PEER REVIEW HISTORY

ARTICLE DETAILS

TITLE (PROVISIONAL)	Effects of chess-based cognitive remediation training as therapy add-on in alcohol and tobacco use disorders: protocol of a randomized, controlled clinical fMRI trial
AUTHORS	Gerhardt, Sarah; Lex, Gereon; Holzammer, Jennifer; Karl, Damian; Wieland, Alfred; Schmitt, Roland; Recuero, Ainoa Jiménez; Montero, Juan Antonio; Weber, Tillmann; Vollstädt-Klein, Sabine

VERSION 1 – REVIEW

REVIEWER	Gobet, Fernand University of Liverpool
REVIEW RETURNED	21-Nov-2021

GENERAL COMMENTS	This protocol presents the design of a study using chess-based cognitive remediation training as therapy add-on for individuals with alcohol or tobacco use disorders. It is overall clearly written and in general presents sufficient details. My main concern is that the wrong design is used. To control for placebo effects, you need to use a design with three groups, one of which is an active control group: - treatment as usual - treatment as usual + Chess-based cognitive remediation training - treatment as usual + Some other novel activity [active control group] Without this third group, it is logically impossible to draw any conclusions for possible differences between the first two groups after treatment, as chess specific effects and placebo effects are confounded. Another concern is the a priori likelihood that chess has positive effects (beyond placebo effects) is low. In general, research on cognitive training has found few, if any, reliable positive effects with respect to cognitive and scholastic outcomes (Sala et al., 2019). One exception is chess, but most of the research on this topic uses a two-group design (as in the proposal), which is insufficient to draw any reliable conclusion. When the correct three-group design was used, no effect was found, as noted in the proposal.

	The proposal mentions a few studies using chess as therapy, with encouraging results. However, all those studies use a weak design, so it's not possible to reach any trustworthy conclusion. Incidentally, the authors could make their case stronger if they could present data showing that the prevalence of alcohol or tobacco use disorders is lower in the chess population than in the population at large. The proposal plans to collect a large number of measures. What statistical safeguards will be taken against Type 1 errors? With respect to the proposed data analysis, there should be more detail about which analyses will be used for each specific hypothesis. More detail should also be provided about the regression analyses, in particular which variables will be used as predictors and criteria. Finally, in the ANOVAs with repeated measures, which action will be taken if the chess group and the control group differ on some variables on the pre-test? This often makes the results uninterpretable (see e.g. Figure 1 of Demily et al.'s, 2009, mentioned in the proposal, where the pre-test differences invalidate the authors' conclusions). Reference Sala, G., Aksayli, N. D., Tatlidil, K. S., Tatsumi, T., Gondo, Y., & Gobet, F. (2019). Near and far transfer in cognitive training: A second-order meta-analysis. *Collabra: Psychology, 5*, 18. https://doi.org/https://doi.org/10.1525/collabra.203

REVIEWER	Hammarberg, Anders Stockholms Lans Landsting, Centre for Dependency Disorders
REVIEW RETURNED	22-Feb-2022

GENERAL COMMENTS	The aim of this study is to investigate the efficacy of Chess-based cognitive remediation training (CRT) as therapy add-on in patients with alcohol and tobacco use disorders. I thank the authors for the opportunity to read and review this manuscript on an important subject. As the authors state in the introduction, relapse rates in treatment for substance use disorders are comparably high, and there is a need for development of new treatment strategies with the aim of helping affected individuals to reduce their substance use, with a reduction in related negative consequences as a result. The paper is well written and easy to read. The study has many merits, not the least in applying several different research methods (including fmri) to investigate research questions. However, in my opinion, there are several points that needs to be addressed more thoroughly before considering publication. General issues: - The study was registered in march 2019. The journal´s policy (as described in instructions for authors) is to primarily publish protocol manuscripts at an early stage of the study, and the fact that study

	by now ought to be at the end of enrolment might influence the decision to publish. - Study endpoints as described in the manuscript compared to description in registry (Clinical Trials Gov) differs in several respects. For example, in the manuscript it is the impression that primary endpoint is time to first relapse (page 13, line 53), while this endpoint is not at all stated in the registry (cue reactivity is stated as first outcome). If the primary endpoint has changed during study time, this should be addressed in the manuscript. - The former point is related to my concern related to the sample size calculation (page 14, line 14). In a power analysis it should be expressed which outcome measure that forms the basis of the calculation. Different measures most certainly requires different sample sizes in order to show effect, a fact proven not least in the present field of research, where studies more often find effects on measures related to cognitive functioning but not on drinking outcomes, and where power issues are sometimes measured as explanation (e.g. Khemiri et al., 2018). - In the section “Endpoints” (page 13), I guess that ”…the duration until the first severe relapse during the follow-up” is to interpreted as the primary endpoint, however this is not stated. Further, there is no definition of “severe relapse”. The study involves two different substances, for which definitions of relapse have differed considerably. This would need to be clarified. - In the introduction, I think the parts covering prevalence and treatment of SUD, as well as the current knowledge regarding the relation between SUD and aspects of cognition are well covered. Also, I generally agree on that several studies have proven that cognitive training improves cognitive functioning. However, I would say that I need to problematize the authors statement that cognitive training is “As an add-on therapy to treat substance use disorders CRT seems equally promising [35]. (page 7, line 53). My impression is that the research conducted on the effects of cognitive training on substance use related outcomes has been far from promising since 2013 (which is the year of publication of the referred to review). I have had difficulties in finding any study that has shown more than non-significant trends for effects on substance use outcomes. The authors do not provide any evidence for this statement and do not report negative findings for example Khemiri et al., 2018. This section should be expanded, since the proposed benefits of cognitive training on treatment outcomes (not only cognitive functioning) forms theoretical basis of the study. - My general concern with this study relates to the point above. Even if there still is a lack of studies investigating effects on cognitive training, there are some examples (referred to for example in the study by Khemiri et al 2018). They have in common that they involve quite small samples of patients, that they consistently find improvements in cognitive functioning among patients, and that they fail to show effects on substance use outcomes. I can not really see in what way the current study is a development of previous research in this respect, involving only small samples and also two patient groups which are expected to differ quite substantially regarding outcome measures, further complicating comparisons and synthesis of results. I think this question needs to be addressed more thoroughly in the introduction: In what way do the authors expect that this study contribute to existing knowledge? - Suggesting that the authors consider to add a more recent reference, involving cognitive training involving SUD-patients:

	Caetano et al. 2021 (Front Psych): Cognitive Training Effectiveness on Memory, Executive Functioning, and Processing Speed in Individuals With Substance Use Disorders: A Systematic Review. - Issues regarding Methods part: - Page 10, line 53: How is the chess-based intervention used in the study related to other chess-based interventions described in the literature (eg. Sala, G. and Gobet,F. (2016)? - Figure 2a-c. For a non-expert in the method (but being a frequent chess-player), these figures are not so informative in order for the reader to comprehend the method used. The authors could perhaps consider to add some lines describing the task to be solved. - Page 11, line 7: This sentence is complex and it is difficult to understand what is being teached: "…by explicitly teaching different concepts of cognitive functioning, questioning, and identifying the underlying cognitive process, and enhancing the awareness of before mentioned aspects."? Are patients informed about these processes in some form of psychoeducation intervention? If so, this should be more elaborately described, since the method then seems to involve more than "only" chess-skills training. - Page 11, lines 24 and onwards. The study described could/should rather be placed in the introduction section, perhaps in a section relating the method used in the study to previous studies involving CRT. - Page 12, line 13. It is stated that abstinence related goals are to be assessed. This seems to point to that abstinence are required? This is in many cases not the patients´ first choice, and might affect interpretation of (severe) relapse. This should be clarified. - Page 13, line 18. Regarding fmri assessments, the manuscript would benefit from a more detailed description regarding which regions are of interest. Which hypotheses are stated regarding this testing? Fmri generate a substantial amount of data, and the precision of the analyses might benefit from a specific research question (what specific regions does "salience" and "executive control" networks relate to?). - Page 14, line 14, sample size calculation. As previously mentioned, this section is somewhat sketchy. My main question is which endpoint that is chosen for the calculation of the sample size. In my opinion, this study in reality comprise two sub-studies involving two separate samples, with different main outcomes (alcohol and nicotine use), which would probably end up in different sample size requirements. The reader lack information regarding which outcome is expected to have an effect size of 0.2. It seems highly unlikable to me, that 24 participants per group would be enough to detect difference between groups (this opinion based on the fact that previous studies on cognitive training have been negative on drinking outcomes). But this is of course speculations from my side. - Page 15 – hypotheses part: The hypothesis again differ from outcomes describes in the "Endpoint section" as well as from the Gov registry page. For example, psychosocial functioning is mentioned here, but not previously. My suggestion is that these sections are harmonized to make it easier for the reader to follow what has been primary and secondary endpoints/hypotheses. - Lastly, related to previous points, it seems odd to have so many primary endpoints, and only one or two secondary. Usually, one

	single outcome is chosen as primary (based on theoretical considerations described in the Introduction), and the rest is secondary. This in order to avoid the risk of primary endpoints being chosen retrospectively, based on outcome of the study.

VERSION 1 – AUTHOR RESPONSE

Reviewer #1:
This protocol presents the design of a study using chess-based cognitive remediation training as therapy add-on for individuals with alcohol or tobacco use disorders. It is overall clearly written and in general presents sufficient details.
1. My main concern is that the wrong design is used. To control for placebo effects, you need to use a design with three groups, one of which is an active control group:
- treatment as usual
- treatment as usual + Chess-based cognitive remediation training
- treatment as usual + Some other novel activity [active control group]
Without this third group, it is logically impossible to draw any conclusions for possible differences between the first two groups after treatment, as chess specific effects and placebo effects are confounded.
Thank you for your thorough reading and your remark. Firstly, we do not aim to compare different trainings or therapy methods per se as we are mainly interested in neurobiological mechanisms. We aim to examine how neurobiological adaptations and resulting changes within the group receiving treatment as usual (TAU) + therapy add-on in comparison to TAU only can also predict relapse or abstinence. Secondly, and following the official regulations for obtaining funding money by the German Research Foundation, the present study design has already been peer reviewed and was accepted for funding and represents basic research, i.e. it is not included in the German Research Foundation funding program for clinical studies. Therefore, we cannot change the design anymore. Thirdly, we already included study participants since the submission of this study protocol. Changing the study design at the current time point is not feasible.

2. Another concern is the a priori likelihood that chess has positive effects (beyond placebo effects) is low. In general, research on cognitive training has found few, if any, reliable positive effects with respect to cognitive and scholastic outcomes (Sala et al., 2019). One exception is chess, but most of the research on this topic uses a two-group design (as in the proposal), which is insufficient to draw any reliable conclusion. When the correct three-group design was used, no effect was found, as noted in the proposal.
Thank you for this valuable comment. Our study aim is to evaluate neurobiological correlates of addictive behaviour that might be modified in another manner following chess as add-on therapy compared to standard treatment alone which might be also related to treatment outcome. Following the acceptation by the German Research Foundation, we planned the study accordingly and also started the conduct of the study following the peer-reviewed study proposal. We therefore are not able to change the study design in hindsight. However, we included your valuable point in the study protocol and discussed it as a potential limitation.
Please see the new section ‘Discussion‘ for more details
The here presented study aims to examine the effect of CB-CRT as treatment add-on on neurobiological processes but also neuropsychological and psychosocial functioning known to contribute to the development and maintenance of AUD and TUD. The effect of CB-CRT might also results in longer times of abstinence or reduced substance consumption. If CB-CRT as therapy add-on, as examined in this comprehensive study, shows to be more effective than standard treatment alone, this intervention might help to improve health behaviour in affected individuals.

Limitations with respect to the interpretability of the data might derive from the study design. We aim to examine the superior effect of CB-CRT compared to treatment as usual in therapy outcomes that might rely on neurobiological alterations following this training. As postulated by Sala and Gobet (98) a third, active control group might be needed to ultimately evaluate the chess-specific mechanisms and outcomes. Therefore and in case of successfully demonstrating a superior effect of our CB-CRT, a subsequent study might be needed to address this question. Further, even in light of our future results confirming a superior effect of CB-CRT as therapy add-on on neurobiological and neuropsychological processes, these improvements might to translate to longer abstinence or a reduction in the amount of substance consumption. Previously, this has been demonstrated in AUD: Even though an improvement in working memory functioning has been observed following an active working-memory training in patients with AUD, heavy drinking and neuropsychological functioning in other domains remained unchanged [39].
Since the described study includes a cognitive remediation training that exceeds merely training individual domains, we hope to counteract limitations of previous studies. Including social (training in the group) and metacognitive aspects, the CB-CRT might generalize from altering neurobiological processing to behavioural changes, i.e. substance consumption.]

3. The proposal mentions a few studies using chess as therapy, with encouraging results. However, all those studies use a weak design, so it's not possible to reach any trustworthy conclusion. Incidentally, the authors could make their case stronger if they could present data showing that the prevalence of alcohol or tobacco use disorders is lower in the chess population than in the population at large.

4. Thank you very much for this remark. To our knowledge, this question has not been addressed before. To close this research gap, we just started conducting a study on the prevalence of alcohol and tobacco use disorder amongst chess players, but cannot provide data at this stage. It should also be noted that in the study described in this protocol, we are using chess as a tool for cognitive training, but not using chess as a therapy.

5. The proposal plans to collect a large number of measures. What statistical safeguards will be taken against Type 1 errors? With respect to the proposed data analysis, there should be more detail about which analyses will be used for each specific hypothesis. More detail should also be provided about the regression analyses, in particular which variables will be used as predictors and criteria.
Thank you for this remark. We will control for multiple testing within distinct domains such as craving, working-memory, decision-making, or inhibitory control. Further, we will use established correction procedures, e.g., whole brain family-wise error correction (FWE) for fMRI analyses or Bonferroni correction for other statistical analyses.
Main outcomes define adaptations or changes in neurobiological functioning (whole brain measures) and in neuropsychological task (sum scores) that we expect to differ between treatments as usual (TAU) and TAU+chess-based therapy add on. We further expect an influence of these changes on therapy outcome, such as time to first relapse or amount of substance consumed in case of a relapse. The cox-regression will include brain activation (dorsolateral prefrontal or inferior frontal regions during inhibition, or the ventral striatum during cue reactivity tasks) as predictors for relapse.
Please see the corresponding section 'Data analysis plan'.
[Data analysis plan
To analyse psychometric and neuropsychological data, SPSS (Statistics for Windows, Version 2522.0. IBM Corp., Armonk, NY) will be used. The various dependent variables will be evaluated using multivariate analyses of variance with repeated measures. To counteract possible group differences at baseline, a percentage in change (divide by T1 values) or variable values at T1 can be incorporated in subsequent statistical analyses as a covariate. In addition, linear regression models will be calculated to examine the influence of confounding variables (for example, severity of tobacco or alcohol dependence) on the observed change in dependent variables as described previously (e.g.,

craving, task performance, psychosocial well-being). Cox-regression analyses, including, e.g., brain activation in the dorsolateral prefrontal or inferior frontal regions during inhibition and executive functioning, or the ventral striatum during cue reactivity tasks as predictors, will be conducted to examine the association with relapse. To analyse the fMRI data, SPM12 (Wellcome Department of Cognitive Neurology, London, UK) running under Matlab will be used. The pre-processing pipeline will include motion correction, normalization to the Montreal Neurological Institute (MNI) template, and a spatial smoothing with Gaussian kernel of 8 mm full width at half maximum (FWHM) will be conducted. The pre-processed data will then be used for first- and second-level analyses. On the first level (within-subject), neural activation associated with task conditions (contrasts) will be modelled via a convolution with a canonical hemodynamic response function (HFR) following a general linear model (GLM). A high-pass filter to remove low-frequency components of fMRI time-series will be used. Depending on the fMRI tasks, specific contrasts regarding task conditions will be modelledmodeled as described in the above cited literature. On the second level (between-subject) and regarding the effects of group and time, paired t-tests (e.g., pre vs. post intervention within one group) and full factorial models will be used. Additionally, regression models including clinical variables, such as severity of TUD or AUD, will be calculated. To control for multiple statistical testing, we will use established correction procedures, e.g., whole brain family-wise error correction (FWE) for fMRI analyses or Bonferroni correction for other statistical analyses the probability of a family wise error (FWE) will be set to .05.]

6. Finally, in the ANOVAs with repeated measures, which action will be taken if the chess group and the control group differ on some variables on the pre-test? This often makes the results uninterpretable (see e.g. Figure 1 of Demily et al.'s, 2009, mentioned in the proposal, where the pre-test differences invalidate the authors' conclusions).
Thank you for highlighting this issue. To counteract these possible statistical problems, we plan to, firstly, test for group differences in respective variables at T1. Following, a percentage in change (divide by T1) or variable values at T1 can be incorporated in subsequent statistical analyses. We revised this section accordingly. We revised this section accordingly (see comment above).

Reference:
Sala, G., Aksayli, N. D., Tatlidil, K. S., Tatsumi, T., Gondo, Y., & Gobet, F. (2019). Near and far transfer in cognitive training: A second-order meta-analysis. Collabra: Psychology, 5, 18. https://doi.org/https://doi.org/10.1525/collabra.203

Reviewer #2:
The aim of this study is to investigate the efficacy of Chess-based cognitive remediation training (CRT) as therapy add-on in patients with alcohol and tobacco use disorders. I thank the authors for the opportunity to read and review this manuscript on an important subject. As the authors state in the introduction, relapse rates in treatment for substance use disorders are comparably high, and there is a need for development of new treatment strategies with the aim of helping affected individuals to reduce their substance use, with a reduction in related negative consequences as a result. The paper is well written and easy to read. The study has many merits, not the least in applying several different research methods (including fmri) to investigate research questions. However, in my opinion, there are several points that needs to be addressed more thoroughly before considering publication.

General issues:
1. The study was registered in march 2019. The journal´s policy (as described in instructions for authors) is to primarily publish protocol manuscripts at an early stage of the study, and the fact that study by now ought to be at the end of enrolment might influence the decision to publish.
Thank you for this remark. When submitting this manuscript in fall 2021, we were still in the early stage of the study conduct (we included the first participant in March 2020). Due to COVID-restrictions

and several changes in research staff conducting the study, we only included a small number of participants by then. Since the submission of this study protocol, several month passed. Of course, we cannot stop the conduct of the study, since the financing party (German Research Foundation) also specifies a timeframe for the whole project and financing, that has to be respected.

2. Study endpoints as described in the manuscript compared to description in registry (Clinical Trials Gov) differs in several respects. For example, in the manuscript it is the impression that primary endpoint is time to first relapse (page 13, line 53), while this endpoint is not at all stated in the registry (cue reactivity is stated as first outcome). If the primary endpoint has changed during study time, this should be addressed in the manuscript.
Thank you for your thorough reading. The primary endpoint did not change. Please excuse the fact that our manuscript resulted in a confusion regarding this aspect. We revised the corresponding paragraphs accordingly.

3. The former point is related to my concern related to the sample size calculation (page 14, line 14). In a power analysis it should be expressed which outcome measure that forms the basis of the calculation. Different measures most certainly requires different sample sizes in order to show effect, a fact proven not least in the present field of research, where studies more often find effects on measures related to cognitive functioning but not on drinking outcomes, and where power issues are sometimes measured as explanation (e.g. Khemiri et al., 2018).
Thank you for this valuable thought and the suggested literature. As we restructured our endpoints and hypothesis to clarify out aims, we hope that the priority on underlying neural correlated becomes clearer. We agree, that the sample size calculation was restricted to one outcome measure (neurobiological changes). We also added your suggested literature to the study protocol and discussed this aspect in a critical manner. As already mentioned before, our study protocol already underwent peer review (German Research Foundation). We therefore cannot change the study design any more.
Sample size calculation:
Using the software package G*Power[96] the sample size calculation was conducted for the main primary outcomes, i.e., neurobiological correlates underlying adaptations following the CB-CRT, where we expected a minimum effect size of f = 0.2 for all constructs sample size was estimated assuming an effect size of f = 0.2 (ANOVA with repeated measures, within- and between subject factors and interactions). In this case, ideal sample coverage would be 24 individuals per group (at 80% power, alpha-level 5%).

4. In the section “Endpoints” (page 13), I guess that ”…the duration until the first severe relapse during the follow-up” is to interpreted as the primary endpoint, however this is not stated. Further, there is no definition of “severe relapse”. The study involves two different substances, for which definitions of relapse have differed considerably. This would need to be clarified.
Thank you for this remark. We revised sections regarding the endpoints following the comments of both reviewers. We defined ‘severe relapse’ ((daily smoking of at least one cigarette at day, consumption of more than 48 grams (females) or 60 grams (males) alcohol)) in the manuscript.

5. In the introduction, I think the parts covering prevalence and treatment of SUD, as well as the current knowledge regarding the relation between SUD and aspects of cognition are well covered. Also, I generally agree on that several studies have proven that cognitive training improves cognitive functioning. However, I would say that I need to problematize the authors statement that cognitive training is “As an add-on therapy to treat substance use disorders CRT seems equally promising [35]. (page 7, line 53). My impression is that the research conducted on the effects of cognitive training on substance use related outcomes has been far from promising since 2013 (which is the year of publication of the referred to review). I have had difficulties in finding any study that has shown more than non-significant trends for effects on substance use outcomes. The authors do not provide any

evidence for this statement and do not report negative findings for example Khemiri et al., 2018. This section should be expanded, since the proposed benefits of cognitive training on treatment outcomes (not only cognitive functioning) forms theoretical basis of the study.
Thank you for this valuable remark. We revised this paragraph and stated the findings more carefully. In addition, we added a discussion including potential limitations and previous negative findings.
Introduction:
[As an add-on therapy to treat substance use disorders CRT seems equally promising[35] and cognitive training mostly results in improvements within the respective domains [36]. However, there is a lack of studies examining the efficacy of CRT as a modulator of cognition to improve treatment outcomes[37] and findings on the positive outcome following cognitive trainings in AUD are still mixed [38] or not present [39]. A review on AUD[40] discussed that CRT improves split attention, recognition of warning signals, working memory, as well as episodic memory.]
Discussion:
[The here presented study aims to examine the effect of CB-CRT as treatment add-on on neurobiological processes but also neuropsychological and psychosocial functioning known to contribute to the development and maintenance of AUD and TUD. The effect of CB-CRT might also results in longer times of abstinence or reduced substance consumption. If CB-CRT as therapy add-on, as examined in this comprehensive study, shows to be more effective than standard treatment alone, this intervention might help to improve health behaviour in affected individuals.
Limitations with respect to the interpretability of the data might derive from the study design. We aim to examine the superior effect of CB-CRT compared to treatment as usual in therapy outcomes that might rely on neurobiological alterations following this training. As postulated by Sala and Gobet (98) a third, active control group might be needed to ultimately evaluate the chess-specific mechanisms and outcomes. Therefore and in case of successfully demonstrating a superior effect of our CB-CRT, a subsequent study might be needed to address this question. Further, even in light of our future results confirming a superior effect of CB-CRT as therapy add-on on neurobiological and neuropsychological processes, these improvements might to translate to longer abstinence or a reduction in the amount of substance consumption. Previously, this has been demonstrated in AUD: Even though an improvement in working memory functioning has been observed following an active working-memory training in patients with AUD, heavy drinking and neuropsychological functioning in other domains remained unchanged [39].
Since the described study includes a cognitive remediation training that exceeds merely training individual domains, we hope to counteract limitations of previous studies. Including social (training in the group) and metacognitive aspects, the CB-CRT might generalize from altering neurobiological processing to behavioural changes, i.e. substance consumption.]

6. My general concern with this study relates to the point above. Even if there still is a lack of studies investigating effects on cognitive training, there are some examples (referred to for example in the study by Khemiri et al 2018). They have in common that they involve quite small samples of patients, that they consistently find improvements in cognitive functioning among patients, and that they fail to show effects on substance use outcomes. I can not really see in what way the current study is a development of previous research in this respect, involving only small samples and also two patient groups which are expected to differ quite substantially regarding outcome measures, further complicating comparisons and synthesis of results. I think this question needs to be addressed more thoroughly in the introduction: In what way do the authors expect that this study contribute to existing knowledge?
Thank you for this follow-up question. We incorporated previous literature leading to a comprehensive study design that will examine the effect of chess-based cognitive remediation training on several aspects, namely neurobiological correlates of craving, executive functioning and inhibitory control, as well as neuropsychological and psychosocial functioning. Also, treatment outcomes such as abstinence and amount of substances consumed are of interest.

Not only the comprehensiveness of the study design, but also the remediation training that we use represent novel aspects. The social (group therapy) and metacognitive contents of the training might result in more robust findings compared to cognitive trainings focusing on only one cognitive domain, such as reported by Khemiri et al (working memory).
We specified our aim accordingly:
Consequently, our study aims to assess the effects of chess-based CRT (CB-CRT) on treatment outcomes and on underlying neurobiological mechanisms of CB-CRT in AUD and TUD.different aspects of cognition in individuals with AUD and TUD. We will use a novel and structured training program that, besides training cognitive functioning, includes metacognitive methods and social reinforcement. As a result of the comprehensiveness of the proposed study and the novel CB-CRT we will further, we assess the influence of CB-CRT on different aspects of cognition and psychosocial functioning as well as treatment outcome in individuals with AUD and TUD. underlying neurobiological mechanisms of CB-CRT in AUD and TUD also in relation to treatment outcome.

7. Suggesting that the authors consider to add a more recent reference, involving cognitive training involving SUD-patients: Caetano et al. 2021 (Front Psych): Cognitive Training Effectiveness on Memory, Executive Functioning, and Processing Speed in Individuals With Substance Use Disorders: A Systematic Review.
Thank you for this interesting literature. We incorporated it in the study protocol in the introduction.

8. Issues regarding Methods part:
- Page 10, line 53: How is the chess-based intervention used in the study related to other chess-based interventions described in the literature (eg. Sala, G. and Gobet,F. (2016)?
Thank you for this question. As described in ‘chess-based cognitive remediation training’ we conduct a training that is not equal to ‘playing chess’ per se. We follow a standardized manual including specific tasks appropriate to train a specific neurocognitive domains throughout one session (e.g., short-term memory, inhibition, planification skills). These tasks sometimes rely on strategies similar to chess as a game. All tasks use chess figures and the board as tools to train those skills. In addition, we train metacognitive functioning also by explicitly teaching the underlying neurocognitive mechanisms and the corresponding goals of each session and social reinforcement strategies will be applied. We revised this section and hope, that it is now clearer for the prospective reader.
[The training battery, which is administered in a group setting using mainly a chess demonstration board, is designed to strengthen cognitive functioning in specific domains such as selective attention (figure 2a), short-term memory (figure 2b), focal attention, pattern recognition, visuospatial abilities, planification skills (figure 2c), and inhibition. Participants do not need to know the game of chess. They will receive general information about the rules and strategies used for the corresponding training day. Overall, metacognitive abilities are trained as well, e.g., by explicitly teaching giving psychoeducational information regarding different concepts of cognitive functioning, questioning, and identifying the underlying cognitive process, and enhancing the awareness of before mentioned aspects. Participants perform most of the specific tasks in front of the group and, for a social reinforcement effect, everyone will applaud the respective participant.]
Following our study, we plan to publish this standardized manual in order for other researchers or medical, therapeutic staff to use it for subsequent studies or in therapeutic settings.
On a side note, an app for mobile devices resulted from our cooperation with the Spanish workgroup that developed the training (available at http://www.Gymchess.com). This app uses the same tasks as in the training battery used in this study but adapted for mobile use.

- Figure 2a-c. For a non-expert in the method (but being a frequent chess-player), these figures are not so informative in order for the reader to comprehend the method used. The authors could perhaps consider to add some lines describing the task to be solved.
Thank you for this remark. We already specified the solution or the corresponding way to solve the task in the figure legend similar to how we instruct our participants. Since this is a chess-‘based’

cognitive remediation training, no 1:1 comparison can be made between our tasks and an actual game of chess. We specified this in the corresponding section.

- Page 11, line 7: This sentence is complex and it is difficult to understand what is being teached: "…by explicitly teaching different concepts of cognitive functioning, questioning, and identifying the underlying cognitive process, and enhancing the awareness of before mentioned aspects."? Are patients informed about these processes in some form of psychoeducation intervention? If so, this should be more elaborately described, since the method then seems to involve more than "only" chess-skills training.
Thank you for this remark, we rephrased the sentence: Overall, metacognitive abilities are trained as well, e.g., by giving psychoeducational information regarding different concepts of cognitive functioning, questioning, and identifying the underlying cognitive process, and enhancing the awareness of before mentioned aspects.

- Page 11, lines 24 and onwards. The study described could/should rather be placed in the introduction section, perhaps in a section relating the method used in the study to previous studies involving CRT.
Thank you for this suggestion. Since the training and the corresponding study has not been published yet, we suggest to keep it in the section describing the training and not using this information as introductory and hypothesis-generating scientific information. Of course, if this argument would get in the way of publishing this protocol, we are willing to move it to the introduction section.

- Page 12, line 13. It is stated that abstinence related goals are to be assessed. This seems to point to that abstinence are required? This is in many cases not the patients´ first choice, and might affect interpretation of (severe) relapse. This should be clarified.
Thank you for this remark. The goal attainment scale used here includes individual's goals, which could also be a reduction of smoking. It then assesses the satisfaction with respect to this individual goal.

- Page 13, line 18. Regarding fmri assessments, the manuscript would benefit from a more detailed description regarding which regions are of interest. Which hypotheses are stated regarding this testing? Fmri generate a substantial amount of data, and the precision of the analyses might benefit from a specific research question (what specific regions does "salience" and "executive control" networks relate to?).
Thank you for this remark. We specified corresponding brain regions in the 'Endpoints' section, since the 'fMRI assessment section rather focuses on methodological issues.
[Endpoints are changes in neural alcohol and tobacco cue-reactivity[80, 90] (e.g., reduction in substance-related activation of striatal brain regions), neural correlates of inhibition (stop-signal task)[93] (e.g., increased dorsolateral prefrontal neural activation) and working memory (N-back task)[91] (e.g., increased inferior frontal neural activation), as well as functional connectivity within the salience network (SN; insula, anterior cingulate cortex) and executive control network (ECN; dorsolateral frontal and lateral posterior parietal cortices) using resting-state fMRI data.]

- Page 14, line 14, sample size calculation. As previously mentioned, this section is somewhat sketchy. My main question is which endpoint that is chosen for the calculation of the sample size. In my opinion, this study in reality comprise two sub-studies involving two separate samples, with different main outcomes (alcohol and nicotine use), which would probably end up in different sample size requirements. The reader lack information regarding which outcome is expected to have an effect size of 0.2. It seems highly unlikable to me, that 24 participants per group would be enough to detect difference between groups (this opinion based on the fact that previous studies on cognitive training have been negative on drinking outcomes). But this is of course speculations from my side.

Thank you for this follow-up question. We revised this section. Please see our answer to your question #3 for more details.

- Page 15 – hypotheses part: The hypothesis again differ from outcomes describes in the “Endpoint section” as well as from the Gov registry page. For example, psychosocial functioning is mentioned here, but not previously. My suggestion is that these sections are harmonized to make it easier for the reader to follow what has been primary and secondary endpoints/hypotheses
Thank you for this suggestion. We revised the corresponding paragraphs accordingly.

- Lastly, related to previous points, it seems odd to have so many primary endpoints, and only one or two secondary. Usually, one single outcome is chosen as primary (based on theoretical considerations described in the Introduction), and the rest is secondary. This in order to avoid the risk of primary endpoints being chosen retrospectively, based on outcome of the study.
Thank you for this remark. As we defined and published (clinicaltrials.gov) the endpoints beforehand at, we would suggest keeping it that way. Please see our revised manuscript for a clearer description of endpoints since we agree that our description lacked consistency.

VERSION 2 – REVIEW

REVIEWER	Hammarberg, Anders Stockholms Lans Landsting, Centre for Dependency Disorders
REVIEW RETURNED	25-May-2022

GENERAL COMMENTS	Reviewer #2: The aim of this study is to investigate the efficacy of Chess-based cognitive remediation training (CRT) as therapy add-on in patients with alcohol and tobacco use disorders. I thank the authors for the opportunity to read and review this manuscript on an important subject. As the authors state in the introduction, relapse rates in treatment for substance use disorders are comparably high, and there is a need for development of new treatment strategies with the aim of helping affected individuals to reduce their substance use, with a reduction in related negative consequences as a result. The paper is well written and easy to read. The study has many merits, not the least in applying several different research methods (including fmri) to investigate research questions. However, in my opinion, there are several points that needs to be addressed more thoroughly before considering publication. General issues: 1. The study was registered in march 2019. The journal´s policy (as described in instructions for authors) is to primarily publish protocol manuscripts at an early stage of the study, and the fact that study by now ought to be at the end of enrolment might influence the decision to publish. Thank you for this remark. When submitting this manuscript in fall 2021, we were still in the early stage of the study conduct (we included the first participant in March 2020). Due to COVID-restrictions and several changes in research staff conducting the study, we only included a small number of participants by then. Since the submission of this study protocol, several month passed. Of course, we cannot stop the conduct of the study, since the financing party (German Research Foundation) also specifies a timeframe for the whole project and financing, that has to be respected.

	Comment: I think the authors has given a satisfactory answer. 2. Study endpoints as described in the manuscript compared to description in registry (Clinical Trials Gov) differs in several respects. For example, in the manuscript it is the impression that primary endpoint is time to first relapse (page 13, line 53), while this endpoint is not at all stated in the registry (cue reactivity is stated as first outcome). If the primary endpoint has changed during study time, this should be addressed in the manuscript. Thank you for your thorough reading. The primary endpoint did not change. Please excuse the fact that our manuscript resulted in a confusion regarding this aspect. We revised the corresponding paragraphs accordingly. Comment: I think the authors have revised the manuscript satisfactorily. 3. The former point is related to my concern related to the sample size calculation (page 14, line 14). In a power analysis it should be expressed which outcome measure that forms the basis of the calculation. Different measures most certainly requires different sample sizes in order to show effect, a fact proven not least in the present field of research, where studies more often find effects on measures related to cognitive functioning but not on drinking outcomes, and where power issues are sometimes measured as explanation (e.g. Khemiri et al., 2018). Thank you for this valuable thought and the suggested literature. As we restructured our endpoints and hypothesis to clarify out aims, we hope that the priority on underlying neural correlated becomes clearer. We agree, that the sample size calculation was restricted to one outcome measure (neurobiological changes). We also added your suggested literature to the study protocol and discussed this aspect in a critical manner. As already mentioned before, our study protocol already underwent peer review (German Research Foundation). We therefore cannot change the study design any more. Sample size calculation: Using the software package G*Power[96] the sample size calculation was conducted for the main primary outcomes, i.e., neurobiological correlates underlying adaptations following the CB-CRT, where we expected a minimum effect size of f = 0.2 for all constructs sample size was estimated assuming an effect size of f = 0.2 (ANOVA with repeated measures, within- and between subject factors and interactions). In this case, ideal sample coverage would be 24 individuals per group (at 80% power, alpha-level 5%). Comment: I think the authors have revised the script satisfactorily. The more precise definition of primary endpoint clarified this point. 4. In the section “Endpoints” (page 13), I guess that ”…the duration until the first severe relapse during the follow-up” is to interpreted as the primary endpoint, however this is not stated. Further, there is no definition of “severe relapse”. The study involves two different substances, for which definitions of relapse have differed considerably. This would need to be clarified. Thank you for this remark. We revised sections regarding the endpoints following the comments of both reviewers. We defined ‘severe relapse’ ((daily smoking of at least one cigarette at day,

	consumption of more than 48 grams (females) or 60 grams (males) alcohol)) in the manuscript. Comment: I think the authors have revised the manuscript satisfactorily, even though I question the power to detect effects in this regard du to small sample sizes. On the other hand, these are not primary endpoints, and the analyses are hence of an exploratory nature. 5. In the introduction, I think the parts covering prevalence and treatment of SUD, as well as the current knowledge regarding the relation between SUD and aspects of cognition are well covered. Also, I generally agree on that several studies have proven that cognitive training improves cognitive functioning. However, I would say that I need to problematize the authors statement that cognitive training is "As an add-on therapy to treat substance use disorders CRT seems equally promising [35]. (page 7, line 53). My impression is that the research conducted on the effects of cognitive training on substance use related outcomes has been far from promising since 2013 (which is the year of publication of the referred to review). I have had difficulties in finding any study that has shown more than non-significant trends for effects on substance use outcomes. The authors do not provide any evidence for this statement and do not report negative findings for example Khemiri et al., 2018. This section should be expanded, since the proposed benefits of cognitive training on treatment outcomes (not only cognitive functioning) forms theoretical basis of the study. Thank you for this valuable remark. We revised this paragraph and stated the findings more carefully. In addition, we added a discussion including potential limitations and previous negative findings. Introduction: [As an add-on therapy to treat substance use disorders CRT seems equally promising[35] and cognitive training mostly results in improvements within the respective domains [36]. However, there is a lack of studies examining the efficacy of CRT as a modulator of cognition to improve treatment outcomes[37] and findings on the positive outcome following cognitive trainings in AUD are still mixed [38] or not present [39]. A review on AUD[40] discussed that CRT improves split attention, recognition of warning signals, working memory, as well as episodic memory.] Discussion: [The here presented study aims to examine the effect of CB-CRT as treatment add-on on neurobiological processes but also neuropsychological and psychosocial functioning known to contribute to the development and maintenance of AUD and TUD. The effect of CB-CRT might also results in longer times of abstinence or reduced substance consumption. If CB-CRT as therapy add-on, as examined in this comprehensive study, shows to be more effective than standard treatment alone, this intervention might help to improve health behaviour in affected individuals. Limitations with respect to the interpretability of the data might derive from the study design. We aim to examine the superior effect of CB-CRT compared to treatment as usual in therapy outcomes that might rely on neurobiological alterations following this training. As postulated by Sala and Gobet (98) a third, active control group might be needed to ultimately evaluate the chess-specific mechanisms and outcomes. Therefore and in case of

	successfully demonstrating a superior effect of our CB-CRT, a subsequent study might be needed to address this question. Further, even in light of our future results confirming a superior effect of CB-CRT as therapy add-on on neurobiological and neuropsychological processes, these improvements might to translate to longer abstinence or a reduction in the amount of substance consumption. Previously, this has been demonstrated in AUD: Even though an improvement in working memory functioning has been observed following an active working-memory training in patients with AUD, heavy drinking and neuropsychological functioning in other domains remained unchanged [39]. Since the described study includes a cognitive remediation training that exceeds merely training individual domains, we hope to counteract limitations of previous studies. Including social (training in the group) and metacognitive aspects, the CB-CRT might generalize from altering neurobiological processing to behavioural changes, i.e. substance consumption.] Comment: I think the authors have revised the manuscript satisfactorily. 6. My general concern with this study relates to the point above. Even if there still is a lack of studies investigating effects on cognitive training, there are some examples (referred to for example in the study by Khemiri et al 2018). They have in common that they involve quite small samples of patients, that they consistently find improvements in cognitive functioning among patients, and that they fail to show effects on substance use outcomes. I can not really see in what way the current study is a development of previous research in this respect, involving only small samples and also two patient groups which are expected to differ quite substantially regarding outcome measures, further complicating comparisons and synthesis of results. I think this question needs to be addressed more thoroughly in the introduction: In what way do the authors expect that this study contribute to existing knowledge? Thank you for this follow-up question. We incorporated previous literature leading to a comprehensive study design that will examine the effect of chess-based cognitive remediation training on several aspects, namely neurobiological correlates of craving, executive functioning and inhibitory control, as well as neuropsychological and psychosocial functioning. Also, treatment outcomes such as abstinence and amount of substances consumed are of interest. Not only the comprehensiveness of the study design, but also the remediation training that we use represent novel aspects. The social (group therapy) and metacognitive contents of the training might result in more robust findings compared to cognitive trainings focusing on only one cognitive domain, such as reported by Khemiri et al (working memory). We specified our aim accordingly: Consequently, our study aims to assess the effects of chess-based CRT (CB-CRT) on treatment outcomes and on underlying neurobiological mechanisms of CB-CRT in AUD and TUD.different aspects of cognition in individuals with AUD and TUD. We will use a novel and structured training program that, besides training cognitive functioning, includes metacognitive methods and social reinforcement. As a result of the comprehensiveness of the

	proposed study and the novel CB-CRT we will further, we assess the influence of CB-CRT on different aspects of cognition and psychosocial functioning as well as treatment outcome in individuals with AUD and TUD. underlying neurobiological mechanisms of CB-CRT in AUD and TUD also in relation to treatment outcome. Comment: I think the authors have revised the manuscript satisfactorily. The question remains open. 7. Suggesting that the authors consider to add a more recent reference, involving cognitive training involving SUD-patients: Caetano et al. 2021 (Front Psych): Cognitive Training Effectiveness on Memory, Executive Functioning, and Processing Speed in Individuals With Substance Use Disorders: A Systematic Review. Thank you for this interesting literature. We incorporated it in the study protocol in the introduction. Comment: Great! 8. Issues regarding Methods part: - Page 10, line 53: How is the chess-based intervention used in the study related to other chess-based interventions described in the literature (eg. Sala, G. and Gobet,F. (2016)? Thank you for this question. As described in 'chess-based cognitive remediation training' we conduct a training that is not equal to 'playing chess' per se. We follow a standardized manual including specific tasks appropriate to train a specific neurocognitive domains throughout one session (e.g., short-term memory, inhibition, planification skills). These tasks sometimes rely on strategies similar to chess as a game. All tasks use chess figures and the board as tools to train those skills. In addition, we train metacognitive functioning also by explicitly teaching the underlying neurocognitive mechanisms and the corresponding goals of each session and social reinforcement strategies will be applied. We revised this section and hope, that it is now clearer for the prospective reader. [The training battery, which is administered in a group setting using mainly a chess demonstration board, is designed to strengthen cognitive functioning in specific domains such as selective attention (figure 2a), short-term memory (figure 2b), focal attention, pattern recognition, visuospatial abilities, planification skills (figure 2c), and inhibition. Participants do not need to know the game of chess. They will receive general information about the rules and strategies used for the corresponding training day. Overall, metacognitive abilities are trained as well, e.g., by explicitly teaching giving psychoeducational information regarding different concepts of cognitive functioning, questioning, and identifying the underlying cognitive process, and enhancing the awareness of before mentioned aspects. Participants perform most of the specific tasks in front of the group and, for a social reinforcement effect, everyone will applaud the respective participant.] Following our study, we plan to publish this standardized manual in order for other researchers or medical, therapeutic staff to use it for subsequent studies or in therapeutic settings.

	On a side note, an app for mobile devices resulted from our cooperation with the Spanish workgroup that developed the training (available at http://www.Gymchess.com). This app uses the same tasks as in the training battery used in this study but adapted for mobile use. Comment: I think the authors have revised the manuscript satisfactorily. - Figure 2a-c. For a non-expert in the method (but being a frequent chess-player), these figures are not so informative in order for the reader to comprehend the method used. The authors could perhaps consider to add some lines describing the task to be solved. Thank you for this remark. We already specified the solution or the corresponding way to solve the task in the figure legend similar to how we instruct our participants. Since this is a chess-'based' cognitive remediation training, no 1:1 comparison can be made between our tasks and an actual game of chess. We specified this in the corresponding section. Comment: I think the authors have revised the manuscript satisfactorily. - Page 11, line 7: This sentence is complex and it is difficult to understand what is being teached: "…by explicitly teaching different concepts of cognitive functioning, questioning, and identifying the underlying cognitive process, and enhancing the awareness of before mentioned aspects."? Are patients informed about these processes in some form of psychoeducation intervention? If so, this should be more elaborately described, since the method then seems to involve more than "only" chess-skills training. Thank you for this remark, we rephrased the sentence: Overall, metacognitive abilities are trained as well, e.g., by giving psychoeducational information regarding different concepts of cognitive functioning, questioning, and identifying the underlying cognitive process, and enhancing the awareness of before mentioned aspects. Comment: I think the text is more comprehensible now. - Page 11, lines 24 and onwards. The study described could/should rather be placed in the introduction section, perhaps in a section relating the method used in the study to previous studies involving CRT. Thank you for this suggestion. Since the training and the corresponding study has not been published yet, we suggest to keep it in the section describing the training and not using this information as introductory and hypothesis-generating scientific information. Of course, if this argument would get in the way of publishing this protocol, we are willing to move it to the introduction section. Comment: Thank you for your answer – for me the text can remain as in the present form.

	- Page 12, line 13. It is stated that abstinence related goals are to be assessed. This seems to point to that abstinence are required? This is in many cases not the patients´ first choice, and might affect interpretation of (severe) relapse. This should be clarified. Thank you for this remark. The goal attainment scale used here includes individual's goals, which could also be a reduction of smoking. It then assesses the satisfaction with respect to this individual goal. Comment: I think the authors have given a satisfactory answer. - Page 13, line 18. Regarding fmri assessments, the manuscript would benefit from a more detailed description regarding which regions are of interest. Which hypotheses are stated regarding this testing? Fmri generate a substantial amount of data, and the precision of the analyses might benefit from a specific research question (what specific regions does "salience" and "executive control" networks relate to?). Thank you for this remark. We specified corresponding brain regions in the 'Endpoints' section, since the 'fMRI assessment section rather focuses on methodological issues. [Endpoints are changes in neural alcohol and tobacco cue-reactivity[80, 90] (e.g., reduction in substance-related activation of striatal brain regions), neural correlates of inhibition (stop-signal task)[93] (e.g., increased dorsolateral prefrontal neural activation) and working memory (N-back task)[91] (e.g., increased inferior frontal neural activation), as well as functional connectivity within the salience network (SN; insula, anterior cingulate cortex) and executive control network (ECN; dorsolateral frontal and lateral posterior parietal cortices) using resting-state fMRI data.] Comment: I think the authors have given a satisfactory answer. - Page 14, line 14, sample size calculation. As previously mentioned, this section is somewhat sketchy. My main question is which endpoint that is chosen for the calculation of the sample size. In my opinion, this study in reality comprise two sub-studies involving two separate samples, with different main outcomes (alcohol and nicotine use), which would probably end up in different sample size requirements. The reader lack information regarding which outcome is expected to have an effect size of 0.2. It seems highly unlikable to me, that 24 participants per group would be enough to detect difference between groups (this opinion based on the fact that previous studies on cognitive training have been negative on drinking outcomes). But this is of course speculations from my side. Thank you for this follow-up question. We revised this section. Please see our answer to your question #3 for more details. Comment: I think the authors have revised the manuscript satisfactorily. - Page 15 – hypotheses part: The hypothesis again differ from outcomes describes in the "Endpoint section" as well as from the Gov registry page. For example, psychosocial functioning is mentioned here, but not previously. My suggestion is that these sections are harmonized to make it easier for the reader to follow what has been primary and secondary endpoints/hypotheses Thank you for this suggestion. We revised the corresponding paragraphs accordingly.

	Comment: I think the authors have revised the manuscript satisfactorily. - Lastly, related to previous points, it seems odd to have so many primary endpoints, and only one or two secondary. Usually, one single outcome is chosen as primary (based on theoretical considerations described in the Introduction), and the rest is secondary. This in order to avoid the risk of primary endpoints being chosen retrospectively, based on outcome of the study. Thank you for this remark. As we defined and published (clinicaltrials.gov) the endpoints beforehand at, we would suggest keeping it that way. Please see our revised manuscript for a clearer description of endpoints since we agree that our description lacked consistency. Comment: I think the authors have revised the manuscript satisfactorily.

VERSION 2 – AUTHOR RESPONSE

Reviewer #2:
As a comment to the reviewer's remarks: Thank you for your positive evaluation regarding all points.

The aim of this study is to investigate the efficacy of Chess-based cognitive remediation training (CRT) as therapy add-on in patients with alcohol and tobacco use disorders. I thank the authors for the opportunity to read and review this manuscript on an important subject. As the authors state in the introduction, relapse rates in treatment for substance use disorders are comparably high, and there is a need for development of new treatment strategies with the aim of helping affected individuals to reduce their substance use, with a reduction in related negative consequences as a result. The paper is well written and easy to read. The study has many merits, not the least in applying several different research methods (including fmri) to investigate research questions. However, in my opinion, there are several points that needs to be addressed more thoroughly before considering publication.
General issues:
1. The study was registered in march 2019. The journal´s policy (as described in instructions for authors) is to primarily publish protocol manuscripts at an early stage of the study, and the fact that study by now ought to be at the end of enrolment might influence the decision to publish.
Thank you for this remark. When submitting this manuscript in fall 2021, we were still in the early stage of the study conduct (we included the first participant in March 2020). Due to COVID-restrictions and several changes in research staff conducting the study, we only included a small number of participants by then. Since the submission of this study protocol, several month passed. Of course, we cannot stop the conduct of the study, since the financing party (German Research Foundation) also specifies a timeframe for the whole project and financing, that has to be respected.

Comment: I think the authors has given a satisfactory answer.

2. Study endpoints as described in the manuscript compared to description in registry (Clinical Trials Gov) differs in several respects. For example, in the manuscript it is the impression that primary endpoint is time to first relapse (page 13, line 53), while this endpoint is not at all stated in the registry (cue reactivity is stated as first outcome). If the primary endpoint has changed during study time, this should be addressed in the manuscript.

Thank you for your thorough reading. The primary endpoint did not change. Please excuse the fact that our manuscript resulted in a confusion regarding this aspect. We revised the corresponding paragraphs accordingly.

Comment: I think the authors have revised the manuscript satisfactorily.

3. The former point is related to my concern related to the sample size calculation (page 14, line 14). In a power analysis it should be expressed which outcome measure that forms the basis of the calculation. Different measures most certainly requires different sample sizes in order to show effect, a fact proven not least in the present field of research, where studies more often find effects on measures related to cognitive functioning but not on drinking outcomes, and where power issues are sometimes measured as explanation (e.g. Khemiri et al., 2018).

Thank you for this valuable thought and the suggested literature. As we restructured our endpoints and hypothesis to clarify out aims, we hope that the priority on underlying neural correlated becomes clearer. We agree, that the sample size calculation was restricted to one outcome measure (neurobiological changes). We also added your suggested literature to the study protocol and discussed this aspect in a critical manner. As already mentioned before, our study protocol already underwent peer review (German Research Foundation). We therefore cannot change the study design any more.

Sample size calculation:

Using the software package G*Power[96] the sample size calculation was conducted for the main primary outcomes, i.e., neurobiological correlates underlying adaptations following the CB-CRT, where we expected a minimum effect size of f = 0.2 for all constructs sample size was estimated assuming an effect size of f = 0.2 (ANOVA with repeated measures, within- and between subject factors and interactions). In this case, ideal sample coverage would be 24 individuals per group (at 80% power, alpha-level 5%).

Comment: I think the authors have revised the script satisfactorily. The more precise definition of primary endpoint clarified this point.

4. In the section "Endpoints" (page 13), I guess that "…the duration until the first severe relapse during the follow-up" is to interpreted as the primary endpoint, however this is not stated. Further, there is no definition of "severe relapse". The study involves two different substances, for which definitions of relapse have differed considerably. This would need to be clarified.

Thank you for this remark. We revised sections regarding the endpoints following the comments of both reviewers. We defined 'severe relapse' ((daily smoking of at least one cigarette at day, consumption of more than 48 grams (females) or 60 grams (males) alcohol)) in the manuscript.

Comment: I think the authors have revised the manuscript satisfactorily, even though I question the power to detect effects in this regard du to small sample sizes. On the other hand, these are not primary endpoints, and the analyses are hence of an exploratory nature.

5. In the introduction, I think the parts covering prevalence and treatment of SUD, as well as the current knowledge regarding the relation between SUD and aspects of cognition are well covered. Also, I generally agree on that several studies have proven that cognitive training improves cognitive functioning. However, I would say that I need to problematize the authors statement that cognitive training is "As an add-on therapy to treat substance use disorders CRT seems equally promising [35]. (page 7, line 53). My impression is that the research conducted on the effects of cognitive training on substance use related outcomes has been far from promising since 2013 (which is the year of publication of the referred to review). I have had difficulties in finding any study that has shown more than non-significant trends for effects on substance use outcomes. The authors do not provide any evidence for this statement and do not report negative findings for example Khemiri et al., 2018. This section should be expanded, since the proposed benefits of cognitive training on treatment outcomes (not only cognitive functioning) forms theoretical basis of the study.

Thank you for this valuable remark. We revised this paragraph and stated the findings more carefully. In addition, we added a discussion including potential limitations and previous negative findings.

Introduction:

[As an add-on therapy to treat substance use disorders CRT seems equally promising[35] and cognitive training mostly results in improvements within the respective domains [36]. However, there is a lack of studies examining the efficacy of CRT as a modulator of cognition to improve treatment outcomes[37] and findings on the positive outcome following cognitive trainings in AUD are still mixed [38] or not present [39]. A review on AUD[40] discussed that CRT improves split attention, recognition of warning signals, working memory, as well as episodic memory.]
Discussion:
[The here presented study aims to examine the effect of CB-CRT as treatment add-on on neurobiological processes but also neuropsychological and psychosocial functioning known to contribute to the development and maintenance of AUD and TUD. The effect of CB-CRT might also results in longer times of abstinence or reduced substance consumption. If CB-CRT as therapy add-on, as examined in this comprehensive study, shows to be more effective than standard treatment alone, this intervention might help to improve health behaviour in affected individuals.
Limitations with respect to the interpretability of the data might derive from the study design. We aim to examine the superior effect of CB-CRT compared to treatment as usual in therapy outcomes that might rely on neurobiological alterations following this training. As postulated by Sala and Gobet (98) a third, active control group might be needed to ultimately evaluate the chess-specific mechanisms and outcomes. Therefore and in case of successfully demonstrating a superior effect of our CB-CRT, a subsequent study might be needed to address this question. Further, even in light of our future results confirming a superior effect of CB-CRT as therapy add-on on neurobiological and neuropsychological processes, these improvements might to translate to longer abstinence or a reduction in the amount of substance consumption. Previously, this has been demonstrated in AUD: Even though an improvement in working memory functioning has been observed following an active working-memory training in patients with AUD, heavy drinking and neuropsychological functioning in other domains remained unchanged [39].
Since the described study includes a cognitive remediation training that exceeds merely training individual domains, we hope to counteract limitations of previous studies. Including social (training in the group) and metacognitive aspects, the CB-CRT might generalize from altering neurobiological processing to behavioural changes, i.e. substance consumption.]

Comment: I think the authors have revised the manuscript satisfactorily.

6. My general concern with this study relates to the point above. Even if there still is a lack of studies investigating effects on cognitive training, there are some examples (referred to for example in the study by Khemiri et al 2018). They have in common that they involve quite small samples of patients, that they consistently find improvements in cognitive functioning among patients, and that they fail to show effects on substance use outcomes. I can not really see in what way the current study is a development of previous research in this respect, involving only small samples and also two patient groups which are expected to differ quite substantially regarding outcome measures, further complicating comparisons and synthesis of results. I think this question needs to be addressed more thoroughly in the introduction: In what way do the authors expect that this study contribute to existing knowledge?
Thank you for this follow-up question. We incorporated previous literature leading to a comprehensive study design that will examine the effect of chess-based cognitive remediation training on several aspects, namely neurobiological correlates of craving, executive functioning and inhibitory control, as well as neuropsychological and psychosocial functioning. Also, treatment outcomes such as abstinence and amount of substances consumed are of interest.
Not only the comprehensiveness of the study design, but also the remediation training that we use represent novel aspects. The social (group therapy) and metacognitive contents of the training might result in more robust findings compared to cognitive trainings focusing on only one cognitive domain, such as reported by Khemiri et al (working memory).
We specified our aim accordingly:
Consequently, our study aims to assess the effects of chess-based CRT (CB-CRT) on treatment outcomes and on underlying neurobiological mechanisms of CB-CRT in AUD and TUD.different aspects of cognition in individuals with AUD and TUD. We will use a novel and structured training program that, besides training cognitive functioning, includes metacognitive methods and social reinforcement. As a result of the comprehensiveness of the proposed study and the novel CB-CRT we will further, we assess the influence of CB-CRT on different aspects of cognition and psychosocial

functioning as well as treatment outcome in individuals with AUD and TUD. underlying neurobiological mechanisms of CB-CRT in AUD and TUD also in relation to treatment outcome.

Comment: I think the authors have revised the manuscript satisfactorily. The question remains open.

7. Suggesting that the authors consider to add a more recent reference, involving cognitive training involving SUD-patients: Caetano et al. 2021 (Front Psych): Cognitive Training Effectiveness on Memory, Executive Functioning, and Processing Speed in Individuals With Substance Use Disorders: A Systematic Review.
Thank you for this interesting literature. We incorporated it in the study protocol in the introduction.

Comment: Great!

8. Issues regarding Methods part:
- Page 10, line 53: How is the chess-based intervention used in the study related to other chess-based interventions described in the literature (eg. Sala, G. and Gobet,F. (2016)?
Thank you for this question. As described in 'chess-based cognitive remediation training' we conduct a training that is not equal to 'playing chess' per se. We follow a standardized manual including specific tasks appropriate to train a specific neurocognitive domains throughout one session (e.g., short-term memory, inhibition, planification skills). These tasks sometimes rely on strategies similar to chess as a game. All tasks use chess figures and the board as tools to train those skills. In addition, we train metacognitive functioning also by explicitly teaching the underlying neurocognitive mechanisms and the corresponding goals of each session and social reinforcement strategies will be applied. We revised this section and hope, that it is now clearer for the prospective reader.
[The training battery, which is administered in a group setting using mainly a chess demonstration board, is designed to strengthen cognitive functioning in specific domains such as selective attention (figure 2a), short-term memory (figure 2b), focal attention, pattern recognition, visuospatial abilities, planification skills (figure 2c), and inhibition. Participants do not need to know the game of chess. They will receive general information about the rules and strategies used for the corresponding training day. Overall, metacognitive abilities are trained as well, e.g., by explicitly teaching giving psychoeducational information regarding different concepts of cognitive functioning, questioning, and identifying the underlying cognitive process, and enhancing the awareness of before mentioned aspects. Participants perform most of the specific tasks in front of the group and, for a social reinforcement effect, everyone will applaud the respective participant.] Following our study, we plan to publish this standardized manual in order for other researchers or medical, therapeutic staff to use it for subsequent studies or in therapeutic settings.
On a side note, an app for mobile devices resulted from our cooperation with the Spanish workgroup that developed the training (available at https://smex-ctp.trendmicro.com:443/wis/clicktime/v1/query?url=http%3a%2f%2fwww.Gymchess.com&umid=35dd25f4-2969-45c6-b71b-3513953b385f&auth=9bdbfb691f9c6334b2ba2eebd6a53fe50f3dc0ca-59e3a0d12fc3280a0f81b8284fe863d414b8f5f4). This app uses the same tasks as in the training battery used in this study but adapted for mobile use.

Comment: I think the authors have revised the manuscript satisfactorily.

- Figure 2a-c. For a non-expert in the method (but being a frequent chess-player), these figures are not so informative in order for the reader to comprehend the method used. The authors could perhaps consider to add some lines describing the task to be solved.
Thank you for this remark. We already specified the solution or the corresponding way to solve the task in the figure legend similar to how we instruct our participants. Since this is a chess-'based' cognitive remediation training, no 1:1 comparison can be made between our tasks and an actual game of chess. We specified this in the corresponding section.

Comment: I think the authors have revised the manuscript satisfactorily.

- Page 11, line 7: This sentence is complex and it is difficult to understand what is being teached: "…by explicitly teaching different concepts of cognitive functioning, questioning, and identifying the underlying cognitive process, and enhancing the awareness of before mentioned aspects."? Are patients informed about these processes in some form of psychoeducation intervention? If so, this should be more elaborately described, since the method then seems to involve more than "only" chess-skills training.
Thank you for this remark, we rephrased the sentence: Overall, metacognitive abilities are trained as well, e.g., by giving psychoeducational information regarding different concepts of cognitive functioning, questioning, and identifying the underlying cognitive process, and enhancing the awareness of before mentioned aspects.

Comment: I think the text is more comprehensible now.

- Page 11, lines 24 and onwards. The study described could/should rather be placed in the introduction section, perhaps in a section relating the method used in the study to previous studies involving CRT.
Thank you for this suggestion. Since the training and the corresponding study has not been published yet, we suggest to keep it in the section describing the training and not using this information as introductory and hypothesis-generating scientific information. Of course, if this argument would get in the way of publishing this protocol, we are willing to move it to the introduction section.

Comment: Thank you for your answer – for me the text can remain as in the present form.

- Page 12, line 13. It is stated that abstinence related goals are to be assessed. This seems to point to that abstinence are required? This is in many cases not the patients´ first choice, and might affect interpretation of (severe) relapse. This should be clarified.
Thank you for this remark. The goal attainment scale used here includes individual's goals, which could also be a reduction of smoking. It then assesses the satisfaction with respect to this individual goal.

Comment: I think the authors have given a satisfactory answer.

- Page 13, line 18. Regarding fmri assessments, the manuscript would benefit from a more detailed description regarding which regions are of interest. Which hypotheses are stated regarding this testing? Fmri generate a substantial amount of data, and the precision of the analyses might benefit from a specific research question (what specific regions does "salience" and "executive control" networks relate to?).
Thank you for this remark. We specified corresponding brain regions in the 'Endpoints' section, since the 'fMRI assessment section rather focuses on methodological issues.
[Endpoints are changes in neural alcohol and tobacco cue-reactivity[80, 90] (e.g., reduction in substance-related activation of striatal brain regions), neural correlates of inhibition (stop-signal task)[93] (e.g., increased dorsolateral prefrontal neural activation) and working memory (N-back task)[91] (e.g., increased inferior frontal neural activation), as well as functional connectivity within the salience network (SN; insula, anterior cingulate cortex) and executive control network (ECN; dorsolateral frontal and lateral posterior parietal cortices) using resting-state fMRI data.]

Comment: I think the authors have given a satisfactory answer.

- Page 14, line 14, sample size calculation. As previously mentioned, this section is somewhat sketchy. My main question is which endpoint that is chosen for the calculation of the sample size. In my opinion, this study in reality comprise two sub-studies involving two separate samples, with different main outcomes (alcohol and nicotine use), which would probably end up in different sample size requirements. The reader lack information regarding which outcome is expected to have an effect size of 0.2. It seems highly unlikable to me, that 24 participants per group would be enough to detect difference between groups (this opinion based on the fact that previous studies on cognitive training have been negative on drinking outcomes). But this is of course speculations from my side.
Thank you for this follow-up question. We revised this section. Please see our answer to your question #3 for more details.

Comment: I think the authors have revised the manuscript satisfactorily.

- Page 15 – hypotheses part: The hypothesis again differ from outcomes describes in the "Endpoint section" as well as from the Gov registry page. For example, psychosocial functioning is mentioned here, but not previously. My suggestion is that these sections are harmonized to make it easier for the reader to follow what has been primary and secondary endpoints/hypotheses
Thank you for this suggestion. We revised the corresponding paragraphs accordingly.

Comment: I think the authors have revised the manuscript satisfactorily.

- Lastly, related to previous points, it seems odd to have so many primary endpoints, and only one or two secondary. Usually, one single outcome is chosen as primary (based on theoretical considerations described in the Introduction), and the rest is secondary. This in order to avoid the risk of primary endpoints being chosen retrospectively, based on outcome of the study.
Thank you for this remark. As we defined and published (clinicaltrials.gov) the endpoints beforehand at, we would suggest keeping it that way. Please see our revised manuscript for a clearer description of endpoints since we agree that our description lacked consistency.

Comment: I think the authors have revised the manuscript satisfactorily.