Dear Dr. Zahedi,

We would like to thank you and the reviewers, Katie Hobbs and Mariela Mihaylova, for reviewing our registered report and offering such detailed feedback. We appreciate the critical assessment of our paper and the chance to submit an improved manuscript. We address the concerns raised point by point.

<div align="center">

Response to reviewers
</div>

General notes. All relevant changes in the text are highlighted in yellow. Text is not highlighted where only the structure has been changed or minor grammatical/readability changes were made. The order of the hypotheses and research questions have been altered to better reflect our approach and expected effects. Namely, RQ5 is now RQ1 (effects on similarity ratings). We made these changes for readability and because the manipulation checks were separated from the hypotheses.

Response to editor comments

Comment E1. First, the manipulation checks in the proposed study are not clearly distinguished from the experimental manipulations. This needs to be either clarified better or to be developed better.

Response. We have changed the previous H1 and H2 to clear manipulation checks (MC1 and MC2) and describe these in the 'Manipulation Check' section (Pg. 11, line 17): "We expect there to be two observable group effects as a result of the manipulation. First, the ratings of temporal distance on a manipulated timeline (Figure 1) are expected to be lower in the differences condition than in the similarities condition (MC1), which will be tested using an independent groups t-test. This would indicate that the manipulation successfully influences participants to rate the recalled event as closer or more distant from 'today' in the respective similarities and differences conditions. Second, the manipulation should yield significantly higher scores on psychological distance in the differences relative to the similarities group (MC2), which will be tested using an independent groups t-test."

Comment E2. Second, the authors need also to discuss how and when they will interpret their negative results. This is currently missing.

Response. We have introduced the interpretation of negative results into the sixth column of Table 1, 'Interpretation given different outcomes' (pg. 13). For example, for H1 and H2, "Evidence for the null hypothesis would be regarded as moderate with $3 < BF_{01} \leq 10$ for the interaction effects."

Comment E3. Third, the power analysis is somewhat limited and has not been based on an exact effect size; further, it needs to consider the uncertainty of the effect sizes. I suggest the authors include a power analysis that does so, using a more advanced analysis, or at least justify their hypothesized effect sizes better.

Response. The medium effect sizes were based on the interaction effect found in the Hanko et al. (2009) paper. As per the recommendations of the reviewers, we have simulated data based on medium effect sizes and conducted a power analysis using the superpower package in R. Under the section Sampling Plan, we now describe our selection of effect sizes

and the corresponding power analysis (pg. 24, line 5): "Except for the medium interaction effect size reported by Hanko et al. (2009; Study 1) for an interaction involving a similarity focus manipulation ($\eta2p=.07$), there are no similar studies directly manipulating comparison focus in context of aversive memories and/or that closely match our experimental conditions. Moreover, the above-mentioned effect size cannot directly be translated to the present study due to different outcome variables and the availability of baseline assessments for most analyses, which may enhance precision. For our 2 (comparison focus: similarities, differences) × 3 (time: baseline, post-manipulation, post-speech) experimental design, we simulated ANOVAs with a hypothesized experimental group effect size of approximately d = 0.50 from baseline to post-manipulation, and a slightly reduced group effect size of approximately d = 0.45 from baseline to post-speech, whereby repeated measures were correlated with r = .50. The sample size was set at n = 80 per group, resulting in a total sample of N = 160. Using the 'superpower' package in R, we conducted simulations with 1000 iterations to evaluate the statistical power with alpha set at 0.05. For the comparison focus × time interaction in the overall ANOVA, we determined a power of 86.4% (95% CI = [84.1, 88.4%]). The power to detect significant group effects in the contrast from baseline to post-manipulation was estimated at 89.2% (95% CI = [87.3, 91.1%]), and for the contrast from baseline to post-speech it was 80.3% (95% CI = [77.8, 82.5%]).

To assess statistical power of the MANOVA, we conducted a Monte Carlo simulation-based power analysis with two dependent variables: self-rated and experimenter-rated speech performance. We assumed a moderate correlation (r = .30) between these dependent variables. We then evaluated two scenarios. First, we expected relatively uniform moderate effect sizes (d =0.50 and d=0.45). With 80 participants per group, power was estimated at 92.9% (95% CI = [91.2%, 94.5%]). Second, we evaluated a scenario where one effect is medium-sized (d = 0.50) and the other small (d = 0.20). Here power was estimated at 82.2% (95% CI = [79.9%, 84.4%]). The R code for our power calculations can be found on the project's Open Science Framework folder [https://osf.io/39uwb]."

Comment E4. Finally, the authors also need to consider the alternative interpretation of their results as in the current format; they only focus on one possibility and exclude others. I would suggest being more open to alternative interpretations of the results here.

Response. We added alternative interpretations of the results in the final column of Table 1, 'Theory that could be shown wrong by the outcomes' (pg. 13). For example, for RQ1, "Previous research has shown that perceived temporal distance with past selves corresponds to perceived similarity, i.e., closeness-similar, distance-different. Opposite effects would indicate that the selective accessibility model does not account for cognitions in socially anxious individuals and would require further investigation."

Response to Reviewer 1 (Katie Hobbs)

Comment R1.1. Thank you for inviting me to review this interesting stage 1 report. McCarthy et al plan to investigate whether manipulating the temporal distance from a negative social memory by priming participants to focus on differences or similarities produces change in

social anxiety, affect, and performance on a speech task. I have evaluated the paper in line with required criteria as outlined below and have included some additional thoughts at the end.

McCarthy et al plan to investigate whether priming participants to focus on differences or similarities alters the perceived temporal distance of a negative social memory in a sample of individuals with subclinical levels of social anxiety (RQ1). They then plan to investigate whether differences in perceived temporal distance produces differences in perceptions of social anxiety in the past self (RQ2), positive/negative affect when recalling the memory (RQ3), performance on a speech task (RQ4) and perceived similarity with the past self (RQ5). I believe that these research questions are scientifically justifiable and fall within established ethical norms.

Response. Thank you for taking the time to provide such detailed feedback, it is very much appreciated.

Comment R1.2. The hypotheses as stated in Table 1 are coherent and credible. However, it is unclear how McCarthy et al would interpret contradictory results for hypotheses centered around the same research question. For example, if McCarthy et al find evidence of a group difference for H1 but not for H2 would they conclude that the experimental manipulation was successful?

Response. As per our response to the editor above (E1), we have separated the manipulation checks from the hypotheses. Under the 'Manipulation Checks' section (pg. 11, line 25), we have included the following explanation regarding interpretation: "In the case that MC1 is not confirmed but MC2 is, we would assume that the spatiotemporal timeline is not representative of elapsed time. Conversely, if MC1 is confirmed but MC2 is not, we would assume that priming comparison focus has not had the desired effect on perceived psychological distance from the past self. Alternatively, it could imply that effects of the temporal comparison on self-appraisals are stable and not affected by assimilation or contrast effects of comparison focus. If any of the following hypotheses are supported but MC2 is not confirmed, this would suggest that comparison focus effects cannot be attributed to perceived temporal distance. Therefore, we will investigate group differences on outcomes even when MC1 and/or MC2 are not confirmed."

Comment R1.3. I also think that McCarthy et al could be more specific with the interpretation of the hypotheses in table 1. For example, for RQ2 McCarthy et al write "Significant group effects would show that a comparison focus on differences or similarities can influence current and past self-evaluations of social anxiety symptoms.". I think it would be helpful to consider the direction of effects here (particularly in an opposing direction to that hypothesised) and relate directly back to the outcome measure of the hypothesis. It may also be helpful for McCarthy et al to provide figures for their hypothesised results which can later be compared against the actual findings.

Response. We have specified interpretations of the effects of the hypotheses to the relevant outcome measures in the sixth column of Table 1 'Interpretation given outcomes' (pg. 13). For example, for RQ1, "H1a will be confirmed by a group × time interaction effect from baseline to post-manipulation. H1b will be confirmed by the interaction effect of group on

changes in similarity ratings with the past self. This would indicate that the experimental manipulation induced the expected contrast or assimilation effects on perceived similarity." We would avoid including further figures in the paper and hope that the updates to the structure negates this.

Comment R1.4. When stating the hypotheses, this is very clearly done within Table 1 but is less clear in the introductory text. I think it would be helpful to be much more explicit as to the expected group differences within this section rather than just stating a group difference will occur.

Response. We have now explicitly stated the directional hypotheses in the introduction under 'Hypotheses'. E.g. (pg. 12, line 11), "As per RQ1, we expect ratings of perceived similarity with the past self to decrease from baseline to post-manipulation in the differences group and to increase in the similarities group, reflecting a time × condition interaction (H1a). Importantly, we expect higher similarity ratings at post-speech for the similarities group than the differences group (H1b)."

Comment R1.5. In Table 1, H4 "The difference in social anxiety ratings between past and current selves will be higher in the differences group than the similarities group". Could the authors be more explicit about what the difference means here? I understand H3 with the current levels of social anxiety but what does a difference score tell us beyond this?

Response. H4 (now H2b) has now been made more explicit in Table 1 (pg. 13): "H2b: The differences group will exhibit a larger difference between current and past social anxiety ratings than the similarities group, which will exhibit little to no difference between the two ratings."

Comment R1.6. I believe that the study procedure is feasible, and McCarthy have demonstrated this in their preliminary study. It does seem like a complicated and long procedure for participants though, so I wonder about potential fatigue effects. Is it possible for McCarthy et al to measure/mitigate this?

Response. In the preliminary study, the procedure including informed consent took approximately 30 minutes. We do not have any cognitively demanding tasks and no questionnaire requires more than 5 minutes to answer. We therefore do not foresee any potential fatigue effects and none were reported by participants/experimenters in the preliminary study.

Comment R1.7. Regarding the sample size calculation, a fairly rudimentary calculation has been done using a medium effect size and conventional levels of power/alpha to determine a sample size of 128 participants. Firstly, I think it would be helpful to use a more specific effect size rather than a generic medium effect (is there a specific analysis that is relevant in Hanko et al?). Secondly, I am concerned that this sample size is too small given (a) the inconclusive results of the preliminary study, and (b) the large number of hypotheses being tested. I would be more confident in the proposed sample size if McCarthy et al could visualise their expected results as I have suggested in the previous section and simulate data based on these parameters to test the required sample size. I also think it might be helpful

Response. As per our comment to the editor (E3), we based the interaction effect size on data from our source paper Hanko et al. (2009). Additionally, there are no similar published studies directly manipulating comparison focus in context of aversive memories and/or that closely match our experimental conditions. We have simulated data based on the medium and low effects and run power analyses for the repeated measures ANOVAs and a MANOVA, described in the Sampling plan section, pg. 24, line 5, as well as Table 1. We have also reduced the number of hypotheses from 10 to 7, given that 2 of the hypotheses are now labelled as manipulation checks and we only consider changes in negative affect, not positive affect (see Table 1, pg. 13).

Regarding the data quality, as we are using a laboratory-based sample, this does not typically require assumptions of poor quality. Moreover, the sample size estimation refers to completers and all dropouts will be replaced, thus attrition is addressed.

Response. Given that we use the temporal distance measure (the MEQ psychological distancing subscale) for the manipulation check, it would not be viable to include this in further analyses. Also, the only difference between the experimental conditions (differences vs similarities groups) in our design relates to the manipulation of temporal distance from a negative social memory. Consequently, we have reason to expect that any significant changes between the groups in the outcome measures are related to our manipulation.

Response. Thank you for the comment.

Response. We have addressed the manipulation checks in previous comments. We do not expect floor or ceiling effects of our primary and secondary outcome variables as per the preliminary study (except for speech length, see comment R1.14 below), but will assess normal distribution, skewness and kurtosis as described under the Analysis plan section, (pg. 27, line 1).

Comment R1.11. The preliminary study is clearly distinguished from the work that is yet to be done. I find the figure outlining the study procedures to be very helpful.

Response. Thank you for the acknowledgement. Considering later comments from Reviewer 2, we have moved this to the appendices to improve readability.

Comment R1.12. The report is comprehensive; however I think it would be benefit from the language being more accessible. McCarthy et al have used a lot of jargon, particularly in the abstract, which makes the manuscript quite dense and difficult to read if you are not an expert within this field.

Response. We have attempted to explain concepts in line with the literature, given we have borrowed from different theories. We have rewritten several sections of text to improve understanding and use fewer technical terms where possible. For readability we have shortened sentences in the introduction, e.g. (pg. 4 line 12), "Most humans ~~tend to~~ focus on personal improvement~~, with personal change perceived as more~~ and perceive their personal development over time as positive, ~~suggesting that a~~ indicating a self-enhancement motive ~~often drives~~ in temporal comparison." We have also made the abstract more accessible, for example: "Temporal comparisons with past selves have been found to influence current self-appraisals of attributes, including well-being. The comparison process involves using a past self as a standard, while the current self serves as the target. ~~In this study we utilize temporal comparison and instruct participants reporting elevated symptoms of social anxiety to focus on differences or similarities with a past aversive memory to assess changes in affective, cognitive and behavioural outcomes.~~ Previous evidence has shown that focusing on differences (similarities) when comparing with a past extroverted self will lead to lower (higher) current ratings of extraversion, indicating contrast (assimilation) effects."

Comment R1.13. Temporal discrimination task instructions - I'm not sure the temporal manipulation is 'pure' here as it's mixed up with positivity/negativity. Participants primed to focus on similarities are asked about things that have improved, whereas those focused on differences are asked about things that have not improved. I would suggest there's a valanced connotation using the word 'improved', perhaps 'changed' may be a more neutral alternative.

Response. We agree and have amended the instructions from improved/not improved to changed/not changed for the differences and similarities conditions respectively (pg. 33, line 2).

Comment R1.14. In the speech task, participants are allowed to ask to stop. How will the authors handle data in this situation?

Response. Speech length is recorded as a variable. We have added the following explanation on (pg. 22, line 12): "As speech length will have ceiling effects due to the five-minute time limit, we will dichotomize the data into participants who finished prior to five minutes and those who finished early. We use this data only for descriptive purposes of group differences."

Response to Reviewer 2 (Mariela Mihaylova)

Comment R2.1. This is an interesting paper on how we perceive our past and current selves. It investigates how focusing on either similarities or differences with a past self affects psychological distance in individuals with social anxiety. The paper was strongly grounded in theory and well-researched. The hypotheses and methods are logical and strongly linked to the literature. I also appreciated the clear detailing of each hypothesis and analysis in Table 1 and strong research question-to-analysis connection.

Although the topic of the current paper is outside my area of expertise, I will do my best to give some suggestions for improvement to the authors. In the below review, I break down my comments by each section of the paper.

Response. Thank you for your time and efforts in helping us to improve this paper.

Comment R2.2. The introduction is generally logical and well-researched, however I found it very difficult to read overall and saw issues with the flow. All the sentences were very complex, making the paper difficult to understand. It was also difficult to keep all the terminology straight. The whole introduction would benefit from using simpler, shorter sentences, breaking up the walls of text into paragraphs to increase readability and clarity, and forming better connections between all the paragraphs.

One way to make sentences more clear would be for example:

The sentence: "This suggests that dysfunctional comparison habits may contribute to pathological behaviors and cognitions" could be changed to → "Dysfunctional comparison habits may contribute to pathological behaviors and cognitions."

Response. We have shortened sentences in the introduction and reduced similar references, such as using self-appraisal or self-perception exclusively instead of self-evaluation and self-concept. For example (pg. 5, line 2), "Memories of aversive past events shape current self-perceptions and ~~may thereby~~ can provoke current worry or fear of negative evaluations ~~by others in the here and now~~ (Matos et al., 2013)."

Comment R2.3. Importantly, the research gap and objectives are missing from the introductory paragraph. The research background and context is provided, but does not clearly articulate what is missing from the literature that this study will address, nor what the objectives of the study are. This makes the reader not understand why this study matters.

Response. We have rewritten much of the introductory paragraph to better reflect the approach of the study and the research gaps and how this will be addressed (pg. 3, line 12): "The Selective Accessibility Model (SAM; Mussweiler, 2003) suggests that comparison focus on similarities or differences will lead to respective assimilation or contrast effects. For example, perceiving one's current mood as consistent with a negative past memory will trigger a search for evidence (similarities) to support this. Subsequently, manifested assimilation of a negative past self into the current self can lead to reduced self-esteem. Temporal comparison processes therefore influence mental health, and evidence indeed shows that temporal self-appraisals are negatively affected by anxiety and depression (Sokol et al., 2022; Sokol & Serper, 2017). Manipulating comparison focus in temporal comparisons has shown assimilation and contrast effects on subsequent self-appraisals (Hanko et al., 2009), while perceived temporal distance has also been found to produce assimilation and contrast effects with recalled past selves (Broemer et al., 2008). However, to our knowledge, no research has investigated these effects in (sub-)clinical samples, and it remains unclear whether contrast and assimilation effects influence symptoms of clinical social anxiety. Negative perceptions of past selves are indeed relevant to social anxiety, with shame memories often acting as reference points. Memories of aversive past events shape current self-perceptions and can provoke current worry or fear of negative evaluations (Matos et al., 2013). It is therefore important to investigate how manipulating recall of past negative selves may produce assimilation and contrast effects on self-appraisals and subsequent indicators of mental health in a clinically relevant sample. In this paper we present an experimental paradigm to assess whether manipulating comparison focus influences the impact of temporal perceptions of an aversive memory on social anxiety, affect, and behaviour."

Response. As per our response above (R2.3), we have rewritten the first paragraph to introduce the key concepts relevant to the study goals.

Response. We have defined perceived temporal distance in the corresponding section (pg. 6, line 19): "Perceived temporal distance refers to the psychological proximity to a past memory and has a significant impact on the way past selves are perceived and appraised."

Comment R2.6. The methods section also talks about "augmented assimilation-contrast paradigm" but this never defined.

Response. We have removed this reference and instead describe the initial goal of the study (pg. 10, line 2): "In the present study, we investigate whether contrast or assimilation effects with a negative past self can be induced in a student sample with elevated, sub-clinical levels of social anxiety using an experimental manipulation of comparison focus (Hanko et al., 2009; Mussweiler, 2001)."

Comment R2.7. The paper also jumps right into theories and concepts without really defining them. For example, the SAM model is never defined, making it confusing and unclear. Same with the temporal self-appraisal theory. We need clear definitions.

Response. We have included definitions of the following models.

Temporal self-appraisal theory (pg. 4, line 14): "Temporal Self-Appraisal Theory posits that individuals evaluate their past selves in ways that enhance their current self-perception, often perceiving former selves as inferior and temporally distant compared to their present self (Ross & Wilson, 2003). In contrast, more recent past selves will be more favourably appraised and past successes as more recent."

SAM model (pg. 3, line 12): "The Selective Accessibility Model (SAM; Mussweiler, 2003) suggests that comparison focus on similarities or differences will lead to respective assimilation or contrast effects. For example, believing that one's current mood is congruent with a negative past memory will trigger a search for evidence (similarities) to support this."

And later (pg. 5, line 9): "Temporal comparison can be understood as a process of (dis)similarity testing between the current self (i.e., the target) and a past self (i.e., the standard). This process is primed by the selective accessibility of hypothesis-(in)congruent information (Mussweiler, 2003). Specifically, when individuals are primed to focus on similarities versus differences with a comparison standard, assimilation versus contrast effects are likely to occur, respectively (Mussweiler, 2020). Assimilation implies that target evaluations shift toward the standard, whereas contrast means that evaluations shift away from the standard."

Comment R2.8. I found the transitions between each paragraph in the introduction not very well connected. Right now, the introduction section is like mini-essays on each construct but how they are all connected is unclear.

For example, the paper starts talking about comparison, then jumps to temporal distance, then back to the comparison process. The links between them are not straightforward and going back and forth between concepts is confusing.

To make the transitions smoother, you could add transition sentences to paragraphs and connecting sentences such as "building on this concept…" or "expanding these findings to…"

Making the connection with social anxiety clearer for each construct would also help improve this section

Response. We have restructured the introduction for a more intuitive read: Temporal comparison theory → Assimilation contrast effects → Perceived Temporal Distance → Temporal Comparisons Relevant to Social Anxiety and Affective Disorders. This has included improving the coherence between sections and making the connection to social anxiety clearer, e.g. prior to Perceived Temporal Distance (pg. 6, line 12): "These effects on a stable concept such as extraversion have implications for further use of managing perceptions in other psychological concepts, such as anxiety. Expanding these findings in clinical samples could provide a novel mechanism of change for therapeutic interventions. It is therefore important to consider further useful mechanisms in the temporal comparison process, particularly perceptions of temporal distance to temporal selves."

Comment R2.9. Hypotheses and manipulation checks. Did I miss it or is this section only discussing the hypotheses? I did not see any manipulation checks here. It might be helpful to state how RQ1 and 2 serve as manipulation checks in this section.

Response. As per our comments above to the Editor (E1) and Reviewer 1, we have separated the manipulations checks from the hypotheses.

Comment R2.10. Another thing to consider regarding manipulation checks would be to include a specific manipulation check items immediately after the priming tasks to assess whether participants perceive the intended temporal closeness or distance.

Response. The MEQ measure for MC2 will be administered directly after the manipulation.

Comment R2.11. In Table 1, the authors mention what previous research hasn't addressed. Why is this only being stated here but not in the intro where it belongs? Bringing this up sooner would paint a better picture for the reader about what is currently missing in the literature (the gap) and how your study will help address it.

Response. We have addressed this in response to the comment R2.3.

Comment R2.12. Methods (the proposed study). After the intro, the paper jumps into a preliminary study, for which the hypotheses just outlined don't correspond to. This breaks the flow of the paper because everything the reader just learned the current study would be about in the intro, is not what follows in the methods.

I would advise authors to go straight into the current (proposed) study after the methods. This would make the whole paper more coherent. Maybe mention that this was developed following the preliminary study and then remove the preliminary study from the main paper. You can still put the preliminary study in a supplementary section for reference, but in the current paper the preliminary study really doesn't bring much. It is also missing key details necessary for a methods section like the experimental design.

Response. As the preliminary study was key to the development of the current proposed study, we have moved it to the Appendices. This will improve coherence and still keeps the information easily accessible to the reader.

Comment R2.13. Would be helpful to have a table or section in the Methods that describes how each of the constructs from the introduction will be operationalized and measured to remind the reader.

Response. We have added Table 2 on page 19 to describe briefly the operationalization and measurement of relevant constructs related to outcomes.

Comment R2.14. A 3x2 design is mentioned but the factors and levels are not stated. Can you clarify?

Response. We have clarified this further (pg. 15, line 24): "This research protocol is a 2x3 mixed design (Comparison focus x Time), where participant allocation to the comparison focus conditions (differences or similarities group) is randomized and double-blinded, while time refers to measures at baseline (T1), post-manipulation (T2) and post-speech task (T3)."

Comment R2.15. It would be helpful to assess power based on the type of design you have. This will help give you more reliable findings. You can use the Superpower package in R, for example: https://cran.r-project.org/web/packages/Superpower/vignettes/intro_to_superpower.html#specifying-the-design-using-design I would recommend this over using G*Power, which has been shown to not be very robust in all cases. For more info, see Brysbaert, 2019.

Response. Thank you for the advice regarding the R package superpower. We have simulated data and used the Superpower package to run advanced power analyses and have addressed this in previous comments. See E3 and R1.7.

Comment R2.16. The impromptu speech tasks are known to elicit stress in participants (i.e., Shields et al., 2017). Did I miss it but why are you doing this in addition to having an already an anxious sample? What does having participants being "double stressed" bring? The rationale for this needs to be further justified.

Response. The speech task is a common task to elicit stress to assess effects on social performance of manipulation or treatments in clinical or subclinical anxious samples (Edwards et al., 2003; Kampmann et al., 2016), which we describe in our Empirical Approach (pg. 10, line 24). We use an established protocol and do not expect any adverse reactions. Nonetheless, participants will be offered the chance to contact a psychotherapist in the work unit at the end of the study, should they experience significant distress. This is already included in our ethics approval.

Comment R2.17. The participants section needs more clarification. Like which psychiatric disorders were excluded and how were they assessed? What were the cut-off values for exclusion?

Response. Inclusion and exclusion criteria were self-reported via online screenings. To clarify, we have added (pg. 16, line 18): "The inclusion and exclusion criteria questions are self-report and participants will be informed of their eligibility after completing the online screening."

Comment R2.18. Also, for the social screening and depression section, what is the cut-off value for depression above which you will exclude participants? I don't believe this is mentioned.

Response. We do not screen for depression; hence we do not use cut-offs. We describe the assessment of depression in the Empirical Approach and the Protocol (e.g. pg. 11, line 5), "Depression will also be measured at baseline to control for potential baseline group differences when assessing affect". As we assess negative affect at three timepoints of the experiment, we want to control for potential effects of group differences in depression at baseline, if necessary.

Comment R2.19. Page 16 - the authors state the FNE instructions are not validated. How are you planning to validate them to ensure reliability?

Response. We do not use the FNE for diagnostic purposes (i.e., identifying individuals with clinical social anxiety), but to assess the assimilation or contrast effects of the manipulation. We will report internal consistency and do not intend to validate the adjusted FNE items.

Comment R2.20. In the missing data section, it would be helpful to stipulate how other types of missing data will be handled (i.e., if participants write nonsense or illegible responses in the forced choice questionnaires) - will they be excluded or will you include them based on various cut-off values?

Response. We now address this under Aversive Memory Characteristics in the Assessment of Variables section (pg. 21, line 11): "All ratings that are below two standard deviations from the sample mean will be further reviewed. An experimenter will review the memory recall descriptions. This will enable us to screen for events that do not adequately reflect a negative social experience, i.e., do not include feelings of embarrassment, humiliation or shame. Participants reporting lack of such feelings will be excluded from the analysis."

Comment R2.21. Similarly, how will you handle outliers in the dataset?

Response. We have added the following explanation under Missing and Outlier data section (pg. 26, line 12): "Outliers will be identified as values of +/-2.5SD from the mean for all outcome variables. We will employ Winsorizing and replace data points with the sample mean +/-2.5SD prior to the analyses (Rivest, 1994)."

Comment R2.22. Lastly, it doesn't look like this paper corresponds to APA guidelines. Can the authors use APA guidelines please? This is the standard for most journals and papers.

Response. The style has been updated to meet the APA guidelines.