

Impact of a pre-test measurement of alcohol craving in cue-exposure studies: relationship with social desirability and demand effects

Michelle Heck^a, Etienne Quertemont^a, Jessica Simon^a

^a Psychology & Neuroscience of Cognition – PsyNCog, ULiège, Place des orateurs 2, 4000 Liège,

Belgium

m.heck@uliege.be

equertemont@uliege.be

j.simon@uliege.be

Corresponding author: Michelle Heck, m.heck@uliege.be, Place des orateurs 2, 4000 Liège

Abstract

Craving is one of the most studied concepts in the field of addiction. It is often investigated with repeated measure experimental designs using self-reported scales. However, the explicit nature of self-reported craving scales may make them vulnerable to social desirability and demand effects. The aim of the present study was to test whether a pre-experimental measurement of craving affects its post-experimental assessment after an alcohol video exposure and whether these changes relate to social desirability, demand effects and alcohol consumption. Seventy-five healthy volunteers aged 18-30 years were randomly assigned to two experimental groups: a pre-post and a post-only craving assessment group. They were invited to watch an alcohol-related video. Social desirability, demand effects, engagement towards the video and severity of alcohol consumption were assessed in all participants. The results showed a significant effect of alcohol consumption ($p < .001$, $\eta^2 p = .09$) on post-experimental craving. The main effect of the repeated measure of craving was also significant ($p = .038$, $\eta^2 p = .001$), together with the interaction between

these two variables ($p=.03$, $\eta^2_p=.06$). The pre-experimental craving measurement increased its post-experimental levels, but only in heavy drinkers. However, no evidence was found that these changes were related to social desirability or demand effects. Additional exploratory analyses indicated that craving increase in high alcohol drinkers was mediated by a stronger engagement towards the alcohol-related video.

In summary, the repeated measurement of craving with explicit scales has a significant impact on the results of alcohol craving studies and may not always be desirable, especially when the true purpose of the study needs to be hidden from the participants. However, the present results also suggest that the pre-post experimental design is advisable when the experimenters seek to maximize the relationship between the individual levels of alcohol consumption and alcohol cue-exposure craving changes.

Keywords: Alcohol craving assessment – Engagement– Social Desirability – Demand effects – Reactivity

Introduction

Alcohol craving is defined as “*a strong desire or urge to drink*” (American Psychiatric Association, 2013, p. 491). It is considered as a key process in the development, maintenance, and relapse of alcohol use disorders (AUD) (Bordnick et al., 2008; van Lier et al., 2018; Robinson & Berridge, 1993; Schacht, Anton, & Myrick, 2013; Tiffany, 1990; Tiffany & Wray, 2012) and one of the most studied concepts in the field of addiction (Addolorato, Leggio, Abenavoli, Gasbarrini, & Alcoholism Treatment Study Group, 2005; Tiffany & Wray, 2012). Craving is a complex state that has long been a source of debate regarding its definition, its role in addiction, and its assessment (Anton, 1999; Kavanagh et al., 2013; Sayette et al., 2000; Tiffany & Wray, 2012). Because of its clinical relevance, craving has been extensively investigated in observational and experimental studies. However, the induction and the assessment of craving remain challenging, especially in laboratory studies due to many possible experimental biases (Kavanagh et al., 2013).

Regarding the assessment of craving, various tools were developed for clinical and research settings. Broadly, they can be classified in two major categories: implicit measures that indirectly assess craving through behavior, and self-reported measures. Self-reported measures are by far the most popular tools for assessing craving (Kavanagh et al., 2013), while implicit measures remain relatively controversial (for a review see Ooteman et al., 2006; Sayette et al., 2000). Although very practical, cost-effective, and widely used in research and clinical practice, self-reported craving is confronted with a set of possible confounding variables and biases. In particular, the meaning given to specific ratings could vary between individuals and / or over time (Kavanagh et al., 2013). To avoid this problem, the items of craving scales are generally

made as unambiguous as possible. However, the explicit nature of self-reported craving scales in turn increases the risks of other experimental biases, such as reactivity-related biases.

The term “*reactivity*” in that context characterizes “*the phenomenon that occurs when individuals alter their behavior because of the awareness of being studied*” (Jimenez-Buedo & Guala, 2016, p. 10; see also Adair, 1984). According to Jiménez-Buedo (Jiménez-Buedo, 2021), several terms are used to qualify this reactivity depending on the context, such as “Hawthorne effects”, “placebo effects”, “demand effects”, “methodological artifacts”, or “social desirability biases”. Study reactivity therefore includes all these concepts and especially the notion of “demand effects”. In behavioral sciences, it has long been acknowledged that the subjects of experiments are not passive responders to stimuli (Orne, 1962). They usually want to make a useful contribution based on their knowledge or guessing of the purpose of the study. Indeed, it was shown that the participants’ knowledge of the experimental hypothesis significantly impacts their responses, and often leads them to behave in a way designed to confirm this hypothesis (Nichols & Maner, 2008). Accordingly, all explicit or implicit information about the experimental hypothesis may become significant determinants of the participant’s behavior. The combination of these cues about the tested hypothesis is referred to as “demand characteristics” in the sense that experimental subjects use those cues to determine what the experimenter is “demanding” of them. In turn, the consequences on the participants’ behavior are termed “demand effects”. Three behavior profiles have been described according in this context: the “good-” (Orne, 1962), the “negativistic-” (Masling, 1966), and the “faithful” subjects (Fillenbaum, 1966). The first refers to subjects motivated to shape their behavior in order to confirm what they believe is the experimental hypothesis. In contrast, negativistic subjects are motivated to behave in the opposite way and seek to contradict the researcher’s expectations. Finally, faithful subjects seek to behave

truthfully and call upon their personal subjective experiences to best answer the questions (Herbert, 2007). The “Response Behavior Detection Questionnaire” (Herbert, 2007) was developed to assess the personal sensitivity to the demand characteristics. Because of the usually explicit nature of craving assessment and because participants can easily guess that researchers expect alcohol craving to increase after alcohol-cue exposure, demand effects might be a serious concern in alcohol craving experimental studies. However, to our knowledge, this has never been studied so far.

Another factor known to disturb self-reported measures is social desirability bias. It is defined as *“the tendency to give responses that make the respondent look ‘good’”* (Paulhus, 1991, p. 17). It includes two dimensions: impression management (IM), which is the tendency of people to intentionally make themselves look better, and self-deceptive enhancement (SDE), which is the tendency of people to have an overly positive but nonetheless sincere image of themselves (Paulhus, 1998). Because of the potentially sensitive nature of alcohol craving for some study participants, its assessment with self-reported scales might be significantly affected by social desirability. To our knowledge, only two studies have investigated this phenomenon in relation to craving. While Rohsenow and colleagues (1992) did not find social desirability to impact self-reported craving in alcoholics, Marissen et al. (2006) found a significant negative correlation between social desirability and self-reported craving in detoxified heroin users. Socially desirable responding could also differ according to alcohol consumption patterns. This has been suggested by Rohsenow (1992), who found lower levels of social desirability among subjects with higher alcohol dependence scores. Therefore, the potential impact of social desirability on self-reported craving clearly requires further studies.

Finally, the repeated use of alcohol craving scales might also affect its experimental assessment. Most experimental studies use pretest-posttest designs, in which alcohol craving is assessed twice, at baseline and after the experimental manipulation. Typically, in these kinds of experimental designs, change scores are calculated, or baseline craving scores are used as covariates in the analyses (Sayette et al., 2000). These strategies aim to minimize the impact of inter-individual variance in the responses to the measurement scale. However, the fact that the participants are confronted twice to the same questionnaire in a relatively short time interval could affect the quality of the measurement, and that in at least two ways. On one hand, memory effects could significantly affect the measure. If participants are asked the same questions twice, the fact that they remember their first answer (at pretest) may affect the answer given on posttest (Sayette et al., 2000; Tourangeau & Rasinski, 1988). On the other hand, it is well established that self-monitoring alters the monitored behavior (Perlmutter, Noblin, & Hakami, 1983). Hence, the pretest craving measurement could create or modify the conscious experience of craving and therefore affect the results of the experiment (Simon, Etienne, Bouchard, & Quertemont, 2020). Asking participants to report their initial level of craving could force them to probe an internal state not readily accessible to consciousness at that precise moment. This monitoring could incidentally activate alcohol-related thoughts, therefore making the concept of alcohol craving more salient (priming effect). This salience could either serve as an alcohol cue thereby inducing craving (and perhaps triggering a set of conditioned physiological responses, as in traditional cue-exposure designs, (Sinha et al., 2009)), or could activate attempts of cognitive control that would result in an underestimation of craving (Kavanagh et al., 2013). Furthermore, such effects of pretest craving measurement could interact with the personal characteristics of the participants and especially their patterns of alcohol consumption. Indeed, heavy and occasional alcohol

drinkers differ in terms of alcohol cue reactivity and craving responses to such cues (Herrmann et al., 2001; Ryan et al., 2010; Simon et al., 2020).

Additionally, the pretest measurement of craving is part of the demand characteristics of the study, therefore cueing subjects about the experimental hypothesis. Accordingly, a significant interaction might be expected between demand effects and the presence of a pretest measurement of craving. Although highly relevant for craving research, the effects of the pretest measurement on post-test self-reported alcohol craving have never been explicitly tested.

In laboratory studies, other significant factors are related to the craving induction techniques and the media used to expose participants to craving-induced stimuli. Some studies using virtual reality for instance have shown positive correlations between the sense of presence (sense of being in the virtual environment) and emotional responses such as anxiety or craving (Bouchard, St-Jacques, Robillard, & Renaud, 2008; Della Libera, Simon, Larøi, Quertemont, & Wagener, 2023; Felnhofer, Hlavacs, Beutl, Kryspin-Exner, & Kothgassner, 2019; Ferrer-García, García-Rodríguez, Gutiérrez-Maldonado, Pericot-Valverde, & Secades-Villa, 2010; Price & Anderson, 2007). In a previous study (Simon et al., 2020), we found that ecological validity (a dimension of the sense of presence) moderated the relationship between alcohol consumption level (occasional vs heavy drinkers) and post-assessment craving following a virtual reality immersion. In the context of a 2D media exposure, variables such as “engagement” (referring to the feeling of being psychologically involved in the video) is also important. Indeed, participants could watch the video (i.e., the alcohol-related stimuli) in a completely passive way or in a more "engaged" one, projecting themselves emotionally into the scenario, potentially resulting in different levels of emotional responses and craving. As a follow up to previous results (Simon et al., 2020), a measure of engagement was therefore included in the present study.

The main goal of the present study was to test whether demand effects, social desirability and pretest baseline measurement of craving affect the post-test self-reported craving after cue exposure through a video showing alcohol drinking situations. Prior to watching this alcohol video, participants were randomly divided into two experimental groups. A pretest measurement of self-reported alcohol craving was carried out in the first group (“pre-post” group) but not in the second group (“post-only group”). It was expected that this pretest measurement would significantly affect the post-test levels of self-reported craving and that this effect would be dependent upon the levels of alcohol drinking in the participants. A significant statistical interaction was therefore hypothesized between the experimental group and the levels of alcohol drinking. Furthermore, “demand effects” components, as well as the levels of social desirability, were assessed in all participants through questionnaires. Significant relationships were anticipated between the levels of self-reported post-test craving and these variables, especially in the “pre-post” group. Indeed, the participants of this latter group should more easily guess that increases in alcohol craving are expected following the experimental manipulation. In other terms, a significant statistical interaction was hypothesized between the experimental group and these variables.

Finally, in an exploratory perspective, the present study aimed to test whether engagement towards the alcohol-related video mediates or moderates the relationship between the severity of alcohol consumption and post-test craving. Engagement was considered as “*a psychological state experienced as a consequence of focusing one’s energy and attention on a coherent set of stimuli or meaningfully related activities and events*” (Witmer & Singer, 1998, p. 227). According to Witmer and Singer (1998), involvement depends on the degree of importance or significance that the individual attributes to stimuli or events. In the present study, we expect high alcohol drinkers

to be more engaged towards the alcohol-related video clip than low alcohol drinkers. In turn, this higher level of engagement might contribute to increase post-test alcohol craving.

Materials and methods

Participants

Following recruitment via an online screening questionnaire, 93 French-speaking non-clinical social drinkers aged between 18 and 30 years (inclusion criteria) took part in the whole experiment. Exclusion criteria were based on the screening questionnaires and were defined as: 1) full abstinence from alcohol consumption, 2) a history of neurological problems that could affect cognitive functioning, 3) regular use of tranquilizers or illicit drugs, which was defined as using one of these substances more than twice in the last two months, except for cannabis (more than twice a month). Tobacco use was not an exclusion criterion.

Eighteen participants were excluded post-data collection, for various reasons: absence of alcohol consumption (AUDIT=0, N=7), unserious answers to the questionnaires (N=1), total abstinence after heavy drinking (N=1), and non-compliance with the instructions of the experiment (N=9), including 3 participants who did not watch the video and 6 participants who failed to comply with the instructed time of the day for carrying the experiment (start after 5 pm, 30 min difference was accepted). Finally, 75 participants were randomly assigned to one of the two experimental groups: *pre-post group* vs. *post-only group*. The random assignment was done by the internet platform that hosted the survey (50% chance to get one survey link or the other). Participant's demographic, psychological and alcohol use characteristics are summarized in table 1. No financial compensation was given for participation in the study. There were no missing

data, since answering all questions on each page was mandatory on the online platform to move to the next page.

The sample size was calculated using an a priori power analysis using WebPower (Zhang & Yuan, 2018) based on the main hypothesis of the present study, i.e., a significant interaction between the experimental group and the levels of alcohol drinking measured through the AUDIT-C. The following parameters were used for the calculation: $\alpha=.05$, power=.80 and $f^2=.15$. As a similar interaction was never tested before, it was difficult to figure the effect size for the calculation. However, from previous studies in our laboratory, craving differences between heavy and light drinkers were usually of medium or large effect sizes (Simon et al., 2020) and, in a precautionary manner, we therefore adopted a medium effect size as the minimal detectable effect size for the calculation. The result of the analysis indicated that seventy-seven participants were needed to reach a .80 power.

Questionnaires

Alcohol consumption. The French version (Gache et al., 2005) of the *Alcohol Use Disorders Identification Test* (AUDIT (Saunders, Aasland, Babor, de la Fuente, & Grant, 1993)) consists of 10 multiple-choice items measuring alcohol consumption, alcohol dependence and alcohol-related problems. Higher scores reflect a high probability of harmful alcohol use.

As the present study was carried out in the middle of the COVID-19 pandemic, the short version of this test, the AUDIT-C, was preferred. Indeed, this version focuses on the recent alcohol drinking behavior and is effective in detecting "at risk" users in a general population (Aalto, Alho, Halme, & Seppä, 2009). Due to the various lockdowns and restrictions related to the pandemic, the social components of alcohol consumption were generally reduced, while at home

drinking increased (Bollen et al., 2021). Therefore, several items of the full AUDIT might not properly assess alcohol-related risk in the specific circumstances of the COVID-19 pandemic (for example, items such as "*Have you or someone else been injured as a result of your drinking?*" or "*Has a relative or friend, doctor or other health worker been concerned about your drinking or suggested you cut down?*")

Questions about current consumption. To gain insight into participants' recent drinking behavior, we included the following two questions into the protocol: "*How many drinks did you have in the past 7 days?*" and "*On how many occasions did you drink alcohol in the past 7 days?*"

Craving. Alcohol craving was assessed via four visual analogue scales (VAS; Kreusch, Billieux, & Quertemont, 2017, adapted from the ACQ, (Singleton, Tiffany, & Henningfield, 1994)) ranging from 0 to 100. VAS evaluated (1) the expectancy of positive reinforcements (i.e., "drinking a glass of alcohol would have made things just perfect"), (2) craving intensity (i.e. "the intensity of my urge to drink alcohol is very strong right now"), (3) the willingness to consume (i.e., "if I could drink a glass of alcohol right now, I would"), and (4) the lack of control (i.e., "it would be hard to turn down a drink right now"). A global craving score was calculated by summing the 4 VAS scores.

Substance use. This questionnaire assessed the frequency of use of cannabis, cocaine, ecstasy, hallucinogens, amphetamines, opiates, and volatile solvents in the past two months on a 4-point Likert scale (never - once/twice - more than twice - regular use).

Immersive tendencies in media. The French version (Robillard, Bouchard, Renaud, & Cournoyer, 2002) of the Immersive Tendencies Questionnaire (ITQ; (Witmer & Singer, 1998b)) assesses the tendency to shut out external distractions in order to focus on different tasks in daily

life. It contains 18 items where participants rate their level of agreement on a 7-point scale. Four dimensions are derived: focus, involvement, emotion, and game.

Engagement. A French version of the short User Engagement Scale (UES, (O'Brien, Cairns, & Hall, 2018) was used to measure participants' engagement while watching the video. The questionnaire is scored on a 5-point Likert scale (ranging from "strongly disagree" to "strongly agree"). This questionnaire was originally developed to measure user engagement in interactive systems such as video games. It is composed of four dimensions: focused attention (i.e., feeling absorbed by the interaction and losing track of time), aesthetics (i.e., the attractiveness and visual appeal of the interface), reward (i.e., a dimension that includes the concepts of novelty, curiosity and interest in the interactive task as well as perceived involvement (the feeling of being "absorbed" and having fun) and "endurability" (overall success of the interaction and the willingness of users to recommend the application to others)) and perceived usability (i.e., the negative affect felt as a result of the interaction and the degree of control and effort deployed). As no French version was available, we used a translated version consisting of 22 items and including 3 of the 4 dimensions initially considered (focused attention, aesthetics, and reward). The perceived usability dimension was not relevant in the present context since the task consisted of watching a video. Therefore, there was no "ergonomics" component. The translated version of the scale presented a rather mediocre factor structure ($\chi^2/DF=1.34$; RMSEA=.07; SRMR=.09). After removal of seven bad items (based on their R^2 and relevance), a satisfying factor structure was obtained ($\chi^2/DF=1.73$; RMSEA=.099; SRMR=.08). Internal consistency of the 3 scales were good (focused attention: $\omega=.92$; aesthetics: $\omega=.89$; reward: $\omega=.86$). The translated version of the scale is available in the supplementary data (*supplemental 4*).

Social Desirability. The DS36 social desirability scale (Tournois, Mesnil, & Kop, 2000) is a self-reported scale that measures two components of social desirability: self-deceptive enhancement (i.e., the act of deceiving oneself in good faith) and impression management (i.e., the act of consciously deceiving others). It consists of a 7-modality Likert-type scale (from "completely false" to "completely true"). Higher scores indicate higher social desirability.

Response Behaviors. The Response Behavior Detection Questionnaire (RBDQ, (Herbert, 2007)) measures the sensitivity to demand characteristics through three major scales: (1) a sensitivity to demand characteristics scale (i.e., *an expression of interest on behalf of the respondent, seeking to make sense of the questioning situation with which they are confronted*"), (2) a cue interpretation scale (i.e., *the analysis and interpretation of these cues by the participants to produce a set of conjectures about the purpose of the research*"), and (3) a response motivation scale (i.e., *to the willingness of the individual being questioned to rely on their interpretation of the demand characteristics in order to engage in specific response behaviors that may be "negativistic", "faithful", or "good"*). The first scale consists of three 11-point Likert items. The second scale is made of 2 items, an 8-point multiple-choice question that asks subjects about their level of recognition of the study's objectives (ranging from "I didn't wonder" to "I'm sure I guessed correctly what it was about"), and one open question related to the recognition of the experimental hypotheses (which we coded 0 = "absence of or incorrect hypothesis", 1 = "hypothesis related to alcohol", 2 = "hypothesis related to craving for alcohol", 3 = "hypothesis related to the biases that might affect the measurement [of alcohol craving]"). This measure was used as an ordinal scale because of the degree of graduation in the elaboration on hypotheses. The third scale (response motivation scale) is made of six 11-point Likert items, with smaller values indicating "negativistic" responding (i.e., motivation to behave in the opposite way and to

contradict the researcher's expectations) and higher scores indicating “good” responding” (i.e., the motivation to behave in a way that confirms the researcher's expectations).

Anxiety. The Trait scale of the State-Trait Anxiety Inventory (STAI, Spielberger et al., 1993), assesses trait anxiety through 20 items. Higher scores indicate greater anxiety.

Additionally, several questionnaires were included to make the fake study title believable and to mask the real purpose of the study: The Karolinska Sleepiness Scale (KSS, (Kaida et al., 2006)), the Center for Epidemiologic Studies – Depression scale (CES-D; (Radloff, 1977)), the Rosenberg's Self-Esteem Scale (RSES, Rosenberg, 1965), the short version of the French Impulsive Behavior Scale (UPPS-P (Billieux et al., 2012)), and the short version of the Big Five Inventory (BFI-10, (Rammstedt & John, 2007); French translation Courtois et al., 2020)).

Video material

In order to induce craving, a short video clip showing various drinking situations was used. It featured people drinking alcoholic beverages in restaurants, bartenders preparing cocktails, people doing “cheers”, etc. (Figure 1). The video sequences were accompanied by background music. The video lasted a total of 4 minutes and 26 seconds.

Figure 1: Screenshots of the video

Design and procedures

This study was approved by the Ethics Committee of the Faculty of Psychology, Speech and Language Therapy and Education of the University and was conducted in accordance with the ethical standards described in the Declaration of Helsinki (1964).

Participants were recruited by "word-of-mouth", via social networks and the internal platform of the University but also by poster advertising. The ad informed the participants that the experiment was intended to evaluate the impact of personality traits on the feeling of presence. People interested in the study were invited to fill in an online screening survey (socio-demographic and consumption data, AUDIT-C, ITQ, DS36). Recruitment and testing took place from January 2021 to July 2021.

Participants who met the inclusion criteria were invited by e-mail to complete an online survey that lasted approximately 35 minutes. Among the instructions mentioned in this e-mail, participants were asked to start the experiment after 5 p.m., in order to ensure that all participants were in relatively the same drinking mood (i.e., few people feel an urge to drink alcohol in the morning). After reading the information form and giving informed consent, participants were asked to complete several pre-experimental questionnaires (UPPS-P, STAI-T, KSS, Craving-VAS, EES, BFI-10). All participants had to answer the same pre-experimental questionnaires, except for the pretest alcohol craving scales. Whereas baseline alcohol craving was assessed before the video in the pre-post group, no questions about alcohol were asked to the post-only group in the first part of the experiment.

The second phase of the study then involved watching a video containing various alcohol-related cues on the participants' personal computer at home. To ensure the proper functioning of their equipment, a five-second "test" video (sunset) preceded the experimental video. At this stage, the participants were asked to adjust the sound and to put the video in full screen. For the experimental video, the viewing instructions were as follows: *"You are now going to see several video sequences taken from commercials. Watch them (with the volume turned up) and do not pause the video. Watch the video carefully and soak up the atmosphere as best as you can."*

Finally, all participants were asked to complete a series of post-experimental questionnaires (KSS, Craving, UES, CES-D, AUDIT, questions about drinking in the past 7 days, RBDQ).

The real purpose of the study was then revealed to the participants before asking them to sign a second informed consent.

Statistical analysis

Statistical analyses were performed using Jamovi (The jamovi project, 2021) and R (R Core Team, 2021). All the analyses set the alpha level for statistical significance at .05.

Demographics and manipulation check

Differences between the two experimental groups were first tested on the main demographic (gender, age, educational level), psychological (social desirability, immersive tendencies, engagement during video) and alcohol use (AUDIT-C, number of alcoholic drinks consumed during the last week) variables using Pearson's chi-square test for independence and t-tests for independent groups (table 1). In case of severe violation of the homogeneity of variance verified by Levene's test, the Satterthwaite's approximation was preferred. In case of severe violation of the normality condition, Mann-Whitney U test was preferred. In order to confirm that the alcohol-related video elicits craving, a Student's t test was performed to compare pre-experimental and post-experimental craving scores in the pre-post group. Cohen's d, Cramer's V, and Rank biserial correlation were used as measure of effect size.

Primary hypotheses

To test for the effects of the experimental condition and alcohol consumption on posttest craving, a multiple regression analysis was computed with the craving score as the dependent variable and the AUDIT-C score, experimental condition (*pre-post vs post-only*), and their interaction as

explanatory variables. Partial η^2 was used as a measure of effect size. Moreover, Kendall's correlations were calculated between the post-experimental craving and AUDIT-C scores, separately for each group. The correlation coefficients were compared using the Fisher r -to- z transformation and discussed according to the effect size of the q -index derived according to Cohen's classification (Cohen, 1988). A multiple regression analysis was computed to determine whether demand characteristics and experimental condition could jointly explain the self-reported craving. This regression included 6 explanatory variables: Experimental group, Sensitivity to demand characteristics, Cue interpretation, Response behavior, and the interactions between experimental group and the 3 processes of demand effects.

To test the effect of SDE and IM sub-dimensions of social desirability in conjunction to experimental condition on craving, non-parametric correlations were preferred instead of linear models, since the condition of normality of residuals was not met.

A Mann-Whitney U test was used to determine whether the two experimental groups differed in terms of their elaboration on experimental hypotheses. Rank biserial correlation was used as a measure of effect size.

Exploratory hypotheses

Finally, mediation and moderation analysis were performed to explore the impact of the dimensions of engagement in the video on the relationship between AUDIT-C and self-reported craving. In a moderation analysis, the effect of the moderator is statistically characterized as a significant interaction affecting the direction and/or the magnitude of the relationship between the dependent and independent variables. Therefore, we tested a linear hierarchical regression, including in the first block the AUDIT-C and the dimension of engagement. In the second block,

the interaction term between the AUDIT-C and the dimension of engagement was added.

Similarly, a mediation model was performed to test whether engagement mediates the relationship between AUDIT-C and the self-reported craving. Since the associations between engagement and consumption level (AUDIT-C) and between craving and consumption level turned out to be very different in the two experimental groups (pre-post vs post-only), we performed moderation analyses separately for the two groups on one hand, and a moderated mediation analysis on the other.

Finally, to ensure the significance of the tested effects, we performed 5000 bootstraps.

Results

Descriptive data and manipulation check

Groups did not significantly differ on the main socio-demographic, psychological and alcohol use variables (Table 1).

Table 1

The Student's t test in the "pre-post" group showed a significant increase in self-reported craving following the alcohol-related video ($t(35)=-3.27$, $p=.002$, $d = -.55$) (Figure 2).

Figure 2

Primary hypotheses

Effects of the pretest measurement of craving

Table 2 reports the results of the multiple regression analysis that was computed with the post-test craving score as the dependent variable and the AUDIT-C score and experimental group as explanatory variables. The regression model accounted for 14.87% of the total variance in self-reported craving [$F(3,71)=4.14$; $p=.009$, adjusted $R^2=.11$]. As shown on figure 3, globally,

participants with higher scores on the AUDIT-C had higher levels of post-test craving. However, the relationship between AUDIT-C and craving was significantly different between the two groups with a statistically significant relationship only in the pre-post group.

Accordingly, the correlation between AUDIT-C and craving was statistically significant in the “pre-post” group ($r=.55$, $p<.001$) but not in the “post only” group ($r=.02$, $p=.90$). The q index indicated that the difference between the correlations is of large size ($q = .60$).

Table 2

Figure 3

Demand Effects

Table 3 reports the results of the multiple regression analysis that was computed to test the effect of demand characteristics on post-experimental craving. The linear regression with the post-craving score as the dependent variable and the three demand effect scores and experimental group as explanatory variables did not show any significant effect (Model: $F(7,67) = .54$, $p=.80$, $R^2=.05$, $\text{adj } R^2 = -.05$). Neither the simple effects nor the interaction between demand effect processes and experimental condition had a significant impact on self-reported craving (Table 3).

Table 3

The Mann Whitney U test on experimental hypothesis discovery from the participants indicated no statistically significant difference between the two groups (*pre-post vs post-only*) in terms of elaboration on experimental hypotheses ($w=538$, $p=.06$, $r_{tb} = .23$). Table 4 illustrates the degree of recognition of the experiment's hypotheses among the different experimental groups.

Table 4

Social Desirability Bias

Kendall's correlations between posttest craving and the two dimensions of social desirability were computed in both groups. No significant correlation was found in either group (Pre-post group:

Impression management $\tau=.03$, $p=.81$; self-deceptive enhancement $\tau=.16$, $p=.17$; Post-only group: Impression management $\tau=-.15$, $p=.19$; self-deceptive enhancement $\tau=-.11$, $p=.33$).

Exploratory hypotheses

Role of engagement in the relationship between alcohol consumption and craving

When testing the moderator effect of engagement on the relationship between alcohol consumption level (AUDIT-C) and self-reported craving in the two groups, no significant moderating effect was obtained (Table 5).

Table 5.

A moderated mediation model was tested to determine whether engagement mediated the relationship between alcohol consumption and self-reported craving, and whether the experimental group moderated these relationships. Since the reward dimension of engagement was the most strongly associated with post-exposure craving ($r=.44$, $p<.001$), the model was tested with this dimension.

In the context of a simple mediation effect of engagement (reward dimension) on the relationship between alcohol consumption (AUDIT-C) and craving, the indirect effect of consumption level (AUDIT-C) was significant ($z=2.23$, $p=.02$) (5000 bootstraps).

Moreover, in this model, the direct effect of AUDIT-C was no longer significant ($z=1.29$, $p=.19$) (Figure 4), indicating a complete mediation effect. When taking into account the moderating effect of the experimental group (pre-post vs. post-only) on each path of this model, the effect of AUDIT-C on engagement was moderated by the experimental group (i.e., whether or not participants were exposed to a pre-experimental craving measure) (Figure 5). More specifically, in the pre-post group, the Average Causal Mediation Effect (ACME, [total effect-direct effect]) was significant (estimate= 9.51, 95% CI = [1.14, 23.20], $p=.02$ (using Nonparametric Bootstrap

Confidence Intervals with the Percentile Method, 5000 simulations)). Moreover, the Average Direct Effect (ADE, [total effect-indirect effect]) was non-significant (estimate= 15.12, 95% CI = [-4.57, 33.12], $p=.13$), and the Total Effect was significant (estimate=24.63, 95% CI = [11.62, 38.93], $p=.0008$). The proportion mediated was 39% (95% CI = [.04, 1.28], $p=.02$). In the post-only group however, there was no significant effect (ACME: estimate=1.75, 95% CI =[-4.54, 8.98], $p=.55$; ADE: estimate=-.65, 95% CI=[-15.61, 17.84], $p=.95$; Total Effect: estimate=1.10, 95% CI=[-13.22, 20.07], $p=.89$; proportion mediated: estimate=1.59, 95% CI=[-5.33, 5.13], $p=.88$).

Figure 4.

Figure 5.

To get an integrated idea of the phenomenon at play, correlation matrices were computed for the dimensions of engagement, post experimental craving, and AUDIT-C, separately in the two groups. Results show that almost all engagement dimensions significantly correlate with post experimental craving in both groups. However, while the correlation between the reward dimension and craving is of medium size in the post-only group ($r=.38$, $p=.01$), it is of large size in the pre-post group ($r=.55$, $p<.001$). The correlation between focused attention and craving is non-significant in the pre-post group ($r=.33$, $p=.05$), but significant and of medium size in the post-only group ($r=.42$, $p<.01$). In both groups, the correlations between the aesthetics dimension and craving are significant and of medium size (pre-post: $r=.36$, $p=.02$, post-only: $r=.34$, $p=.03$). Means, standard deviations, and Pearson correlation matrices for the main measures of interest are available in the supplemental data (Supplemental 1, 2 and 3).

Discussion

Summary and interpretation of main findings

The purpose of this study was twofold. Firstly, to take a critical look at the use of pre-post designs to assess craving based on self-reported scales and secondly, to investigate the impact of social desirability and demand effects associated with craving measurement. As expected, alcohol cue exposure through a video increased self-reported craving in the pre-post group (Adams, Rapinda, Frohlich, O'Connor, & Keough, 2019; Kreusch et al., 2017; Ryan et al., 2010). However, the levels of elicited craving were dependent upon the levels of alcohol use. Indeed, the participants with higher AUDIT-C scores were those who reported the stronger levels of alcohol craving, but only in case of a baseline measurement of alcohol craving. Contrary to our hypothesis, we did not find any evidence that self-reported craving is related to demand effects or to social desirability. Finally, the reward dimension of the engagement questionnaire mediated the relationship between alcohol consumption levels and self-reported craving in the pre-post group, but not in the post-only group.

Induction of craving

Our results suggest that the substance-related video significantly increased craving in some participants. Our results also show that a “simple” 2D video may be sufficient to elicit craving. In an effort to more closely match the complexity of real life, researchers are increasingly using virtual reality to induce craving (for a review, see Bordnick and Washburn, 2019; Segawa et al., 2020). While this approach is of great interest in a clinical population, the present results indicate that more simplistic designs such as exposure to visual stimuli in a 2D video could be as effective as more sophisticated techniques in social drinkers.

Assessing craving once or twice

As sometimes discussed in the literature, informing participants about the study objectives, or assessing craving prior to its induction could alter the self-reported alcohol craving levels

(Kavanagh et al., 2013; Sayette et al., 2000; Simon et al., 2020). This is confirmed in the present study in which higher mean levels of post-experimental self-reported craving were recorded when a baseline pre-experimental measure of craving was performed. However, there was also a significant interaction between the level of alcohol consumption and the group (*pre-post* vs. *post-only*). When a baseline pre-experimental measure of craving was performed, i.e., in the *pre-post* group, stronger levels of post-experimental craving were recorded in high drinkers relative to low drinkers, with a medium-size correlation between the measures of AUDIT-C and craving. In contrast, in the *post-only* group, there was no significant relationship between the levels of alcohol craving and the levels of alcohol consumption.

A possible explanation of this effect is that the baseline pre-experimental measure of craving may already act as a salient cue in high alcohol drinkers. Indeed, it has been shown that very subtle cues could already trigger motivation by unconsciously drawing attention on drug consumption (Kavanagh et al., 2013; Roehrich & Goldman, 1995; Stein, Goldman, & Del Boca, 2000). In particular, the Elaborated Intrusion Theory of desires suggests that simply asking participants about their subjective experience of craving may activate spontaneous, intrusive thoughts and therefore act as a salient cue for substance use (Kavanagh, Andrade, & May, 2005). These intrusive thoughts may then potentiate the effects of subsequent alcohol cues by adding a top-down process to the bottom-up effect of the video (“priming” effect).

An alternative explanation of the present results is that the pre-experimental measure sets a reference point for the subsequent post-experimental self-reported craving. Indeed, it is well known that when participants are asked the same question twice (or more), the memory of their first answer affects their subsequent responses (Schwarz, Revilla, & Weber, 2020). Therefore, higher levels of post-experimental craving in heavy drinkers submitted to a pre-experimental

assessment might reflect their subjective experience of an increase in their motivation to drink alcohol after viewing the video. In contrast, low drinkers would not share such a subjective experience. If true, the ‘reference point’ explanation has both positive and negative implications. From a positive point of view, a pre-experimental measure of craving might allow a better assessment of the subjective experience of craving changes. However, on the negative side, the pre-experimental measure of craving might expose the results of the study to additional memory biases. Further studies will be required to investigate these questions.

Impact of social desirability and demand characteristics on self-reported craving

The present results did not support the hypothesis of a significant impact of either demand effects or social desirability biases on craving assessment, although they confirmed that a pre-experimental assessment of craving alters the self-reported post-experimental levels of craving. Since human beings have a natural tendency to ascribe purpose and meaning even in the absence of purpose and meaning, there can be no experiment without demand effects (Orne, 1962). It therefore seems logical that participants with more insights into the real purpose of a study are more prone to adopt specific response behaviors (according to perceived demand characteristics). In the present study, we did not find a relationship between the sensitivity of the participants to demand characteristics as assessed by the Response Behavior Detection Questionnaire and their levels of self-reported craving. This might be explained either by the insufficient sensitivity of the demand effect scale for the present kind of experiment or by the fact that post-experimental craving is affected by a different mechanism. For example, memory and priming effects as evoked above could explain how the post-experimental measurement of craving is affected by its pre-experimental assessment without the involvement of demand effects per se. The same logic might apply to the impact of social desirability. However, it has been suggested that social desirability bias is triggered by two factors: (a) the nature of the questions related to sensitive

and/or highly personal topics and (b) the presence of an interviewer which activates social norms (Nass, Moon, & Carney, 1999). Therefore, given that this study took place online, and tested a non-clinical sample, the context and the sample used could also have reduced the expression of social desirability biases.

The importance of engagement for craving induction

In the present study, the self-reported post-experimental craving was related to the participants' levels of alcohol consumption only when a pre-experimental measurement of craving was included, i.e., in the pre-post experimental group. Interestingly, only in that group, the relationship between the levels of alcohol consumption and post-experimental craving was mediated by the reward dimension of engagement (Figure 5). The reward dimension assessed in the present study is relatively close to the concept of cognitive involvement of Webster and colleagues (Webster & Hackley, 1997; Webster & Ho, 1997), defined as the subjective experience of human–technology interaction including the intrinsic interest, the curiosity and the attention focus dimensions. Engagement is dependent upon the degree of importance or significance that the individual attributes to the depicted stimuli or events (Witmer & Singer, 1998a). Globally, the more individuals pay attention to the stimuli, the more they become "engaged", resulting in higher levels of immersion in a media. Conversely, the more a participant is distracted by personal thoughts or by things irrelevant to the current activity, the less involved they will be (Nicovich, 2005; Witmer & Singer, 1998a). This moderated mediation provides an interesting explanation to the observations that post-experimental craving is affected by the presence of a pre-experimental measurement and that it is related to alcohol consumption levels of the participants only in that case. Indeed, in the pre-post group, the pre-experimental craving measurement informed the participants that the study was about alcohol drinking motivation, probably increasing the interest and immersion of participants in the video, differentially

according to their usual levels of alcohol consumption. In turn, these levels of engagement resulted in higher post-experimental levels of craving in heavy alcohol drinkers. In contrast, in the post-only group, there was no significant relationship between the levels of alcohol consumption of the participants and their engagement in the video, which was determined by other factors. As a result, post-experimental craving was still related to the video engagement but unrelated to the levels of alcohol consumption of the participants. Therefore, engagement in the alcohol-related video appears to be an essential factor to induce alcohol craving and seems to be affected by a pre-experimental measurement of alcohol craving through an explicit self-reported scale (possibly acting as an alcohol cue). However, as the study of the relationships between these variables was exploratory in the present experiment, further investigations will be required to dig deeper into these phenomena.

Limits

It should be noted that the present study had several limitations. The covid-19 health crisis in which this experiment took place might have affected the participants' behaviors. The closure of bars and restaurants and the successive lockdowns have changed people's drinking habits (Garcia & Sanchez, 2020). While some individuals consumed more alcohol, others may have engaged in light or moderate drinking patterns at home resulting in a decrease of consumption (Nicholls & Conroy, 2021). Furthermore, the online nature of the present study involved both advantages and limitations. As the participants carried out the study at home, there was less control over the experimental environment and especially the day on which participants had watched the video, although the time of the day was controlled. For example, it is likely that participants are not in the same state of mind and will not feel the same desire to drink alcohol on a Monday compared to a Friday. In our sample, 40% of the participants completed the study on a "typical student drinking day" (Thursday, Friday, or Saturday), although post-hoc statistical analyses found no

significant differences in the levels of self-reported craving between the "typical drinking days" and the working days ($U=670$, $p=.95$).

Conversely, the fact that participants were not in the sterile context of a research laboratory when assessing their craving, but in an environment potentially closer to their actual drinking environments, may have been an advantage rather than a disadvantage. Indeed, research has shown that perceived substance availability can modulate alcohol cue reactivity (Papachristou et al., 2012). Since substance availability was most likely higher in a home setting than in a laboratory, the online context might have improved the ecological validity of the design. Future studies should therefore compare the impact of online relative to laboratory settings for the investigation of alcohol craving. The use of a "simple" self-reported measure of craving is another possible limitation of the present study. As mentioned in the introduction, craving is a complex multidimensional concept (Ghiță et al., 2019; van Lier et al., 2018; Sayette et al., 2000) that is difficult to capture with such a simple scale. However, as the main aim of the present study was precisely to investigate the experimental biases possibly affecting self-reported craving, the use of such a simple measure of craving was inherent to the purpose of the present experiment. Finally, although the participants were randomly assigned to the two experimental groups, the use of a convenience sample of volunteers is inherent to online studies and is a limitation to the generalization of the present results.

Conclusions

In agreement with Orne's premise (1962), it is important to recognize the subject as an active participant in any experiment. As shown in the present results, a pre-experimental measurement of craving is not trivial and produces several effects on the participants: they react in a different way to the post-experimental assessment of craving. However, no evidence was found that these

changes in post-experimental self-reported craving are related to individual differences in social desirability or respondent biases. Interestingly, the present results suggest that the pre-experimental measurement of craving affects the engagement and interest for the alcohol-related video in a different way for low and high alcohol drinkers, leading to an increase of post-experimental craving in the latter. In conclusion, the repeated measurement of craving with explicit scales may not always be desirable, especially when the true purpose of the study needs to be hidden from the participants. However, the present results also suggest that the pre-post experimental design is advisable if the experimenters seek to maximize the relationship between the individual levels of alcohol consumption and alcohol cue-exposure craving changes.

Role of funding source

This work was supported by the University of Liège. Michelle Heck is a Research Fellow of the Fonds de la Recherche Scientifique – FNRS.

Declaration of interest

None.

Author contributions

EQ: Conceptualization, Writing – Review & Editing, Supervision; **JS:** Methodology, Conceptualization, Writing – Review & Editing, Project administration, Formal analysis; **MH:** Methodology, Investigation, Data curation, Formal Analysis, Writing – Original Draft, Visualization.

Competing Interests

Authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Data availability statement

All data are available at https://osf.io/26dty/?view_only=0336e93e1a014fe8ae7614d4fbc7b68b.

Ethics statement

This study was accepted by the local Ethical Committee of the Faculty of Psychology, Speech Therapy, and Educational Science of the University of Liège and has been carried out in accordance with the Declaration of Helsinki. All authors have approved the final article.

Acknowledgements

Thanks to Tomáš Králik and Josip Boban for their permission to use their videos.

References

- Aalto, M., Alho, H., Halme, J. T., & Seppä, K. (2009). AUDIT and its abbreviated versions in detecting heavy and binge drinking in a general population survey. *Drug and Alcohol Dependence, 103*(1-2), 25-29.
- Adair, J. G. (1984). The Hawthorne effect: A reconsideration of the methodological artifact. *Journal of Applied Psychology, 69*(2), 334-345.
- Adams, T., Rapinda, K. K., Frohlich, J. R., O'Connor, R. M., & Keough, M. T. (2019). Impulsivity moderates the effect of social anxiety on in-lab alcohol craving. *Addictive Behaviors, 97*, 70-76.
- Addolorato, G., Leggio, L., Abenavoli, L., Gasbarrini, G., & Alcoholism Treatment Study Group. (2005). Neurobiochemical and clinical aspects of craving in alcohol addiction: A review. *Addictive Behaviors, 30*(6), 1209-1224.
- American Psychiatric Association. (2013). *Diagnostic and Statistical Manual of Mental Disorders* (Fifth Edition.). American Psychiatric Association.

- Anton, R. F. (1999). What is craving? Models and implications for treatment. *Alcohol Research & Health*, 23(3), 165-173.
- Billieux, J., Rochat, L., Ceschi, G., Carré, A., Offerlin-Meyer, I., Defeldre, A.-C., Khazaal, Y., et al. (2012). Validation of a short French version of the UPPS-P Impulsive Behavior Scale. *Comprehensive Psychiatry*, 53(5), 609-615.
- Bollen, Z., Pabst, A., Creupelandt, C., Fontesse, S., Lannoy, S., Pinon, N., & Maurage, P. (2021). Prior drinking motives predict alcohol consumption during the COVID-19 lockdown : A cross-sectional online survey among Belgian college students. *Addictive Behaviors*, 115, 106772.
- Bordnick, P. S., Traylor, A., Copp, H. L., Graap, K. M., Carter, B., Ferrer, M., & Walton, A. P. (2008). Assessing reactivity to virtual reality alcohol based cues. *Addictive Behaviors*, 33(6), 743-756.
- Bordnick, P. S., & Washburn, M. (2019). Virtual Environments for Substance Abuse Assessment and Treatment. In A. « Skip » Rizzo & S. Bouchard (Éds.), *Virtual Reality for Psychological and Neurocognitive Interventions*, Virtual Reality Technologies for Health and Clinical Applications (p. 131-161). New York, NY: Springer New York.
- Bouchard, S., St-Jacques, J., Robillard, G., & Renaud, P. (2008). Anxiety Increases the Feeling of Presence in Virtual Reality. *Presence: Teleoperators and Virtual Environments*, 17(4), 376-391.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.
- Courtois, R., Petot, J.-M., Plaisant, O., Allibe, B., Lignier, B., Réveillère, C., Lecocq, G., et al. (2020). Validation française du Big Five Inventory à 10 items (BFI-10). *L'Encéphale*, 46(6), 455-462.
- Della Libera, C., Simon, J., Larøi, F., Quertemont, E., & Wagener, A. (2023). Using 360-degree immersive videos to assess multiple transdiagnostic symptoms: A study focusing on fear of negative evaluation, paranoid thoughts, negative automatic thoughts, and craving. *Virtual Reality*, 1-16

- Felinhofer, A., Hlavacs, H., Beutl, L., Kryspin-Exner, I., & Kothgassner, O. D. (2019). Physical Presence, Social Presence, and Anxiety in Participants with Social Anxiety Disorder During Virtual Cue Exposure. *Cyberpsychology, Behavior, and Social Networking*, 22(1), 46-50. Mary Ann Liebert, Inc., publishers.
- Ferrer-García, M., García-Rodríguez, O., Gutiérrez-Maldonado, J., Pericot-Valverde, I., & Secades-Villa, R. (2010). Efficacy of virtual reality in triggering the craving to smoke : Its relation to level of presence and nicotine dependence. *Annual Review of CyberTherapy and Telemedicine*, 8(1), 99-102.
- Fillenbaum, S. (1966). Prior deception and subsequent experimental performance : The « faithful » subject. *Journal of Personality and Social Psychology*, 4(5), 532-537. US: American Psychological Association.
- Gache, P., Michaud, P., Landry, U., Accietto, C., Arfaoui, S., Wenger, O., & Daeppen, J.-B. (2005). The Alcohol Use Disorders Identification Test (AUDIT) as a screening tool for excessive drinking in primary care : Reliability and validity of a French version. *Alcoholism, Clinical and Experimental Research*, 29(11), 2001-2007.
- Garcia, L. P., & Sanchez, Z. M. (2020). Consumo de álcool durante a pandemia da COVID-19 : Uma reflexão necessária para o enfrentamento da situação. *Cadernos de Saúde Pública*, 36(10), e00124520.
- Ghiță, A., Teixidor, L., Monras, M., Ortega, L., Mondon, S., Gual, A., Paredes, S. M., et al. (2019). Identifying triggers of alcohol craving to develop effective virtual environments for cue exposure therapy. *Frontiers in Psychology*, 10(JAN).
- Herbert, M. (2007). Que se passe-t-il lorsque les répondants à un questionnaire tentent de deviner l'objectif de recherche ? Le biais du répondant : Conceptualisation, mesure et étude d'impact. Présenté à Congrès International de l'AFM, Aix-les-Bains (France).

- Herrmann, M. J., Weijers, H. G., Wiesbeck, G. A., Böning, J., & Fallgatter, A. J. (2001). Alcohol cue-reactivity in heavy and light social drinkers as revealed by event-related potentials. *Alcohol and Alcoholism*, 36(6), 588–593.
- The Jamovi Project (2021). jamovi (Version 1.6) [Computer Software]. Retrieved from <https://www.jamovi.org>
- Jiménez-Buedo, M. (2021). Reactivity in social scientific experiments : What is it and how is it different (and worse) than a Placebo effect? *European Journal for Philosophy of Science*, 11(2), 42.
- Jimenez-Buedo, M., & Guala, F. (2016). Artificiality, Reactivity, and Demand Effects in Experimental Economics. *Philosophy of the Social Sciences*, 46(1), 3-23.
- Kaida, K., Takahashi, M., Akerstedt, T., Nakata, A., Otsuka, Y., Haratani, T., & Fukasawa, K. (2006). Validation of the Karolinska sleepiness scale against performance and EEG variables. *Clinical Neurophysiology : Official Journal of the International Federation of Clinical Neurophysiology*, 117(7), 1574-1581.
- Kavanagh, D. J., Andrade, J., & May, J. (2005). Imaginary Relish and Exquisite Torture : The Elaborated Intrusion Theory of Desire. *Psychological Review*, 112(2), 446-467.
- Kavanagh, D. J., Statham, D. J., Feeney, G. F. X., Young, R. McD., May, J., Andrade, J., & Connor, J. P. (2013). Measurement of alcohol craving. *Addictive Behaviors*, 38(2), 1572-1584.
- Kreusch, F., Billieux, J., & Quertemont, E. (2017). Alcohol-cue exposure decreases response inhibition towards alcohol-related stimuli in detoxified alcohol-dependent patients. *Psychiatry Research*, 249, 232-239.
- Marissen, M. A. E., Franken, I. H. A., Blanken, P., van den Brink, W., & Hendriks, V. M. (2006). The Relation Between Social Desirability and Different Measures of Heroin Craving. *Journal of Addictive Diseases*, 24(4), 91-103.

- Masling, J. (1966). Role-related behavior of the subject and psychologist and its effects upon psychological data. *Nebraska Symposium on Motivation*, 14, 67-103. US: University of Nebraska Press.
- Nass, C., Moon, Y., & Carney, P. (1999). Are People Polite to Computers? Responses to Computer-Based Interviewing Systems¹. *Journal of Applied Social Psychology*, 29(5), 1093-1109.
- Nicholls, E., & Conroy, D. (2021). Possibilities and pitfalls? Moderate drinking and alcohol abstinence at home since the COVID-19 lockdown. *International Journal of Drug Policy*, 88, 103025.
- Nichols, A. L., & Maner, J. K. (2008). The Good-Subject Effect : Investigating Participant Demand Characteristics. *The Journal of General Psychology*, 135(2), 151-166.
- Nicovich, S. G. (2005). The Effect of Involvement on Ad Judgment in a Video Game Environment : The Mediating Role of Presence. *Journal of Interactive Advertising*, 6(1), 29-39.
- O'Brien, H. L., Cairns, P., & Hall, M. (2018). A practical approach to measuring user engagement with the refined user engagement scale (UES) and new UES short form. *International Journal of Human-Computer Studies*, 112, 28-39.
- Ooteman, W., Koeter, M. W. J., Vserheul, R., Schippers, G. M., & Brink, W. (2006). Measuring Craving : An Attempt to Connect Subjective Craving with Cue Reactivity. *Alcoholism : Clinical and Experimental Research*, 30(1), 57-69.
- Orne, M. T. (1962). On the social psychology of the psychological experiment : With particular reference to demand characteristics and their implications. *American Psychologist*, 17(11), 776-783. US: American Psychological Association.
- Papachristou, H., Nederkoorn, C., Corstjens, J., & Jansen, A. (2012). The role of impulsivity and perceived availability on cue-elicited craving for alcohol in social drinkers. *Psychopharmacology*, 224(1), 145-153.

- Paulhus, D. L. (1991). Measurement and Control of Response Bias. *Measures of Personality and Social Psychological Attitudes* (p. 17-59). Elsevier.
- Paulhus, D. L. (1998). Interpersonal and intrapsychic adaptiveness of trait self-enhancement : A mixed blessing? *Journal of Personality and Social Psychology*, *74*(5), 1197-1208.
- Perlmutter, L. C., Noblin, C. D., & Hakami, M. (1983). Reactive Effects of Tests for Depression : Theoretical and Methodological Considerations. *Journal of Social and Clinical Psychology*, *1*(2), 128-139. Guilford Publications Inc.
- Price, M., & Anderson, P. (2007). The role of presence in virtual reality exposure therapy. *Journal of Anxiety Disorders*, *21*(5), 742-751.
- R Core Team (2021). *_R: A Language and Environment for Statistical Computing_*. R Foundation for Statistical Computing, Vienna, Austria. Available at <https://www.R-project.org/>
- Radloff, L. S. (1977). The CES-D Scale : A self-report depression scale for research in the general population. *Applied Psychological Measurement*, *1*(3), 385-401.
- Rammstedt, B., & John, O. P. (2007). Measuring personality in one minute or less : A 10-item short version of the Big Five Inventory in English and German. *Journal of Research in Personality*, *41*(1), 203-212.
- Robillard, G., Bouchard, S., Renaud, P., & Cournoyer, L. G. (2002). Validation canadienne-française de deux mesures importantes en réalité virtuelle : L'Immersive Tendencies Questionnaire et le Presence Questionnaire. Présenté à 25e congrès annuel de la Société Québécoise pour la Recherche en Psychologie (SQRP), Trois-Rivières: Canada.
- Robinson, T., & Berridge, K. C. (1993). The neural basis of drug craving : An incentive-sensitization theory of addiction. *Brain Research Reviews*, *18*(3), 247-291.

- Roehrich, L., & Goldman, M. S. (1995). Implicit priming of alcohol expectancy memory processes and subsequent drinking behavior. *Experimental and Clinical Psychopharmacology*, 3, 402-410. US: American Psychological Association.
- Rohsenow, D. J., Monti, P. M., Abrams, D. B., Rubonis, A. V., Niaura, R. S., Sirota, A. D., & Colby, S. M. (1992). Cue elicited urge to drink and salivation in alcoholics : Relationship to individual differences. *Advances in Behaviour Research and Therapy, Urges and Cravings*, 14(3), 195-210.
- Rosenberg, M. (1965). *Society and the Adolescent Self-Image*. Princeton University Press.
<https://www.jstor.org/stable/j.ctt183pjhh>
- Ryan, J. J., Kreiner, D. S., Chapman, M. D., & Stark-Wroblewski, K. (2010). Virtual Reality Cues for Binge Drinking in College Students. *Cyberpsychology, Behavior, and Social Networking*, 13(2), 159-162.
- 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(6), 791-804.
- Sayette, M. A., Shiffman, S., Tiffany, S. T., Niaura, R. S., Martin, C. S., & Schadel, W. G. (2000). The measurement of drug craving. *Addiction*, 95(8), 189-210.
- Schacht, J. P., Anton, R. F., & Myrick, H. (2013). Functional neuroimaging studies of alcohol cue reactivity : A quantitative meta-analysis and systematic review. *Addiction Biology*, 18(1), 121-133.
- Schwarz, H., Revilla, M., & Weber, W. (2020). Memory Effects in Repeated Survey Questions : Reviving the Empirical Investigation of the Independent Measurements Assumption. *Survey Research Methods*, 325-344
- Segawa, T., Baudry, T., Bourla, A., Blanc, J.-V., Peretti, C.-S., Mouchabac, S., & Ferreri, F. (2020). Virtual Reality (VR) in Assessment and Treatment of Addictive Disorders : A Systematic Review. *Frontiers in Neuroscience*, 13, 1409.

- Simon, J., Etienne, A.-M., Bouchard, S., & Quertemont, E. (2020). Alcohol Craving in Heavy and Occasional Alcohol Drinkers After Cue Exposure in a Virtual Environment : The Role of the Sense of Presence. *Frontiers in Human Neuroscience*, *14*, 124.
- Singleton, E., Tiffany, S., & Henningfield, J. (1994). Alcohol Craving Questionnaire (ACQ-NOW) : Background, scoring, and administration. *Intramural Research Program, National Institute on Drug Abuse, Baltimore, MD*.
- Sinha, R., Fox, H. C., Hong, K. A., Bergquist, K., Bhagwagar, Z., & Siedlarz, K. M. (2009). Enhanced Negative Emotion and Alcohol Craving, and Altered Physiological Responses Following Stress and Cue Exposure in Alcohol Dependent Individuals. *Neuropsychopharmacology*, *34*(5), 1198-1208.
- Spielberger, C. D., Bruchon-Schweitzer, M., & Paulhan, I. (1993). *STAI-Y : Inventaire d'anxiété état-trait forme Y*. Paris, France: Éditions du centre de psychologie appliquée, DL 1993.
- Stein, K. D., Goldman, M. S., & Del Boca, F. K. (2000). The influence of alcohol expectancy priming and mood manipulation on subsequent alcohol consumption. *Journal of Abnormal Psychology*, *109*(1), 106-115.
- Tiffany, S. T. (1990). A cognitive model of drug urges and drug-use behavior : Role of automatic and nonautomatic processes. *Psychological Review*, *97*(2), 147-168.
- Tiffany, S. T., & Wray, J. M. (2012). The clinical significance of drug craving : Tiffany & Wray. *Annals of the New York Academy of Sciences*, *1248*(1), 1-17.
- Tourangeau, R., & Rasinski, K. A. (1988). Cognitive processes underlying context effects in attitude measurement. *Psychological Bulletin*, *103*(3), 299-314.
- Tournois, J., Mesnil, F., & Kop, J.-L. (2000). Autoduperie et hétéroduperie : Un instrument de mesure de la désirabilité sociale. *European Review of Applied Sociology*, *50*(1), 219-233.
- van Lier, H. G., Pieterse, M. E., Schraagen, J. M. C., Postel, M. G., Vollenbroek-Hutten, M. M. R., de Haan, H. A., & Noordzij, M. L. (2018). Identifying viable theoretical frameworks with essential

parameters for real-time and real world alcohol craving research : A systematic review of craving models. *Addiction Research & Theory*, 26(1), 35-51.

Webster, J., & Hackley, P. (1997). Teaching Effectiveness in Technology-Mediated Distance Learning. *The Academy of Management Journal*, 40(6), 1282-1309. Academy of Management.

Webster, J., & Ho, H. (1997). Audience engagement in multimedia presentations. *ACM SIGMIS Database: The DATABASE for Advances in Information Systems*, 28(2), 63-77.

Witmer, B. G., & Singer, M. J. (1998b). Measuring presence in virtual environments: A presence questionnaire. *Presence*, 7(3), 225-240.

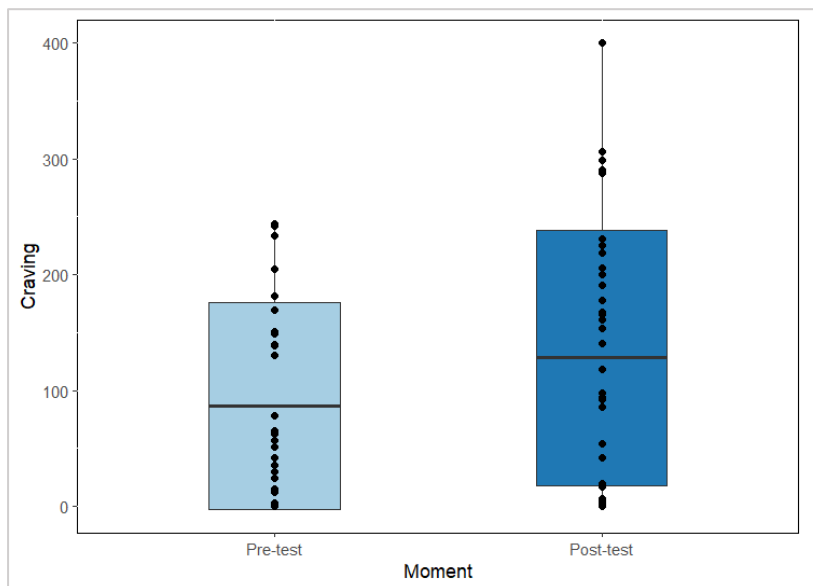
Zhang, Z., & Yuan, K.-H. (2018). *Practical Statistical Power Analysis Using Webpower and R* (Eds). Granger, IN: ISDSA Press. [<https://webpower.psychstat.org>]

Tables and figures

Figure 1. Screenshots of the video

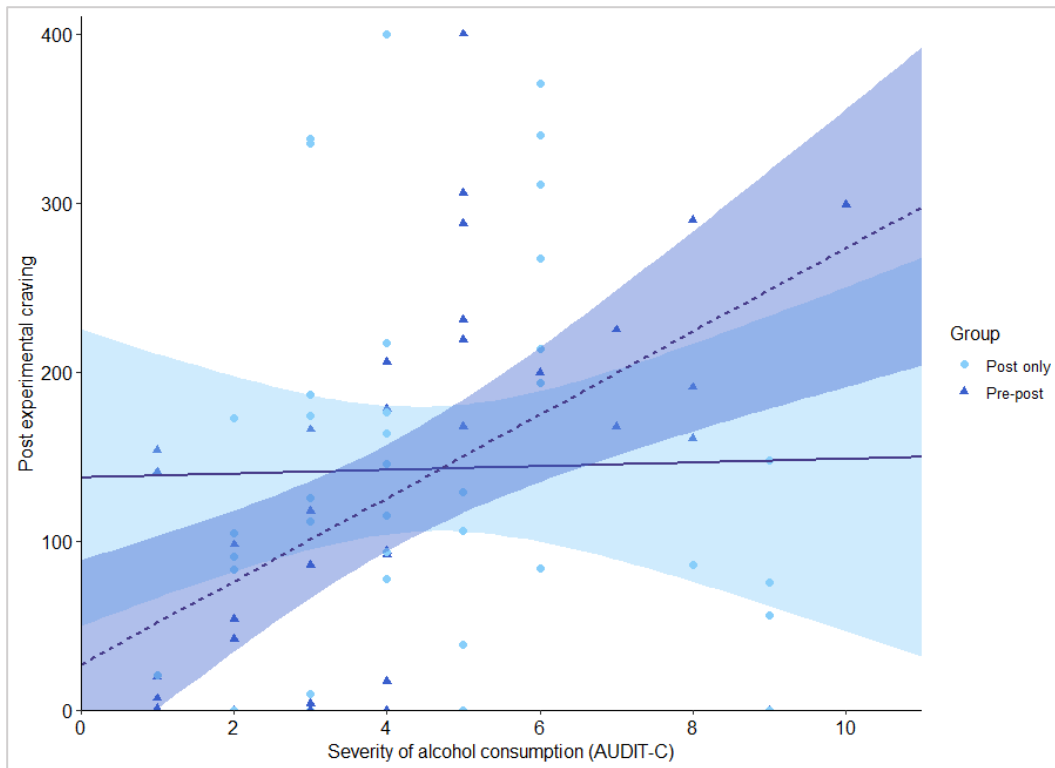


Figure 2. Alcohol craving levels before and after watching an alcohol-related video (pre-post group)



Note. The horizontal line represents the mean. The box represents the mean \pm STD and the lines around it are the minimum and maximum. Black dots are individual data points.

Figure 3. Relationship between the severity of alcohol consumption (AUDIT-C score) and post-test alcohol craving in the two groups



Note. The light blue and dark blue areas around the regression lines represent the confidence intervals at 95 %. The blue triangles and dots represent individual data.

Figure 4. Mediation model of the reward dimension of engagement on the association between alcohol consumption level (AUDIT-C) and (post-exposure) alcohol craving.

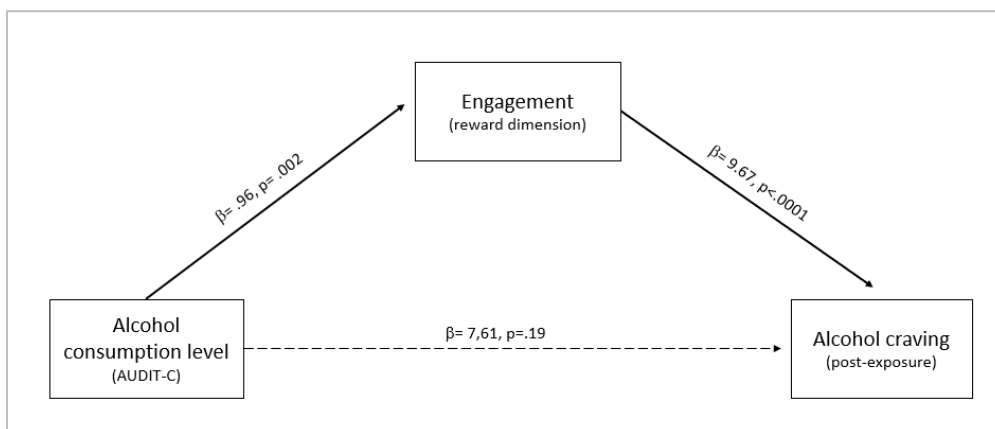


Figure 5. Moderated mediation model

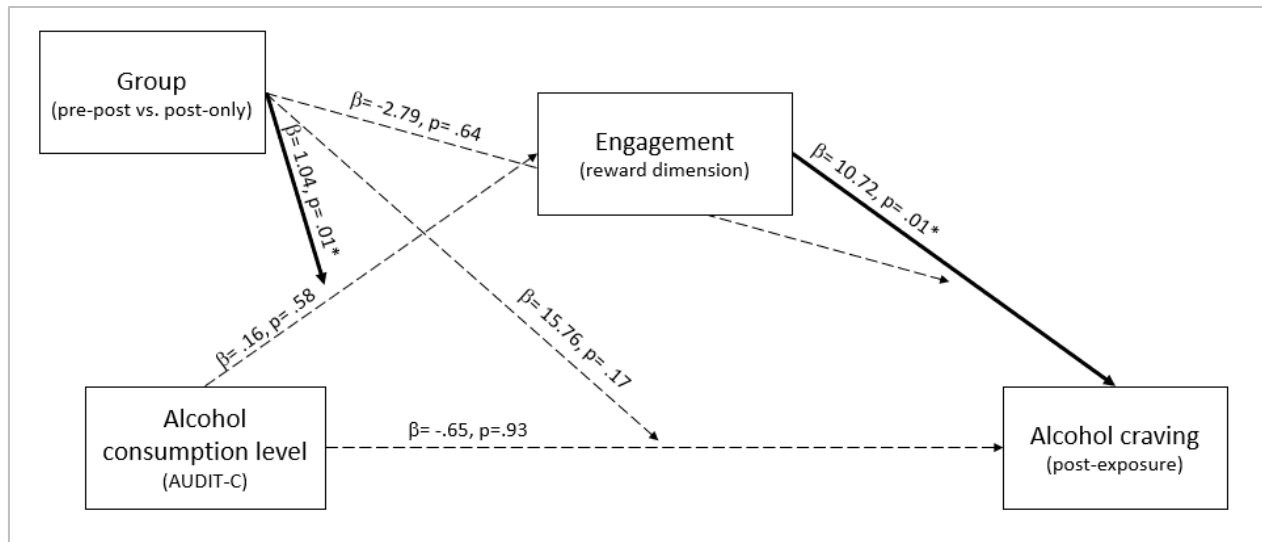


Table 1. Demographic, psychological and alcohol use variables in the two experimental groups.

Variable	Pre-post group	Post-only group	Statistic	p-value	Effect size
Gender (F:M:O)	20:15:1	24:15:0	/	/	/
Age	23.20 (4.30)	24.10 (4.20)	t(73)=-.92	.35	d=-.21
Education level^a	14.00 (2.47)	15.00 (2.58)	U = 565	.14	r _{rb} =.20
Social desirability					
SDE	61.50 (16.30)	63.20 (16.00)	t(73)=-.46	.64	d=-.11
IM	78.40 (13.30)	82.50 (12.80)	U=541	.08	r _{rb} =.23
Immersive tendencies					
Focus	23.60 (4.47)	22.60 (4.27)	t(73)=1.04	.30	d=.24
Involvement	21.20 (5.41)	20.30 (5.37)	t(73)=.78	.44	d=.18
Emotion	18.40 (5.85)	17.00 (4.33)	U=598	.26	r _{rb} =.15
Game	9.17 (4.72)	8.72 (4.10)	U=674	.76	r _{rb} =.04
AUDIT-C	4.11 (2.48)	4.56 (2.14)	U= 604	.29	r _{rb} = .14

No. of drinks^b	3.56 (4.21)	3.87 (7.41)	U=668	.71	r _{rb} = .05
Engagement					
Reward	21.10 (4.75)	19.30 (3.96)	t(73)=1.79	.07	d=.41
Focused attention	16.30 (3.16)	10.80 (4.68)	U=637	.49	r _{rb} = .09
Aesthetics		15.30 (3.51)	U=580	.19	r _{rb} = .17
Trait anxiety	52.30 (10.70)	47.40 (12.90)	U= 554	.11	r _{rb} = .21

Note. Standard deviations are in parentheses. SDE= “self-deceptive enhancement”; IM= “impression management”; AUDIT-C= “alcohol use disorder identification test” (short version).
^a number of completed academic years; ^b number of alcoholic drinks consumed during the last week.

Table 2. Results of the multiple regression analysis with the craving score as the dependent variable and group and alcohol consumption levels as explanatory variables.

Factor	Estimate	SE	t	p	η²p
AUDIT-C	24.63	7.10	3.47	<.001**	.09
Group	110.49	52.30	2.11	.038*	.001
AUDIT-C*group	-23.52	10.63	-2.21	.030*	.06

Note. AUDIT-C= alcohol use disorder identification test (short version).

* p<.05, ** p<.01.

Table 3. Results of the multiple regression analysis with the craving score as the dependent variable and the experimental group and demand effects as explanatory variables.

Factor	Estimate	SE	t	p	η ² p
--------	----------	----	---	---	------------------

Group	89.99	268.21	.34	.74	.005
CI	-2.74	11.28	-.24	.80	<.001
STDC	3.08	2.47	1.24	.22	.046
RB	.83	2.58	.32	.74	<.001
Group*CI	2.63	16.67	.16	.88	<.001
Group*STDC	.71	3.79	.19	.85	<.001
Group*RB	-1.48	3.81	-.39	.69	.002

Note. CI= Cue interpretation, STDC= Sensitivity to Demand characteristics, RB= Response

Behavior.

Table 4. *Percentage of individuals who reported more or less accurate assumptions about the experimental hypothesis in each group*

Elaboration on experimental hypothesis	Pre-post group	Post-only group
No hypothesis made	25%	41%
Hypothesis related to alcohol	44.44%	43.6%
Hypothesis related to craving for alcohol	25%	15.4%
Hypothesis related to biases that might affect the measurement (of craving)	5.56%	0.0%

Table 5. Estimates of the moderation effects of the dimensions of engagement in the two groups.

Effect	Pre-post Group				Post-only Group			
	Estimate	SE	t	p	Estimate	SE	t	p
<i>Reward</i>								
AUDIT-C	15.12	7.84	1.93	.06	-.65	8.11	-.08	.93
RW	7.92	4.09	1.94	.06	10.72	4.38	2.45	.01*
AUDIT-C*RW	1.41	1.58	.89	.37	-2.11	2.31	-.91	.37
<i>Focused Attention</i>								
AUDIT-C	22.63	6.33	3.58	.001**	3.42	7.93	.43	.67
FA	4.60	2.86	1.61	.11	10.32	3.62	2.85	.007**
AUDIT-C*FA	-.28	1.39	-.20	.84	-.16	2.45	-.07	.94
<i>Aesthetics</i>								
AUDIT-C	21.53	6.99	3.08	.004**	.50	8.21	.06	.95
AE	5.77	5.48	1.05	.30	10.82	5.00	2.16	.03*
AUDIT-C*AE	-.05	2.48	-.02	.98	-.10	2.44	-.04	.96

Note. RW= reward, FA= focused attention, AE=aesthetics.

*p<.05, **p<.01, ***p<.001.