**Author’s response to reviews**

**Title:** “Free won't” after a beer or two: Chronic and acute effects of alcohol on neural and behavioral indices of intentional inhibition

**Authors:**

Yang Liu (y.liu3@uva.nl)

Wery van den Wildenberg (w.p.m.vandenwildenberg@uva.nl)

Gorka Fraga González (gorkafraga@gmail.com)

Davide Rigoni (davide@impulseresearch.be)

Marcel Brass (marcel.brass@ugent.be)

Reinout Wiers (r.wiers@uva.nl)

Richard Ridderinkhof (k.r.ridderinkhof@uva.nl)

**Version:** 1 **Date:** 03 Oct 2019

**Author’s response to reviews:**

Dear Dr. Pozza,

We are grateful for the possibility to submit a revised version of our manuscript entitled “ “Free won't” after a beer or two: Chronic and acute effects of alcohol on neural and behavioral indices of intentional inhibition”, authored by Liu, van den Wildenberg, González, Rigoni, Brass, Wiers and Ridderinkhof (Manuscript number: PSYO-D-19-00135).

We thank the reviewers for their many positive comments and helpful detailed suggestions. We have addressed each of their issues pointwise below. To enhance the readability of our rebuttal, the comments of the reviewers are in blue and sequentially numbered. Our replies are shown in black.

We believe our manuscript has greatly benefited from their valuable comments, and we hope you will find the manuscript to be in good shape. We look forward to learning your further decision.

Sincerely,

Yang Liu (on behalf of the co-authors)
REVIEWER 1 (Ann-Kathrin Stock)

The manuscript submitted by Liu et al. reports the effects of lifetime/long-term alcohol use and acute alcohol intoxication on stimulus-driven vs. intentional inhibition. In order to assess these two concepts of inhibition, the authors used a "Chasing Memo" task and a stop-signal task in two unrelated (?) healthy young student samples. The analyses are appropriate and the use of additional Bayesian analyses should be commended. The study question is both timely and interesting, as the effects of alcohol onto volitional inhibition have been much less researched than cued inhibition. However, I am still a bit skeptical, whether the initially proposed research questions can be adequately investigated with the available measures. Yet, it should be well possible to address this issue. In my opinion, the most feasible solution would be make some rather small, but relevant amendments to the wording of the research question and somewhat tone down the deducted conclusions.

We thank the reviewer for the compliments. We will address her questions and concerns listwise below. Regarding the questions posed here:

- “two unrelated (?) healthy young student samples”:
Samples of these two studies were indeed independent. To make this explicit, we modified this sentence “Twenty right-handed male adults independent from Experiment I participated…” (p. 16)

- “small, but relevant amendments to the wording of the research question and somewhat tone down the deducted conclusions”:
research questions and conclusions were rephrased; see our replies to Q1, Q2, and Q4 below.

Q1: General Background:
In the general background section, as well as in some parts of the following manuscript, the authors state that "In terms of drinking, the priming dose effect of alcohol, i.e., loss of control over further consumption after a priming dosage, reflects the insufficiency of intentional inhibition rather than stimulus-driven inhibition (Field, Wiers, Christiansen, Fillmore, & Verster, 2010)." While this is most certainly a big issue in individuals with AUD, a few alternative explanations could also be conceivable in individuals with non-pathological drinking habits: In the investigated samples, the priming dose could also lead to a shift in plans/motivational values, or simply to changes in delay discounting. So while investigating alcohol effects in healthy samples is of course interesting and valid research question, I am skeptical whether this allows for valid conclusions on AUD mechanisms. -After all, loss of control over drinking is one of the key symptoms in AUD, which likely sets affected individuals apart from the general population.

We thank the reviewer for pointing out this important issue. Indeed, three criteria of DSM-5 in diagnosing AUD are relevant to the loss-of-control behavior. However, loss-of-control is thought to play important roles in different stages of addiction, including initial use of a substance and transition from recreational use to heavier use and abuse, rather than only when one is diagnosed with AUD. For instance, if a binge drink occasion happened on a Friday night in the student bar,
this event is unlikely to add to a diagnosis of loss of control related to AUD. However, the same behavior on a Thursday night, before big exams on Friday could add to the diagnosis. Therefore, we would argue that loss of control over drinking among healthy samples can still be relevant, especially when there is no adaptive adjustment to (changing) circumstances (cf., Albertella, Watson, Yücel, & Le, 2019). Second, loss-of-control is the behavioral output, and shifting in plans/motivational values, or changes in delay discounting could all be relevant underlying processes. Such reasons may or may not differ between people with AUD and healthy recreational users.

In sum, we admit that it is necessary to clarify that we focused on healthy recreational users and that findings cannot be generalized to people with AUD.

In the Abstract under the Conclusion (p. 2)

“These findings suggest that both past-year increases in risky alcohol consumption and moderate acute alcohol use have limited effects on stimulus-driven inhibition and intentional inhibition. These conclusions cannot be generalized to alcohol use disorder and high intoxication levels.”

In the general discussion

p. 22 “Accordingly, our conclusions cannot be generalized to the population with AUD.”

Also in our reply to Q3.

Q2: Experiment 1:
The authors stated that they wanted to investigate the effects of "long-term" or "lifetime" alcohol use on inhibition These should not be used as interchangeable terms as lifetime use usually refers to lifetime prevalence, while long-term use is more commonly defined as prolonged and continuous consumption.

Unfortunately, the assessed measures of alcohol consumption do not allow to answer this particular research question. More specifically, neither the AUDIT, nor the CORE provide proper measures of long-term alcohol use (supposing that all participants reported lifetime use, I assumed that the authors most likely wanted to focus on long-term use): The AUDIT was initially developed to distinguish heavy drinkers with AUD from rather strictly abstinent controls. Nowadays, this fact is mostly disregarded and it has become a widely tool which provides increasing scores with increasingly risky drinking behavior. While this application of the test seems warranted, I doubt that this score provides an accurate measure of long-term alcohol, especially as the AUDIT neither assesses the time of first consumption, nor any changes in consumption frequency and patterns over time. Compared to this, the long version of the CORE assesses age of first consumption, binge drinking within the past two weeks, and recent drinking prevalence and. Yet, the CORE questions are not neutral, but slightly biased by moralizing and judgmental undertones commonly found in US campus drug politics, and the test mainly assesses consumption within the past 12 months. Given that the available literature on long-term use often reports on the effects of much longer time intervals, and further given that the students who were included in the current study are unlikely to have started their substance consumption within the 12 months prior to testing, I would suggest to either clearly define long-term use as pertaining to this 12-month time interval, or use different term (like "consumption habits", or the like). In summary, I would recommend that the authors would need to EITHER
rectify their hypotheses to what is actually assessed by the AUDIT, OR explicitly define long-term use as a 12 month period and then focus on the CORE, OR recollect data with a more detailed and appropriate questionnaire.

We are grateful to the reviewer for bringing up these valuable comments and suggestions. First, wording such as lifetime and long-term alcohol use have been replaced by past-year alcohol use.

Second, we did not administer the long version of CORE and only know in the past year and past month on how many occasions the participants used a certain kind of substance. Therefore, we prefer to continue using the AUDIT score and rephrase the research question and conclusions achieved from it:

- the hypothesis of study 1
p. 5 “we tested the hypothesis that higher AUDIT scores (i.e., increase in problematic alcohol use) were associated with prolonged SSRTs”

- discussion of study 1
p. 12 “In the first experiment, past-year increase in risky drinking showed no relationship with any of the inhibition-related tasks and questionnaires.”
In addition, the first three items in AUDIT (also known as AUDIT-C) assess alcohol consumption. We, therefore, replicated our analysis with AUDIT-C.
p. 9 footnotes: “In addition, we replicated these analyses by replacing AUDIT total score by AUDIT-C (the first three items of AUDIT).”
p.12 “Results were very similar when AUDIT-C was used (see Supplementary Materials).”
Accordingly, in the Supplementary Materials, all those results were reported.

Q3: Against this background, I also found some of the current conclusions to be only partly justified (e.g. "it's time to ask whether the stimulus-driven inhibition deficit is a real finding among drinkers"). These should be specified or toned down a bit.
Yes, we agree with the reviewer, this statement may be a bit strong. We now rephrased it into

p. 12 “It’s time to re-assess the connection between recreational moderate alcohol use and stimulus-driven inhibition impairment.”

p. 12 “Relatedly, in our recent individual-level mega-analysis, very limited evidence supporting such deteriorating relationship was found across a broad range of substances (see Liu et al., 2019). As only a small proportion of the participants are diagnosed with SUD, the conclusions cannot be applied to SUD.”

Q4: Experiment 2:
The objective of experiment II was to test "whether and how acute alcohol use influence intentional inhibition". While inhibition is known to be most reliably impaired at high intoxication levels (i.e. typically more than 0.08 %), the authors only induced moderate intoxication of ~0.06 %. The intoxication level might hence have been too low to observe valid intoxication effects. While the authors already clearly mentioned this in their discussion, it
should also be recognized more explicitly (e.g. in the abstract) that the experimental setup does therefore not allow for general conclusions on the full range of acute intoxication, as often seen in the field.

Yes, we totally agree with the reviewer.
- Abstract: see reply to Q1 above.

- Discussion of study 2

p. 21 “Our conclusions cannot be generalized to the full range of acute intoxication.”

Q5: Furthermore, it should be more critically debated whether testing the effects of moderate alcohol intoxication on general domains of inhibition in a non-addicted, highly homogenous group of 16 healthy young psychology students without AUD truly allows for general conclusions on loss of control over drinking. This is relevant to Q1 above. This is also in line with the conclusion from our recent mega-analysis (Liu et al, 2019): we found no evidence in non-SUD samples, but of course there might still be an association in SUD samples. We rephrased the conclusion:
p. 24 “Caution should be taken when extending these conclusions to AUD populations and higher intoxication levels (e.g., 0.08%).”

Q6: Lastly, I was surprised that the N2 and P3 were not quantified at all. -After all, they are the most commonly analyzed ERPs in the context of inhibition and previous studies have explicitly reported intoxication effects on these measures of inhibition. Instead, the focus was put on the somewhat less popular AUC and RP. It would be helpful if the authors could elaborate more on their reasons for doing so.

We acknowledge that these are useful suggestions. Our choices were based on a priori considerations, and these were perhaps not explained clearly enough. So we have now clarified our focus on the Readiness Potential (RP) (as quantified with, a.o., Area Under Curve (AUC)), and elaborated on why we did not quantify N2/P3, on p. 15:

“First, for intentional inhibition we focused on neural activities preceding rather than after intentional inhibition, as 1) this can help predict when intentional inhibition is likely to happen; 2) for voluntarily chosen action/inhibition, nearly all cognitive processes happened before execution of the action; 3) there is no external stop-signal to be time-locked to, which makes the comparison with cued-inhibition on N2/P3 less relevant. Second, N2/P3 comprises a complex of well-known EEG component that is typically associated with cued-inhibition. Since the focus here is not on replicating previous findings of cued inhibition but on exploring the neural activities relevant to intentional inhibition as compared to cued inhibition, and since no N2/P3 could be expected (or indeed observed) for intentional inhibition, our focus was on the RP rather than the N2/P3 complex.”
REVIEWER 2 (Janette Smith)

The manuscript presents two studies; the first, examining whether performance on two types of inhibitory tasks is associated with chronic alcohol use, and the second examining whether performance and brain electrical activity during one type of inhibitory task differs under acute alcohol administration.

The authors are to be commended for tackling the issue of the real-world significance of inhibitory problems associated with alcohol use - the link between performance on inhibitory tasks in the lab, and risky drinking in the real world, is not always clear, and the Chasing Memo task is a nice attempt to operationalise intentional inhibition. The manuscript is generally well written, and my comments are mostly minor suggestions.

We thank the reviewer for the compliments.

p5: the reference to Bo and Landro (2017) here is incorrect; they in fact found drinkers as a whole had *shorter* SSRT than teetotallers, with no further differences associated with units of alcohol/week

Thanks for this correction. Now this reference was replaced by (Smith & Mattick, 2013) on p. 5.

p6: please specify whether the exclusion criterion for depression was current, former, or both

This has been changed into ‘no prior and current diagnosis of depression’ on p. 6.

p6: it is not clear whether participants completed Study 1 in a or online; how were the problems associated with refresh rate detected, and what was the decision rule to exclude these participants?

Participant completed study 1 in the lab, as now clarified on p.9:

“All participants signed an informed consent prior to the laboratory session”

This Chasing Memo task was programmed in a computer with a refresh rate of 60Hz. However, some of the computers in our labs are of 120Hz. We haven’t realized this would influence the moving speed of Memo until a student reported the difference she felt during the practice session (i.e., it was very hard for some participants to track Memo as it moved quite fast). We clarified this point by adding a footnote on p. 6:

“For computers with incorrect refresh rate settings, Memo moved at twice of the speed as planned, which made the moment of disengagement incomparable.”

p7: please clarify the time frame for "successful tracking" - was tracking a failure the moment the cursor moved beyond the 2cm radius? Was it successful the moment it moved back, or was there a minimum time before accumulation began again? Please can you provide some measure of task difficulty (e.g., the average percentage of time that a participant was accumulating points)?

Successful tracking means as long as the cursor is within the 2cm radius. In that way, the accumulation started immediately when moved back. We now made this clearer on p. 7:

“…during successful tracking (i.e., as long as the cursor is within this green zone)…”
Task difficulty is indeed an important issue here. In our task, it comprised a complicated subjective feeling. For a rough objective estimation, the total number of times participant lost the star was selected as the indicator, as we have now added to p. 10:

“Task difficulty was assessed by the number of times one lost the star. Out of the 120 trials, on average participants lost the star 31 times (SD = 21), ranging from 6 to 145. This indicates that most of the participants have a good mastery of the task and should be able to allocate attention to their behavioral intentions.”

p7: I could not find any analysis of the counter in the cued condition - is it only included to be similar to the free condition? How similar was the actual count to participant's reported counts (and if quite different, how does this affect estimation of the W-moment)? Is this value affected by alcohol (whether chronic or acute use)?

In the cued condition, the counter was used to calculate timing accuracy (the difference between the reported and the actual appearance moment of the stop signal), as was reported on p. 9. In the model of W-interval, the influence of timing accuracy has been controlled for (p. 10).

We now analyzed and reported how long-term alcohol use might influence timing accuracy:

p. 9 “The possible association between past-year alcohol use and timing accuracy was examined by Pearson correlation.”

p. 11 “Past-year risky alcohol consumption is not associated with alteration in timing accuracy (r = -0.21, p = 0.10, BF01 = 1.66).

Acute alcohol use did not influence timing accuracy as was reported on p. 19.

p12: were the authors expecting to observe no correlation between the two types of inhibition in experiment 1 - i.e., the SSRT and the disengage RT?

Thank you for pointing out this important question. It was found in previous fMRI studies that intentional inhibition and stimulus-driven inhibition engage a common inhibition network, but intentional inhibition is also characterized by additional context-dependent neural activation in medial prefrontal cortex (Schel et al., 2014). How behavioral output can reflect such overlap is unknown. Being an exploratory analysis, we can expect only a small to moderate relationship.

p. 12 footnotes “We only expect a small to moderate relationship between SSRT and disengage RT as intentional inhibition engaged additional neural activation albeit common inhibition network with stimulus-driven inhibition (Schel et al., 2014).”

p16: 1) there is no report of a manipulation check for the placebo vs. alcohol condition (e.g., participants had to guess which condition they did each day, with correct guesses at chance levels). 2) Although a second experimenter prepared the drinks and measured the breath alcohol concentration, it is possible that they interacted with the participant differently in the two conditions. 3) Further, participants were required to stay until their breath alcohol concentration was under 0.02%, providing a further clue as to which drink had been given.

1) Our alcohol administration procedure and manipulation check are the same as Korucuoglu et al (2015). We have now clarified as follows:

p. 17 “A short manipulation check interview was performed at the end of each session to make sure participants are aware of the alcohol content of the drink.”

Although in the placebo condition participants all believed there is alcohol in the drink, the amount reported is less than that in the alcohol condition. Together with question 2) we now mention this as a limitation in the general discussion:
Fourth, there is room for alcohol administration and placebo conditions to be improved, given that although all participants reported they received alcohol in the placebo condition, the amount is less than that in the alcohol condition; the experimenter blind to alcohol condition may interact with participants differently in two conditions (alcohol/placebo) due to the participants’ status (drunk/sober). We acknowledge this as a potential shortcoming, although these are common issues in this field, and generally not considered overly detrimental to interpretation.”

3) We thank the reviewer for bringing this up; we failed to mention that in line with Korucuoglu et al (2015), participants were informed in advance that they once received a low dose of alcohol and once a high dose for two sessions. Thereby, the presence of expectancy effects is independent of whether they left the lab immediately or not.

p. 16 “They were informed that they would receive a low dose and a high dose of alcohol for two sessions. This assured the presence of expectancy effects in both sessions.”

p17: ‘v.v.’ is not a common abbreviation for (presumably) 'vice-versa'; please spell it out. Thanks, it has been spelled out.

p20: I am surprised that the alcohol-related difference visible in Figure 3 is not significant, since it looks quite large. Please could the authors report an effect size for this comparison; it is possible that a true effect has been missed due to low power from the small sample size. We thank the reviewer for pointing this out. We now reported the effect size of factor Alcohol in the RP onset model and RP AUC model (p. 20). Indeed, the effect sizes are rather small.

p. 22 “Furthermore, a true effect might have been missed due to low power from the small sample size.”

p23: please briefly describe the Marble Task - what does the participant see/hear/do? what psychological processes does it tap? Why is the Chasing Memo task better?
We now introduced the Marble task in more detail on p. 4 where it was mentioned for the first time and its limitation has been noted.

“For instance, in the Marble task, participants view a white marble rolling down a ramp. In 50% of the trials, the marble turns green and participants have to stop it from crashing as fast as possible by pressing the button. If the marble remains white, the participants can choose between performing the prepared action (i.e., stop the marble) and execution of intentional inhibition (i.e., do not stop the marble).”

Figure 1: Please add a display of the digits within the circle, which participants had to note and enter at the end of each trial. I initially misread the methods and thought this number was in the bottom right, which would cause issues with eye movement activity to the bottom right corner coincident with the W-moment. As it is, this is not an issue after all, because the counter was actually presented very close to where the participant would likely be looking. But to save others making the same mistake, please add a display of the counter within the circle as it appeared to the participant. Thanks for pointing this out. This is now added in Fig. 1.
The following additional changes were made

- First, author Gorka Fraga González’s name should be Gorka Fraga González

- Second, as the status of Liu et al (2019) paper changed from under review to published, corresponding changes were made in both the text and the reference list.

- Third, we reformulated the way in which Bayesian results were reported in experiment I (p. 10-12), to make it consistent throughout the paper.