescents, although the specific circumstances under which this works optimally will need further study.