Free-viewing multi-stimulus eye tracking task to index attention bias for alcohol versus soda cues

Soleymani, Ali; Ivanov, Yavor; Mathot, Sebastiaan; de Jong, Peter J.

Published in:
Addictive Behaviors

DOI:
10.1016/j.addbeh.2019.106117

Addictive Behaviors 100 (2020) 106117

Contents lists available at ScienceDirect

Addictive Behaviors

journal homepage: www.elsevier.com/locate/addictbeh

Free-viewing multi-stimulus eye tracking task to index attention bias for alcohol versus soda cues: Satisfactory reliability and criterion validity

Ali Soleymani[a,*], Yavor Ivanov[b], Sebastiaan Mathot[b], Peter J. de Jong[a]

[a] Department of Clinical Psychology & Experimental Psychopathology, University of Groningen, Groningen, the Netherlands
[b] Department of Experimental Psychology, University of Groningen, Groningen, the Netherlands

HIGHLIGHTS

- Cognitive theories emphasize the relevance of attentional bias (AB) in addiction.
- We need psychometrically sound measures of AB.
- The current free viewing eye tracking task showed satisfactory reliability.
- The AB measure also showed criterion validity.
- The stronger craving and alcohol problems, the stronger the AB for alcohol.

ARTICLE INFO

Keywords:
Attention bias
Assessment
Reliability
Criterion validity
Craving
Alcohol use

ABSTRACT

Cognitive -motivational models point to attention bias (AB) as an important factor in the persistence of problematic drinking behavior. Unfortunately, the measures that have been used to examine AB in addiction typically showed poor psychometric properties. To bring research on AB a critical step further it would be crucial to develop tasks with acceptable reliability and construct validity. Recently, Lazarov and colleagues (2016) developed a multi-stimulus free-viewing task (participants were free to look at any part of the screen and there was no secondary task involved) that showed excellent psychometric properties in the context of social anxiety as well as depression. We, therefore, adapted this task and examined its psychometric quality within the context of alcohol use. Participants with varying levels of alcohol use (N = 100) were presented with 54 matrices each containing 8 alcoholic and 8 non-alcoholic drinks. Each matrix was presented for 6 s. First fixation (100 ms) location and latency and total dwell time were assessed for alcohol and soda pictures. Assessment of AB, craving, and alcohol use (problems) was repeated after 3–8 days. Specifically, the dwell-time based AB-measure showed excellent internal reliability and considerable stability. Supporting the validity of the current AB-measures, it was found that participants with higher scores on craving and alcohol problems (i) dwelt longer on alcohol stimuli, and (ii) more often showed a first fixation on alcohol, whereas (iii) stronger craving was associated with shorter latency of first alcohol fixations. The AB-measure showed promising psychometric properties. Thus, this free-viewing eye-tracking task seems a welcome new tool for being used in future research on AB in addiction.

1. Introduction

Cognitive theories of addiction emphasize the relevance of heightened attention for substance cues as a causal agent in the development and persistence of problematic drinking behavior (Fadardi, Cox, & Klinger, 2006; Field & Cox, 2008; Franken, 2003). Studies that thus far examined whether indeed attentional bias might be involved in substance use disorder most often relied on indirect, reaction time (RT) based measures of attention such as the visual probe task (e.g., van Hemel-Ruiter, Wiers, Brook, & de Jong, 2016). Unfortunately, the currently available RT-based assessment tools typically show unacceptably low reliability as estimated by indices of internal consistency and stability (e.g., Ataya et al., 2012; Cisler, Bacon, & Williams, 2009). The low psychometric quality of the current AB measures poses a serious problem for a proper interpretation of the available findings and hampers scientific progress in this field of research (for critical

* Corresponding author at: Faculty of Behavioural and Social Sciences Experimental Psychopathology — Department Clinical Psych. & Exp. Psychopathology, Grote Kruisstraat 2/1, 9712 TS, Groningen, the Netherlands.
E-mail address: a.soleymani@rug.nl (A. Soleymani).

https://doi.org/10.1016/j.addbeh.2019.106117
Received 6 May 2019; Received in revised form 20 August 2019; Accepted 3 September 2019
Available online 05 September 2019

reviews see; Christiansen, Schoenmakers, & Field, 2015; Field et al., 2016; Rodebaugh et al., 2016).

The low reliability of RT-based measures might at least partly be due to the fact that the behavioral output (i.e., key presses) is quite distal from the to-be-measured attentional processes and may propound distracting factors like motor preparation and response execution (Armstrong & Olatunji, 2012). Moreover, RT-based AB measures typically reflect snapshots of attention thereby ignoring the dynamic nature of attentional processes (Bar-Haim, 2010; Shechner et al., 2013). Finally, it has been questioned whether the presentation of just one pair of stimuli (as is typically the case in the visual probe task) can be considered as an adequate reflection of the complexity of real-life substance use relevant situations (Hertel & Mathews, 2011). To more adequately mimic the complexity of daily life, and to challenge the attentional system, it seems essential to use a task with a more complex stimulus configuration.

To address these methodological limitations of RT-based AB paradigms, some researchers used eye tracking techniques to more directly assess overt attention to alcoholic stimuli whilst participants completed another RT task such as the visual probe task (Miller & Fillmore, 2010; Schoenmakers, Wiers, & Field, 2008; Weafer & Fillmore, 2013). Following such approach, several studies found that gaze measures tended to be more robust and showed a stronger relationship with actual drinking behavior than RT-based measures (e.g., Miller & Fillmore, 2010; Schoenmakers et al., 2008). In an attempt to further improve the reliability and construct validity of the visual probe task, Christiansen and colleagues used personalized next to general stimuli (Christiansen, Mansfield, Duckworth, Field, & Jones, 2015). Using personalized stimuli indeed resulted in higher internal consistency than general stimuli for both the RT and eye movement bias indices. However, these indices of AB were not associated with individual differences in alcohol use or craving.

In addition, some studies used eye tracking without a concurrent RT-task to measure attention bias in the context of alcohol use. In a typical task, two types of stimuli (alcoholic vs. non-alcoholic pictures) are presented side by side for a prolonged period (e.g., 2500 ms) while initial fixation and total dwell time per stimulus type are measured. Supporting the validity of eye tracking-based measures of attention using such procedure, it has been found that total fixation time on alcoholic stimuli was associated with the level of alcohol consumption (McAteer, Hanna, & Curran, 2018) and positive alcohol expectancies (McAteer, Curran, & Hanna, 2015). Eye tracking has also been used to index attentional bias for alcohol use during concurrent tasks such as an operant choice task in which participants had to make a choice between two simultaneously presented pictures (one presenting an alcoholic drink and one soda) to win soda or alcohol reward points (e.g., Rose, Brown, Field, & Hogarth, 2013; Rose, Brown, MacKillop, Field, & Hogarth, 2018), or a gaze contingency task in which participants were instructed to focus on a neural target while alcohol or non-alcoholic distractors were concurrently presented on the screen (e.g., Qureshi, Monk, Pennington, Wilcockson, & Heim, 2019; Wilcockson & Pothos, 2015). Unfortunately, none of these studies reported reliability estimates of the AB measures used in these studies, thus the reliability of these eye tracking based attention measures remains to be tested. In addition, all of these studies used a relatively simple stimulus configuration. Thus, these studies did not address the criticism that such configuration (e.g., presenting two stimuli side by side) may present a relatively poor reflection of real-life complexity.

As one way to improve the ecological validity of AB measures, a recent study successfully used portable eye tracking glasses to test AB for alcohol in a more naturalistic setting in which bottles of alcoholic beverages were positioned next to bottles of sodas (Monem & Fillmore, 2017). Results indicated that relatively long fixation on alcoholic beverages compared to sodas was associated with reported alcohol consumption during the past 90 days. Using another approach to increase the ecological validity of AB measures, Pennington and colleagues used a relatively complex stimulus configuration within the context of a conjunction visual search eye tracking task. In this task, participants had to detect an alcoholic or non-alcoholic target picture in an array of matching and non-matching distractors (Pennington, Qureshi, Monk, Greenwood, & Heim, 2019). Supporting the validity of this task, it was found that social drinkers (N = 30) were quicker to detect alcohol than non-alcohol targets, whereas proportional dwell time was relatively low for alcoholic distractors, suggesting that alcohol targets were relatively easy to detect. Yet, reliability estimates as indexed by Cronbach's alpha were still well below the conventional threshold of 0.70, and it remains to be tested whether AB, as indexed by this task, is sensitive to individual differences in alcohol use and problems. Furthermore, although the task used a relatively complex stimulus configuration, the alcohol and non-alcoholic stimuli were represented by a single stimulus (all alcohol/non-alcoholic stimuli in an array were similar). Moreover, because in this task participants are instructed to search for alcohol or non-alcohol cues, task performance may not provide an accurate reflection of individuals' spontaneous behaviors in the absence of such task requirements.

Therefore, the current study used a complex stimulus figuration within the context of a free-viewing task (i.e., participants were free to look at any part of the screen without any secondary task involved) that was adopted from research within the context of affective disorders. In this previous research, eye gaze was tracked while participants freely viewed 60 different matrices comprised of eight disgusted (or sad) and eight neutral facial expressions. Initial fixations and total dwell time to threatening versus neutral faces were used to index AB. In the context of social anxiety disorder (SAD) (Lazarov, Abend, & Bar-Haim, 2016) and major depressive disorder (MDD) (Lazarov, Ben-Zion, Shamaic, Pine, & Bar-Haim, 2018), this free-viewing task showed excellent reliability in terms of internal consistency (Cronbach alpha ranging between 0.85 and 0.95) and considerable test re-test reliability (ranging from 0.62 to 0.74). Importantly, it was also able to successfully differentiate between participants with and without (symptoms of) SAD, as well as between people with and without (symptoms of) MDD.

Based on these promising findings, the current study was designed to examine whether this free-viewing eye-tracking task could also be used to assess AB within the context of alcohol (mis)use. Consistent with Lazarov et al. (2016) eye-tracking was used to assess overall dwell time on alcoholic versus non-alcoholic (soda) pictures, participants' first fixations (on alcoholic vs. non-alcoholic pictures), and how much time passed before a first fixation could be identified (first fixation latency). To investigate the stability of the AB measures, AB was assessed twice, 3 to 8 days apart. Additionally, alcohol use, craving, and problematic behavior related to alcohol use were measured to test the criterion validity of the AB measures. We expected that participants with relatively high alcohol use, craving problems, and problematic behavior related to alcohol use would show: (i) greater total dwell time on alcoholic stimuli compared to non-alcoholic stimuli; (ii) more frequent first fixations on alcoholic stimuli relative to non-alcoholic stimuli; and (iii) shorter latency of first fixations on alcoholic than on non-alcoholic stimuli.

2. Methods

2.1. Participants

To reach a statistical power of at least 0.80 to reliably detect a correlation of medium effect size, we aim to enroll at least 85 participants (Cohen, 1988). Eventually, one hundred students (74 females, mean age = 22.97 years, SD = 3.82, range = 18–34) took part in the study. The only condition for enrollment in the study was uncorrected eyesight to prevent eye-tracking calibration difficulties. Participants received a small financial compensation. The study was approved by the ethical committee of the psychology department of the University of Groningen (ECP approval code = 17,372-S-NE (a, b)).

2.2. Materials

2.2.1. Questionnaires

The Alcohol Use Disorders Identification Test (AUDIT; Saunders, Aasland, Babor, de la Fuente, & Grant, 1993) is a 10-item assessment tool developed by the World Health Organization (WHO) to evaluate alcohol consumption (items 1–3), drinking behaviors and dependence symptoms (items 4–6), and alcohol-related problems (items 7–10). The sum score provides a measure of alcohol use disorder symptoms.

Rutgers Alcohol Problem Index (RAPI – 23-items; White & Labouvie, 1989) is a self-administered assessment questionnaire for evaluating drinking problems in adolescence. Participants were instructed to answer how many times each of the alcohol drinking-related problems happened to them in the last year (none, 1–2 times, 3–5 times, more than 5 times). Responses were coded 0–3 and summed for each person to make the total score. The higher score shows more alcohol-related problems associated with the participant.

The 5-item craving subscale of the Measurements in the Addictions for Triage and Evaluation (MATE -version 2.1) (Schippers & Broekman, 2014) was used to index craving for alcohol.

A Timeline Alcohol Use Calendar was used to assess alcohol use in the last 7 days based on the standard drink scale.

2.2.2. The free-viewing eye-tracking task

The free-viewing attention bias assessment task was designed in OpenSesame (Mathôt, Schreij, & Theeuwes, 2012). Eye-gaze data was collected with an Eyetribe ET1000 eye-tracker. Regions of Interest (ROI) were defined as the area in pixels of each of the displayed images. Thus, in our task, we have 16 ROIs per trial. The eye-tracking plugin collected eye-gaze position (i.e., x, y coordinates) at 30 Hz, and determined whether the current eye-gaze position falls within a ROI, or not. Before the start of each session, a 9-point eye-gaze calibration was performed. The task consisted of 3 blocks of 18 trials. During each trial, a 4 × 4 matrix was presented for 6 s. Each matrix consisted of 8 alcoholic pictures and 8 non-alcoholic pictures. The four inner pictures always consisted of two alcoholic and two non-alcoholic pictures. These pictures were taken from the Amsterdam Beverage Picture Set (Pronk, van Deursen, Beraha, Larsen, & Wiers, 2015). For each matrix, participants were required to gaze at a fixation dot in the center of the screen for 100 ms for that matrix to appear. After a subsequent 2000 ms interval, the next fixation dot appeared. To support participants' task engagement, we used three different types of stimuli: The first block consisted of matrices of bottle stimuli; the second of bottles together with an empty glass, and the last one of bottles with a filled glass. So, after eighteen matrices, the participant was prompted to move on the next set of matrices by pressing any key on the keyboard which mostly they moved to the next block immediately or after a few seconds. Eye tracking task took around 6–10 min in total Fig. 1.

2.2.3. Eye tracking measures

A custom-built plugin for OpenSesame was used to calculate for each trial the time a participant spent looking at each of the predefined regions of interest (i.e., each picture). Our analysis utilizes three output variables of the ROI analysis plugin: Dwell time per ROI; the ROI position within the matrix of the trial's first registered fixation; the first fixation's latency (i.e., the time difference between the first registered fixation within a ROI and the presentation of the matrix).

2.3. Procedure

First, participants completed the calibration. Then they were instructed that a series of matrices would be presented and that they could freely look at these images. It was explained that before the presentation of each picture matrix a dot would appear in the middle of the screen and that they needed to keep their eyes on the dot to be able to go to the next trial. The same order of matrix presentation was used for all participants to optimize the measurement procedure to index individual differences. Following the free-viewing task, participants completed the self-report measures. Finally, the second session was planned, 3–8 days after the first session (M = 5.3; SD = 1.5, Median = 6). The second session was similar to the first (same eye-tracking task followed by the MATE-craving problem questionnaire and self-reported alcohol use over the last 7 days).

2.4. Data reduction and analysis

Three variables were of interest when analyzing the output of the ROI analysis plugin: Fixation location; latency to detect the first fixated picture; and total dwell time on each of both picture categories. Attentional bias was indexed by the mean dwell time on alcoholic minus the mean dwell time on non-alcoholic pictures. Higher bias index (BI) scores thus indicate stronger AB for alcohol. As an estimate of reliability, we computed Cronbach's alpha for the 54 trial-level BI scores. As a second (and often used) index of reliability, we calculated the Spearman-Brown corrected correlation between the BI scores of the odd and even trials of the task (split-half reliability). To estimate the reliability of first fixation location (on alcohol or on soda) we also computed Cronbach's alpha for the 54 trials and assessed split-half reliability as indexed by the Spearman-Brown corrected correlation of the number of first alcohol fixations between odd and even trials.

To evaluate test re-test reliability, we computed the (Pearson's p-m) correlation between the bias indices of the first and second session. To examine the criterion validity of the current AB measures, Pearson correlations were computed between AB and alcohol (mis) use as indexed by standard alcohol use index (to measure last 7 days alcohol use), craving problem (measured with MATE-craving subtest), and problematic behavior related to alcohol use (evaluated using AUDIT and RAPI questionnaires).

3. Results

3.1. Descriptive

Table 1 presents a summary of the means (SDs) of all variables that were assessed per session. The results of the AUDIT questionnaire revealed that 63 of our participants were in low-risk level (They scored 7 or less) which 9 of them scored 0 and reported no alcohol use. In addition, 37 participants scored 8 or more which 34 of them were in the hazardous level (scored 8–15), and two were in the harmful or high-risk level. Finally, only one participant scored more than 20 which represents complete dependence to alcohol use.

3.2. Internal consistency and test re-test reliability

Internal consistency for the dwell-time based BI measure as indexed by Cronbach's alpha was excellent for both session 1 and 2 (alpha = 0.90 and 0.88, respectively). In addition, results of the Spearman-Brown corrected correlation indicated that there was a significant positive association between first and second half of the task's BI scores of a medium to large effect size: In the first session (rs (100) = 0.78, $p < .001$) and second session (rs (100) = 0.66, $p < .001$). A paired-samples t-test indicated that the BI index did not differ between sessions 1 and 2 [t (100) = 1.35, $p = .180$], means being 84.81 (SD = 524) and 22.13 (SD = 501), respectively (time interval between two sessions could vary between 3 and 8 days [M = 5.3; SD = 1.5, Median = 6]). The correlation between the BI index of session 1 and 2 was substantial and statistically significant; r (100) = 0.59, $p < .001$, though below the conventional lower limit of satisfactory test-retest reliability (i.e., 0.70).

Internal consistency for the type of the first fixation as indexed by Cronbach's alpha scores was very low for both session 1 and 2 (alpha = 0.11 and 0.02, respectively). The Spearman-Brown corrected

Fig. 1. An example of how different Regions of Interest defined (Blue areas defined as non-alcoholic and red area as alcoholic stimuli). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Table 1
Descriptive statistics of first and second assessment.

	First session	Second session
AUDIT	5.92 (4.13)	–
MATE (craving)	1.40 (1.85)	1.42 (1.77)
RAPI	5.08 (5.60)	–
Alcohol use (last 7 days)	6.09 (5.60)	5.76 (6.00)
Bias index	84 (524)	22 (501)
First fixation location (Alcohol)[a]	25.21 (3.56)	24.80 (3.72)
First fixation location (non-alcohol)	28.79 (3.56)	29.20 (3.72)
Mean Latency (Alcohol)[b]	411 (127)	438 (168)
Mean Latency (Non-alcohol)	415 (136)	437 (162)

[a] Number of trials.
[b] ms.

correlation of the number of alcohol fixations between odd and even trials during the first session was negative (instead of positive) and of small effect size (rs (100) = −0.172, $p < .09$). Also, during the second session the correlation was small and negative (rs (100) = −0.256, $p < .01$). Together, these findings indicate that the (estimated) reliability of the first fixations was low. The mean number of first alcohol fixations did not differ between sessions [t (100) = 0.84, $p = .39$], means being 25.21 (SD = 3.56) and 24.80 (SD = 3.72). However, the correlation between the number of the first alcohol fixations during the first and second session was very small and not statistically significant; r (100) = 0.12, $p = .22$, indicating low test-retest reliability. Also, the first alcohol fixation latencies did not differ between session [t (100) = −1.3, $p = .17$], means being 411.66 and 438.05. The correlation between the first and second session was however very small and not significant r (100) = 0.18, $p = .06$, indicating low reliability in terms of test-retest reliability. For the first soda fixation latency the correlation between session 1 and 2 was also between small and medium effect size though reached statistically significance, r (100) = 0.29, $p < .003$.

3.3. Criterion validity

Tables 2 and 3 presents a summary of the bivariate correlations and Fig. 2 shows the relationship between the dwell-time based attentional bias index and the AUDIT scores separately for session 1 and 2.

For session 1, the dwell-time based attentional bias index (BI) showed significant (and positive) correlations with the index of alcohol disorder symptoms (AUDIT), alcohol use-related problems (RAPI), and craving as indexed by MATE (see Table 2). Although the correlation between BI and alcohol use over the last 7 days was in the same direction, it did not reach the conventional level of significance. The pattern for attentional bias as indexed by the number of first alcohol fixations showed a similar pattern with (borderline) significant correlations with AUDIT, MATE, and RAPI, but not for past 7 days alcohol use (see Table 2). Attentional bias as indexed by the latency of first alcohol fixations showed only a significant correlation with craving problem (stronger craving problem was associated with shorter

Table 2
Correlations between measure of attentional bias and measures of alcohol (mis) use – first session.

		AUDIT[c]	MATE-craving[d]	RAPI[e]	Alcohol use[f]
Dwell time-based Bias Index[g]		0.35[b]	0.28[b]	0.27[b]	0.18
First alcohol fixations		0.25[a]	0.20[a]	0.19	0.09
Latency	Alcohol	−0.14	−0.23[a]	−0.14	−0.01
	Non-Alcohol	−0.21[a]	−0.35[b]	−0.20[a]	−0.07

[a] Correlation is significant at the 0.05 level (2-tailed).
[b] Correlation is significant at the 0.01 level (2-tailed).
[c] AUDIT (The Alcohol Use Disorders Identification Test) evaluate alcohol consumption, drinking behaviors and dependence symptoms, and alcohol-related problems.
[d] Measurements in the Addictions for Triage and Evaluation measures craving for alcohol.
[e] Rutgers Alcohol Problem Index evaluate drinking problems in adolescence.
[f] The last 7 days alcohol use based on the standard drinking scale.
[g] Bias Index; the mean dwell time on alcohol minus the mean dwell time on soda pictures.

Table 3
Correlations between measure of attentional bias and measures of alcohol (mis) use – second session.

		AUDIT[c]	MATE-craving[d]	RAPI[e]	Alcohol use[f]
Dwell time-based Bias Index[g]		0.14	0.20[a]	0.17	0.15
First alcohol fixations		0.19[a]	−0.01	−0.008	0.11
Latency	Alcohol	0.08[b]	−0.01	0.003	0.03
	Non-Alcohol	−0.04	−0.11	−0.11	−0.08

[a] Correlation is significant at the 0.05 level (2-tailed).
[b] Correlation is significant at the 0.01 level (2-tailed).
[c] AUDIT (The Alcohol Use Disorders Identification Test) evaluate alcohol consumption, drinking behaviors and dependence symptoms, and alcohol-related problems.
[d] Measurements in the Addictions for Triage and Evaluation measures craving for alcohol.
[e] Rutgers Alcohol Problem Index evaluate drinking problems in adolescence.
[f] The last 7 days alcohol use based on the standard drinking scale.
[g] Bias Index; the mean dwell time on alcohol minus the mean dwell time on soda pictures.

latency). Unexpectedly, shorter latency for first soda fixations was associated with higher AUDIT, MATE-craving, and RAPI scores (see Table 2).

For session 2, evidence for criterion validity was much weaker than for session 1. The dwell-time based BI only showed a significant relationship with craving problem, whereas AB, as indexed by the proportion of first alcohol fixations, was only associated with AUDIT score. The latency of the first fixation showed no significant relationships with measures of alcohol (mis) use.

4. Discussion

The aim of this study was to investigate the psychometric properties of attentional bias (AB) measures derived from a newly devised free-viewing eye tracking task that was adapted to assess AB for alcohol cues. The main findings can be summarized as follows: The dwell-time based AB measure showed high reliability as indexed by internal consistency and modest stability as indexed by test-retest reliability, whereas AB as indexed by proportion first alcohol fixations or first alcohol fixation latencies showed low internal consistency and low stability. In addition, both the dwell-time and first fixation-based AB measures showed satisfactory criterion validity as indicated by a positive correlation between AB and scores of craving problem and problems behaviors related to alcohol use.

Currently available RT-based AB measures (e.g., derived from the visual probe task) typically show unacceptably low reliability as estimated by indices of internal consistency and stability (e.g., Ataya et al., 2012; Cisler et al., 2009; Field et al., 2016). Low reliability of the current AB measures poses a serious problem for the interpretation of the available (non) findings and hampers scientific progress in this field of research (for critical reviews see Field et al., 2016; Rodebaugh et al., 2016; Jones, Christiansen, & Field, 2018). Therefore, the current study examined the psychometric properties of AB measures derived from a newly devised free-viewing eye tracking task that has shown adequate internal consistency, substantial test-retest reliability, and satisfactory criterion validity within the context of social anxiety disorder (Lazarov et al., 2016) and depression (Lazarov et al., 2018). In line with the previous findings of Lazarov et al. (2016, 2018) indicating that especially the dwell-time based AB measures showed acceptable reliability, the current findings indicated that also within the context of alcohol cues, the dwell-time based AB measure showed excellent internal consistency. Similar to the previous findings of Lazarov et al. (2016, 2018),

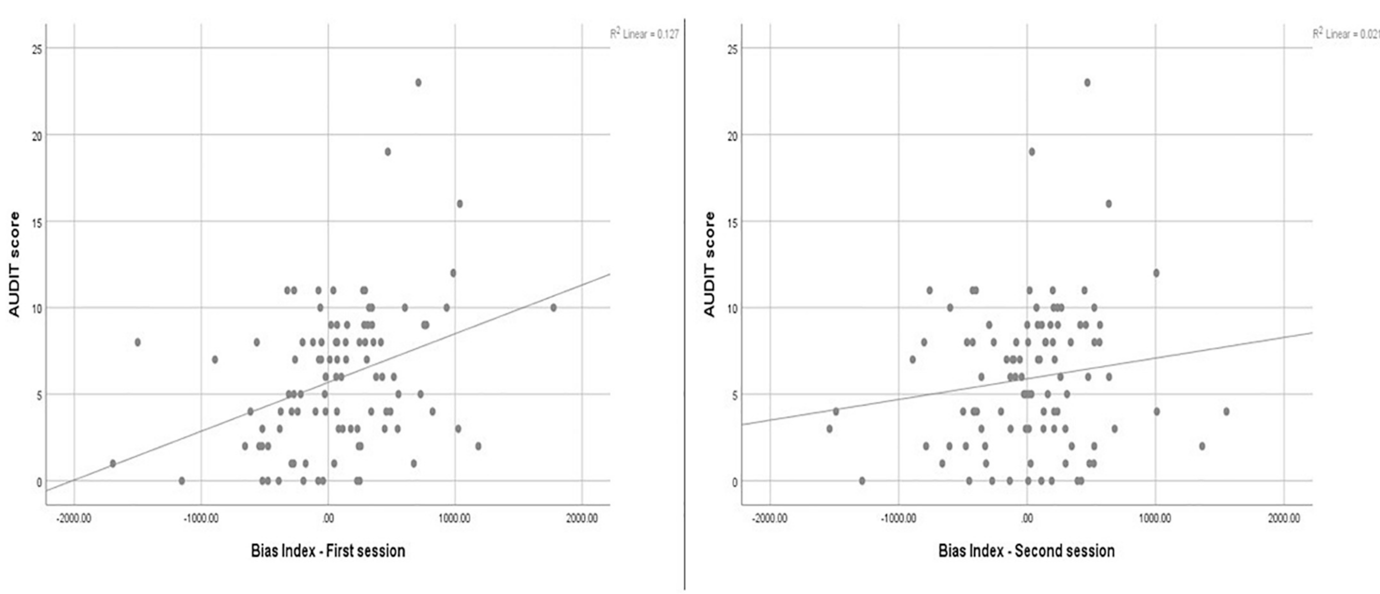

Fig. 2. Relationship between AUDIT score and Bias Index score in first and second sessions.

the AB measure based on the proportion of first fixations (in this study on alcohol stimuli) or the latency of the first (alcohol) fixations showed low internal consistency. On the one hand, this may indicate that the first fixation-based measures are not psychometrically sound. It could, however, also be that AB (also in the context of alcohol cues) is mainly driven by a difficulty to redirect attention away from alcohol cues and less so to the initial, attention-grabbing properties of alcohol cues (cf. Field et al., 2016). In other words, the first fixation-based measures may predominantly pick up noise instead of a systematic attentional pattern.

Consistent with such interpretation, it was found that the relationship with craving problems and problematic behaviors related to alcohol use was most pronounced for the dwell-time based AB measure. In addition, the current results are consistent with earlier findings within the context of alcohol use showing a more powerful correlation between craving and AB measures relying on direct (e.g., eye tracking) than on indirect (RTs) assessments of attentional processes (Field et al., 2016). The failure to find a correlation between RT-based attentional bias measures and craving problems may be the result of the relatively poor psychometric quality of the RT-tasks (Rodebaugh et al., 2016).

The medium-sized correlation between the current dwell-time based AB measure and the behavioral measures of alcohol use-related problems not only supports the psychometric quality of the current measure but also supports the view that AB may be involved in alcohol misuse. Thus, the current pattern of findings may also be taken to indicate that this newly developed eye-tracking task is a helpful and promising new tool to test further the hypothesized impact of AB on alcohol misuse (e.g., Franken, 2003; Field, Mogg, Bradley, & B., 2006; Field et al., 2016).

In further support of the usefulness of the current dwell-time based measure, the correlation between the measure of session 1 and 2 was substantial and similar to the test-retest correlation found by Lazarov et al. (2016) (0.59 vs. 0.63). In other words, compared to the often used RT-based AB measures derived from the visual probe task, the stability of the current dwell-time based measure was much higher, although it should be acknowledged that the current level of stability was still below the conventional lower limit of satisfactory test-retest reliability (0.59 vs. 0.70). One explanation for the still suboptimal stability of the current AB measure might be that habituation effects or carry-over effects have influenced the results during the second assessment, thereby lowering the correlation of the AB measure between both test sessions. Another explanation might be that by exposing participants to the same set of images on more than one session, there could be some boredom or fatigue in fixating on the same alcohol-related images that could lead to increased fixation on the other (non-alcoholic) items. Although this seems at odds with the relatively strong stability over time, it may be considered that this robustness may partly reflect variance that is not really related to the motivational salience of the stimuli but related to more general performance characteristics that are unrelated to the interest in the stimuli and independent of people's drinking behavior. Finally, it may be that AB can at least partly be best seen as a more state-like phenomenon, as AB has been shown to be sensitive to internal factors (e.g., stress, Field & Powell, 2007) and external factors (e.g., cue reactivity, Ramirez, Monti, & Colwill, 2015). Thus, even mild contextual differences across session may contribute to lowering the correlation between two test occasions. Clearly, this also points to the critical importance of keeping internal and external contexts as similar as possible across individuals when interested in using AB as a measure of individual differences.

The AB measure based on first (alcohol) fixation also significantly correlated with craving problem and alcohol use-related problems. This finding thus provided further support for the relevance of AB in alcohol abuse. However, because the reliability of this measure was not supported by the relevant reliability estimates this finding should be interpreted with care.

Finally, it was found that behavioral problems related to alcohol use and craving problem were associated with shorter first fixation latencies. This small to medium-sized relationship was, however, not specific for alcohol cues, and this pattern of results is therefore difficult to explain. One way to go would be to include a neutral non-drink category in the design (e.g., Pennington et al., 2019). Such a strategy might also more generally enhance the sensitivity of the design because an AB for alcoholic stimuli might be partly independent of an AB for sodas. People with an AB for drinks in general, would now fail to show an AB for alcohol because the AB measure is computed as the difference in dwell time between alcohol and soda stimuli (or the proportion of first alcohol fixations) between alcohol and soda stimuli.

The current study has several strengths including the relatively large sample size and the inclusion of two test sessions. However, there are also some important limitations that need to be acknowledged. First, the current study was restricted to a non-clinical student population and they were predominantly female. It would be important for future research to test whether the current findings are robust for both genders, and to replicate the current study in a clinical sample of people with substance use disorders. Second, AB for alcoholic cues may vary as a function of time of the day. For example, in the morning AB may be less pronounced than at the end of the afternoon. Because participants were measured throughout the whole day, this might have reduced the sensitivity of our design. One way to go would be to test participants at the point of the day during which they usually experience most craving. As a related limitation, the participants were measured in a 'sterile' lab-context. It seems reasonable to assume that the current type of AB for alcohol cues is context dependent. Perhaps, then, the sensitivity of the task to find reliable indices of AB can be further enhanced by testing in a relevant (e.g., bar) context. Along the same line, it could be argued that it would be helpful to test in a relevant mood state (internal context). Third, in this study, we only indexed alcohol consumption by assessing alcohol use during the last seven days. Although this may render our measure less sensitive to memory biases, it may not always reflect an accurate representation of the typical alcohol use pattern of participants. In addition, it is relevant to emphasize that in the current study the craving measure reflected mean craving over the last 7 days. Given the theoretically hypothesized relations between state craving and attentional bias (Franken, Zijlstra, Booij, & van der Brink, 2006), it would be relevant for future research to investigate the relationship between AB as indexed with the current task and the strength of momentary craving.

In conclusion, especially the dwell time-based AB measure that was derived from the newly devised free-viewing multiple-stimulus eye tracking tool showed excellent internal consistency and promising test re-test reliability. In addition, this measure showed satisfactory criterion validity that not only confirmed the psychometrical soundness of the AB measure but also the view that AB is involved in alcohol use problems. Taken together, this newly developed free-viewing assessment task seems a welcome addition to the available measures of AB and may be a helpful tool to further investigate the role of AB in the development and persistence of alcohol use problems.

References

Armstrong, T., & Olatunji, B. O. (2012). Eye tracking of attention in the affective disorders: a meta analytic review and synthesis. *Clinical Psychology Review, 32*, 704–723.

Ataya, F. A., Adamsa, S., Mullingsa, E., Cooperb, R. M., Attwooda, A. S., & Munafòa, M. R. (2012). Internal reliability of measures of substance-related cognitive bias. *Drug and Alcohol Dependence, 121*, 148–151.

Bar-Haim, Y. (2010). Research review: Attention bias modification (ABM): A novel treatment for anxiety disorders. *Journal of Child Psychology and Psychiatry and Allied Disciplines, 51*, 859–870.

Christiansen, P., Mansfield, R., Duckworth, J., Field, M., & Jones, A. (2015). Internal reliability of the alcohol-related visual probe task is increased by utilising personalised stimuli and eye-tracking. *Drug and Alcohol Dependence, 155*, 170–174. https://doi.org/10.1016/j.drugalcdep.2015.07.672.

Christiansen, P., Schoenmakers, T. M., & Field, M. (2015). Less than meets the eye: reappraising the clinical relevance of attentional bias in addiction. *Addictive Behaviors, 44*, 43–50.

Cisler, J. M., Bacon, A. K., & Williams, N. L. (2009). Phenomenological characteristics of

attentional biases towards threat: A critical review. *Cognitive Therapy and Research, 33*, 221–234.
Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum Associates.
Fadardi, J. S., Cox, W. M., & Klinger, E. (2006). Individualized versus general measures of addiction-related implicit cognitions. In R. W. Wiers, & A. W. Stacy (Eds.). *Handbook of implicit cognition and addiction* (pp. 121–133). Thousand Oaks, CA, USA: Sage Publications, Inc.
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. https://doi.org/10.1016/j.drugalcdep.2008.03.030.
Field, M., Mogg, K., Bradley, P., & B. (2006). Attention to drug-related cue in drug abuse and addiction: Component processes. In R. W. Wiers, & A. Stacy (Eds.). *Handbook of implicit cognition and addiction* (pp. 151–161). SAGE.
Field, M., & Powell, H. (2007). Stress increases attentional bias for alcohol cues in social drinkers who drink to cope. *Alcohol and Alcoholism, 42*, 560–566.
Field, M., Werthmann, J., Franken, I., Hofmann, W., Hogarth, L., & Roefs, A. (2016). The Role of Attentional Bias in Obesity and Addiction. *Health Psychology, 35*, 767–780.
Franken, I. H. A. (2003). Drug craving and addiction: integrating psychological and neuropsychopharmacological approaches. *Progress in Neuro-Psychopharmacology and Biological Psychiatry, 27*, 563–579.
Franken, I. H. A., Zijlstra, C., Booij, J., & van der Brink, W. (2006). Imaging the addicted brain: Reward, Craving, and Cognitive processes. In R. W. Wiers, & A. W. Stacy (Eds.). *Handbook of implicit cognition, and addiction* (pp. 185–199). SAGE Publication.
van Hemel-Ruiter, M. E., Wiers, R. W., Brook, F., & de Jong, P. J. (2016). Attentional biases and executive control in treatment-seeking substance-dependent adolescents: a cross-sectional and follow-up study. *Drug and Alcohol Dependence, 159*, 133–141.
Hertel, P. T., & Mathews, A. (2011). Cognitive bias modification: Past perspectives, current findings, and future applications. *Perspectives on Psychological Science, 6*, 521–536.
Jones, A., Christiansen, P., & Field, M. (2018). Failed attempts to improve the reliability of the alcohol visual probe task following empirical recommendations. *Psychology of Addictive Behaviors*. https://doi.org/10.1037/adb0000414 Advance online publication.
Lazarov, A., Abend, R., & Bar-Haim, A. (2016). Social anxiety is related to increased dwell time on socially threatening faces. *Journal of Affective Disorders, 193*, 282–288.
Lazarov, A., Ben-Zion, Z., Shamaic, D., Pine, D. S., & Bar-Haim, A. (2018). Free viewing of sad and happy faces in depression: A potential target for attention bias modification. *Journal of Affective Disorders, 238*, 94–100.
Mathôt, S., Schreij, D., & Theeuwes, J. (2012). OpenSesame: An open-source, graphical experiment builder for the social sciences. *Behavior Research Methods, 44*, 314–324. https://doi.org/10.3758/s13428-011-0168-7.
McAteer, A. M., Curran, D., & Hanna, D. (2015). Alcohol attention bias in adolescent social drinkers: an eye tracking study. *Psychopharmacology, 232*, 3183–3191. https://doi.org/10.1007/s00213-015-3969-z.
McAteer, A. M., Hanna, D., & Curran, D. (2018). Age-related differences in alcohol attention bias: a cross-sectional study. *Psychopharmacology, 235*, 2387–2393.
Miller, M. A., & Fillmore, M. T. (2010). The effect of image complexity on attentional bias towards alcohol-related images in adult drinkers. *Addiction, 105*, 883–890.
Monem, R. G., & Fillmore, M. T. (2017). Measuring heightened attention to alcohol in a naturalistic setting: A validation study. *Experimental and Clinical Psychopharmacology, 25*, 496–502.
Pennington, C. R., Qureshi, A. W., Monk, R. L., Greenwood, K., & Heim, D. (2019). Beer? Over here! Examining attentional bias towards alcoholic and appetitive stimuli in a visual search eye-tracking task. *Psychopharmacology*. https://doi.org/10.1007/s00213-019-05313-0.
Pronk, M., van Deursen, S., Beraha, D. M., Larsen, E. H., & Wiers, W. R. (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*, 2047–2055.
Qureshi, A., Monk, R. L., Pennington, C. R., Wilcockson, T. D. W., & Heim, D. (2019). Alcohol-related attentional bias in a gaze contingency task: Comparing appetitive and non-appetitive cues. *Addictive Behaviors, 90*, 312–317. https://doi.org/10.1016/j.addbeh.2018.11.034.
Ramirez, J. J., Monti, P. M., & Colwill, R. M. (2015). Alcohol cue exposure effects on craving and attentional bias in underage college student drinkers. *Psychology of Addictive Behaviors, 29*, 317–322.
Rodebaugh, T. L., Scullin, R. B., Langer, J. K., Dixon, D. J., Huppert, J. D., Bernstein, A., Zvielli, A., & Lenze, E. J. (2016). Unreliability as a threat to understanding psychopathology: The cautionary tale of attentional bias. *Journal of Abnormal Psychology, 125*, 840–851.
Rose, A. K., Brown, K., Field, M., & Hogarth, L. (2013). The contributions of value-based decision-making and attentional bias to alcohol-seeking following devaluation. *Addiction, 108*, 1241–1249.
Rose, A. K., Brown, K., MacKillop, J., Field, M., & Hogarth, L. (2018). Alcohol devaluation has dissociable effects on distinct components of alcohol behavior. *Psychopharmacology, 235*, 1233–1244.
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.
Schippers, G. M., & Broekman, T. G. (2014). *MATE-Q 2.1. manual.* Bêta Boeken: Nijmegen.
Schoenmakers, T., Wiers, R. W., & Field, M. (2008). Effects of a low dose of alcohol on cognitive biases and craving in heavy drinkers. *Psychopharmacology, 197*, 169–178.
Shechner, T., Jarcho, J. M., Britton, J. C., Leibenluft, E., Pine, D. S., & Nelson, E. E. (2013). Attention bias of anxious youth during extended exposure of emotional face pairs: An eye- tracking study. *Depression and Anxiety, 30*, 14–21.
Weafer, J., & Fillmore, M. T. (2013). Acute alcohol effects on attentional bias in heavy and moderate drinkers. *Psychology of Addictive Behaviors, 27*, 32–41.
White, H. R., & Labouvie, E. W. (1989). Towards the assessment of adolescent problem drinking. *Journal of Studies on Alcohol and Drugs, 50*, 30–37.
Wilcockson, T. D., & Pothos, E. M. (2015). Measuring inhibitory processes for alcohol related attentional biases: introducing a novel attentional bias measure. *Addictive Behaviors, 44*, 88–93. https://doi.org/10.1016/j.addbeh.2014.12.015.