

Reviewer #1

We thank the reviewer for their encouraging comments and insightful suggestions for improvement. The suggestions for additional mental health measures were particularly helpful for evaluating potential confounds.

1. I was wondering whether the authors do expect individual differences in the extent of lonely individuals displaying hypersensitivity and/or hyperalertness? Maybe the authors could clarify this issue.

Response: We thank the reviewer for this insightful comment regarding the potential individual differences in hypersensitivity and hyperalertness among lonely individuals. We agree that this is an important aspect to consider.

In our current study, we primarily focus on establishing whether a general pattern of hypersensitivity and hyperalertness to social stimuli can be identified in lonely individuals compared to non-lonely controls. As such, our analyses are designed to detect overall trends across groups rather than within-group variations.

However, we acknowledge the importance of individual differences within the lonely population. These differences could be influenced by several factors, such as the duration and perceived severity of loneliness, co-occurring mental health concerns, and socio-demographic factors. While our study is not specifically powered to explore these individual variations in detail, we intend to conduct exploratory analyses to investigate any significant variances within the lonely group. This will involve examining potential moderators and mediators that might influence the extent of hypersensitivity and hyperalertness.

Moreover, should we find significant patterns in these exploratory analyses, we propose to mention these findings as preliminary and suggest them as a basis for future studies specifically designed to address individual differences in loneliness and associated cognitive processes.

Changes to the manuscript:

- Page 16, Line 518: "Interpretation of control analyses considering individual differences in hypervigilance and hypersensitivity."

2. Although the authors explain why they plan to include individuals with mood disorders, I remain skeptical as to whether this approach could confound potential findings.

Response: To clarify, we do not include people who are undergoing current treatment for mood disorders. However, we do include people with a lifetime history of mood disorders or elevated scores on questionnaires tapping related constructs. The lifetime prevalence of mood disorders is extremely high, with some estimates indicating a third of the population. Moreover, mood disorders are closely related to loneliness. Arguably, selecting a "pure" group of lonely people without any history or sign of mood disorder would not represent the

typical presentation of loneliness. However, we exclude people who are undergoing treatment for mental health disorders or received treatment recently, because the treatment may affect participants' responses and the study is not intended to investigate treatment effects.

We will highlight the potentially confounding effects of mood disorders in our analysis and discussion. For instance, we will compare the strength of the association between our ERP markers and scores on mood disorder and loneliness scales.

3. The authors may also consider to include the Brief Symptom Inventory to check the effect of comorbid mental health issues, as well as questionnaires on current stress levels (e.g. Perceived Stress Scale; Cohen et al., 1983).

Response: We welcome the reviewer's helpful suggestion regarding the questionnaires and incorporate them in our protocol.

Changes to the manuscript:

- Page 6, Line 235-239: "**Perceived Stress:** Perceived stress can influence people's emotional response and reactivity. To distinguish the effect of loneliness from perceived stress, we will administer the Perceived Stress Scale (PSS). The PSS is a widely used self-report scale designed to measure the degree to which situations in one's life are appraised as stressful. The scale has strong psychometric properties, with a Cronbach's alpha coefficient ranging from 0.74 to 0.86 (Cohen, Kamarck, and Mermelstein 1983)."
- Page 7, Line 262-266: "In addition, we will administer the Brief Symptom Inventory (BSI) as an indicator of general psychological distress (Derogatis and Melisaratos 1983). The Brief Symptom Inventory (BSI) is a comprehensive self-report questionnaire developed to evaluate a broad range of psychological symptoms. The scale shows strong psychometric properties with a high Cronbach's alpha coefficient, ranging from 0.71 to 0.85 (Derogatis & Melisaratos 1983)."
- Page 15, Line 502: Exploratory analyses controlling for social isolation, **mental health symptoms, and perceived stress**
- Page 16, Line 514ff: Interpretation of control analyses to establish the specificity of the association between loneliness and social processing, highlighting potential unique and shared mechanisms in loneliness with reference to the relevant literature in social anxiety, depression, **and perceived stress.**

4. Why were the ERPs averaged over the 5th repetition (and not the 6th)?

The mean amplitude of the ERP within the channel region and time window identified through spatiotemporal clustering decreases with the number of repetitions (see figures below). This decrease is not linear. We considered modelling the decrease and comparing the slope of that function between the lonely and non-lonely group. However, we found that this approach was quite sensitive to noise. Therefore, we decided to pursue a more robust analytic strategy that would still allow us to test our hypotheses. The 6th repetition occurred somewhat less frequently, because of the algorithm that randomly determined the number of repetitions. We thus selected the 5th repetition for analysis, which optimized both number of repetitions and trials in our paradigm.

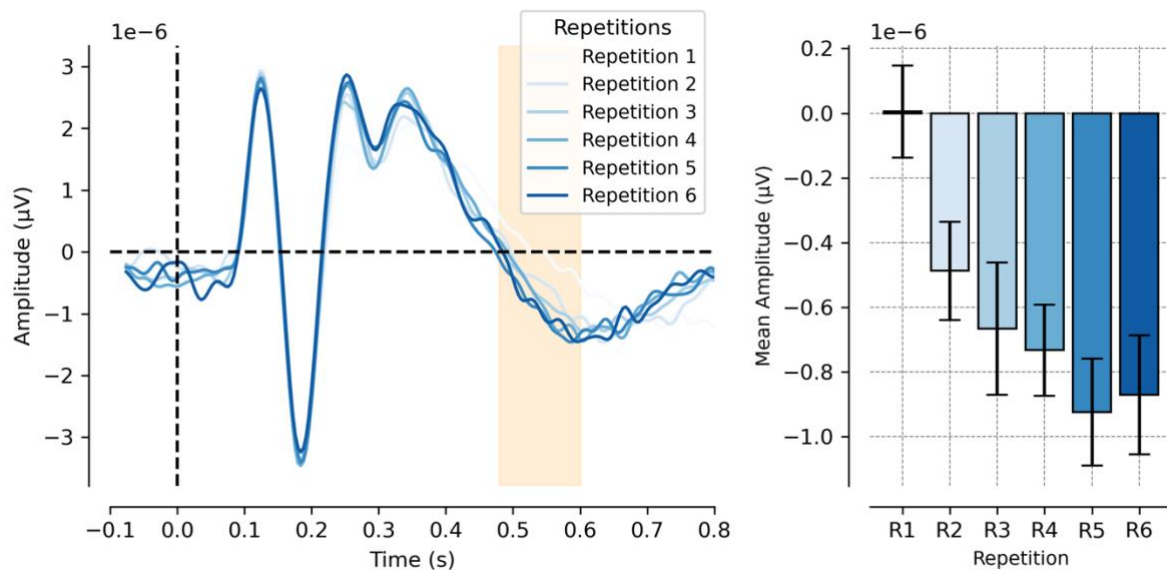


Figure 1 The left panel shows the grand-average event-related potential response to angry facial expressions for each repetition in the sequence. Channels identified through spatiotemporal clustering were averaged together. The shaded area shows the 480-600ms time window also identified through spatiotemporal clustering. The right panel shows the mean amplitude within the channel region and time window of interest for each repetition. The error bars show one standard error.

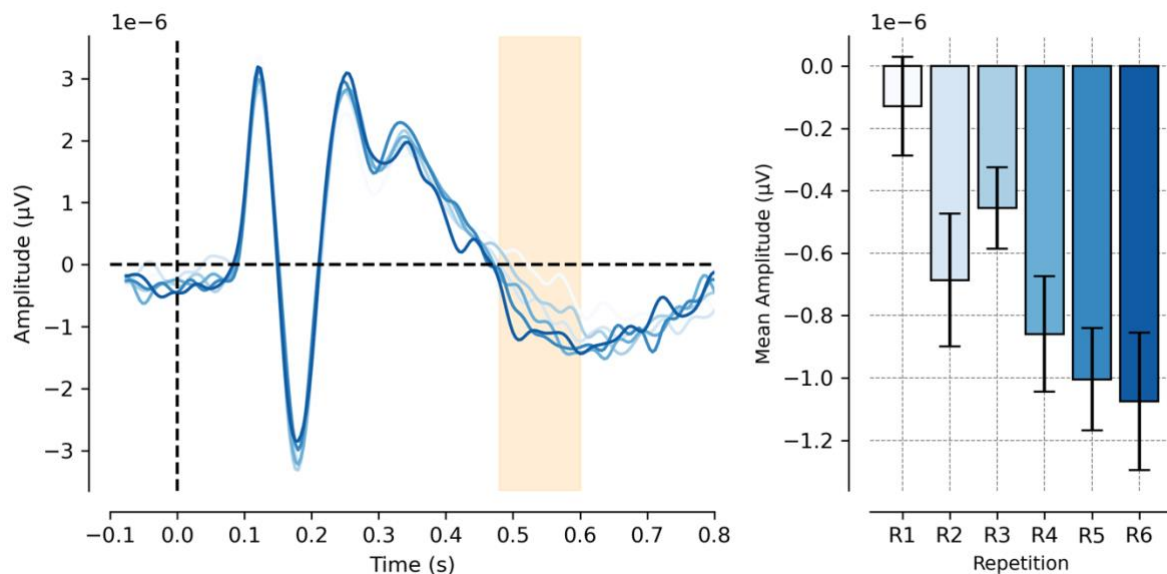


Figure 2 Grand-average ERP and mean amplitude as above for the happy face condition.

Changes to the manuscript:

- Page 13, Line 463-467: “The mean amplitude in response to happy and angry faces decreased with the number of repetitions. To characterise the effect of repetition, we focused on the contrast between the first repetition and the 5th repetition. We chose the 5th over the 6th repetition, since the 6th repetition only occurred in a relatively small subset of trials, due to random allocation of number of repetitions (varying between 6 and 10 repetitions) for each trial.

5. How many trials were available for averaging?

Response: We calculated the number of available trials for each condition and included the results in the revised manuscript.

Changes to the manuscript:

- Page 9, Line 372-378: “We evaluated the number of trials available for analysis after EEG processing. For the angry emotion category, the mean number of trials for the first repetition was 67.08 (SD = 7.03), with a minimum of 34 and a maximum of 73 trials. In the fifth repetition for the same emotion category, the mean number of trials was 68.77 (SD = 7.17), with a range from 37 to 75 trials. Regarding the 'happy' emotion, the first repetition showed a mean of 53.03 trials (SD = 5.30), with the number of trials ranging from 35 to 58. The fifth repetition for 'happy' had a slightly lower mean of 52.90 trials (SD = 6.31), and trials varied between 28 and 58.”
- Page 10, Line 382-385: “The number of trials were equated during the averaging stage. The condition with the least responses determined the number of trials for all conditions, and the other condition(s) were subsampled through random selection of trials.”

6. Finally, I wonder whether the authors may consider to conduct ERP analyses which analyze all available temporal and spatial information (conducting TANOVAs or analyzing GFP; for details, see, e.g., Murray et al., 2008, Brain Topography; Cacioppo et al., 2015, Journal of Neuroscience Methods; Schiller et al., 2023, Brain Topography) rather than a priori disregarding specific spatial and temporal information. Having that said, the approach proposed by the authors seems valid, ...

Response: We appreciate the reviewer's suggestions. Indeed, the proposed analyses could enrich our understanding of the electrophysiological data and offer deeper insights into the underlying neural mechanisms. However, for this registered report, we have chosen to employ the simplest and most well-validated method to test our hypotheses. Consequently, we adopted a two-step approach: initially applying data-driven methods in our first study to establish robust predictions and then testing the group difference based on these predictions. We believe the first study may not be adequately powered for more extensive analyses, and it remains uncertain how well some of the more advanced analyses will replicate in a new sample.

While we intend to conduct more detailed exploratory analyses of the EEG data, including time-frequency and source analyses, our primary goal in this initial analysis is to robustly test our hypotheses.

7. ... controlling for multiple testing (but why was alpha set to 0.02?).

This alpha level is recommended for registered reports and is a requirement for some journals, including *Cortex* (Reference: [Guidelines for Authors](#))

Reviewer #2

We thank the reviewer for their time and consideration. The reviewers' suggestions helped us to clarify important methodological points.

1. In the abstract as well as in the hypotheses, the authors expect greater electro-cortical signal to angry vs. happy faces in those participants, who report high levels of loneliness. Based on the literature (and also as they correctly stated), this pattern of response is to expect also in those with low loneliness. I would therefore slightly reframe the hypotheses mentioning the group differences.

Response: We thank the reviewer for this suggestion. We incorporated the interaction in the description of the hypothesis.

Changes to the manuscript:

- Page 10, Line 413ff: "For hypothesis 1, we expect that lonely people show increased sensitivity to angry **over happy facial** expression. This is operationalised as an increased ERP mean amplitude to deviant angry faces in spatiotemporal clusters that showed significant differences between angry compared to happy faces in Study 1."
- Page 10, Line 417ff: For hypothesis 2, we expect that lonely people show reduced habituation when being repeatedly exposed to angry facial expressions, **while the habituation is expected to be stronger for happy expressions**. This is operationalised as a significant expression-by-repetition interaction in spatiotemporal clusters that either show an effect of emotion or repetition in the Study 1.
- Table 1: "H1: Mean amplitude to angry faces is significantly higher at the *first* repetition and greater in lonely people compared to non-lonely people, **and this difference is greater compared to mean amplitude in the response to happy faces.**"
- Table 1: "H2: Mean amplitude to angry faces at the *fifth* repetition is significantly higher in lonely people compared to non-lonely people, **and this difference is greater compared to mean amplitude in the response to happy faces.**"

2. Were the participants for the screening recruited from the general population or were they psychology students?

Response: We will recruit participants through local advertising using leaflets and posters in public libraries and community centres. We cannot guarantee that we will be able to recruit a fully representative sample, because research volunteers tend to be more educated and affluent than the general population. Recruiting a fully representative sample requires intensive resource investment that we do not have available. However, we will document and report any sample biases. Specifically, we include measures of ethnicity and socioeconomic status. We will discuss sampling biases as a limitation if we find that our sample deviates from the distribution expected from national census surveys.

Changes to the manuscript:

- Page 4, Line 144: “We will recruit participants through local advertising, including in public libraries and community centres, using leaflets and posters.”

3. If I recall correctly, in the DSM-5 there are no mood disorders, anymore.

Response: The reviewer is correct in pointing out that mood disorders are not a distinct diagnostic category in DSM-5. We rephrased the description of the recruitment criteria.

Changes to the manuscript:

- Page 5, Line 171: “... having received or receiving treatment for disorders like anxiety or depression in the last 6 months”.
- Page 5, Line 177: “history of psychiatric disorders, except anxiety and depression.”
- Page 5, Line 187: “Loneliness is highly comorbid with anxiety and depression. Therefore, excluding participants with any history of anxiety or depression would heavily bias the sample.”

4. I have appreciated the clear list of inclusion and exclusion criteria. However, these are partially repeated in the description of the participants. I would avoid repetitions.

Response: We thank the reviewer for noticing this and have removed the duplicated information from the description of Study 2.

5. The criteria of the two researchers for the good data quality are not mentioned and I was wondering what exactly these researchers consider as good or bad data quality.

Response: Thank for pointing this out. We added a description of the quality assessment criteria to the description. We would also like to assure you that all EEG data and their quality rating will be made available to guarantee transparency.

Changes to the manuscript:

- Page 6, Line 219-221: “The inspection assessed if there was high-frequency noise or no signal in more than 10 channels and if there was significant movement- or muscle-related artefact in more than half of the recording.”

6. I could not read any specific description for the blocks in the report and this should be precisely indicated.

Response: In line with the reviewer’s suggestion, we added this to the description of the experiment.

Changes to the manuscript:

- Page 8, Line 320: “The trial sequence was split into 4 blocks of 375 trials to allow participants to rest.”

7. Participants were or will be asked to press the space bar, when the red fixation cross appears over- imposed to the face. Do they have any time restriction for performing this response? If yes, how long did they have? If not, how do the authors consider those trials with very long reaction time?

Response: The response time was not restricted, but since the trial duration was limited to 1.2-1.4s the maximum response time was 1.4s. Trials with a red fixation cross or with a button press were only added as an attention check and were therefore removed from the analysis.

As indicated in the data quality criteria, participants with accuracy scores below 90% will be excluded from further analysis and this will be recorded in the results. Since all participants in the pilot sample had nearly perfect performance (98% to 100%), we do not expect to exclude any, if all, participants based on this task. Therefore, we see the task largely as a global quality control measure. However, we will compare differences in accuracy between the lonely and non-lonely groups to investigate potential attention differences between the groups.

8. The faces are presented for 0.2 sec but the epochs are until 1 sec after stimulus onset. How are the authors sure that the effects observed after face offset are related to it?

Response: The faces in the experiment are presented briefly, for only 0.2 seconds, to minimize eye movements, which could otherwise affect the data. However, the observation period, or epoch, extends to 1 second after the stimulus begins. This longer epoch is used to capture both the immediate neural and cognitive responses to the face and the processes that continue even after the face is no longer visible. Since no other stimuli are introduced during the inter-trial intervals and participants are not engaged in any additional tasks during this time, we can be quite confident that any neural activity observed is related to the face stimuli.

Additionally, the event-related potentials (ERP) are time-locked with the onset of the face stimulus, ensuring that the recorded brain responses are directly linked to the stimulus presentation. Random cognitive processes that are not related to the participant processing the face will be averaged out. Cognitive processes that are structurally related to high-level (non-visual) processing of the faces (such as attention, categorization, emotion processing) are all of essential interest to our research question and should be picked up by the current ERP processing set up. The consistency of the responses across trials reinforces their relevance to the processing of the facial expressions. Moreover, the patterns observed in the ERP waveforms align with those documented in other studies on emotional face processing, lending further credibility to the interpretation that these effects are indeed related to the brief presentation of the faces.

9. The task entailed 1500 trials. In the report, it is read "roughly 50 trials" per condition. I found the word "roughly" somehow inappropriate for the number of trials as this should be precise. Moreover, I count 12 conditions (female vs. male faces with either happy or angry

expression for young, middle and old faces). I might count wrongly, but how exactly did the authors come to 50?

Response: We thank the reviewer for this astute observation. The exact number of trials was indeed not reported. That is because the number of times that a stimulus is presented was determined using a probabilistic rule. For each repetition train, the stimulus is presented at least 5 times. As the number of repetitions increases beyond five, the probability of continuing with the same stimulus decreases by 25% with each additional presentation. This probabilistic rule helps in varying the stimuli exposure and maintaining a degree of unpredictability in the sequence of stimuli presented during the experiment.

Therefore, the number of presentations is not the same for all stimulus repetitions. To provide a more precise description, we added the exact number of trials for the conditions that formed the basis of our analysis. We also clarified the term “condition”. We were referring to groups of stimuli that were important for our contrast, namely happy and angry faces with a set number of repetition, ignoring the sex and age dimensions.

Changes to the manuscript:

- Page 8, Line 309-313: “The exact number of repetitions and the sequence of expressions was randomly determined. For each repetition train, the stimulus is presented at least 5 times. As the number of repetitions increases beyond five, the probability of continuing with the same stimulus decreases by 25% with each additional presentation. This probabilistic rule helps in varying the stimuli exposure and maintaining a degree of unpredictability in the sequence of stimuli presented during the experiment.”
- Page 8, Line 314f: “Trains of angry faces were presented 76 times and trains of happy faces were presented 65 times.”
- Page 8, Line 315ff: “For the analysis, we consider emotion and repetition as experimental conditions, i.e. responses are averaged to collapse other dimensions of the stimuli such as identity, age, and gender.”

10. In the analyses for the electro-cortical signal, the authors will not consider laterality as factors, despite the cluster for the early window has it. Why?

Response: We will discuss the channel regions that show the effects in Study 1 and relate them to the published literature. The right laterality of the repetition effect will be discussed. Further, we include handedness measures for Study 2 to assess if handedness has an impact on the results. To this end, we will assess if there are differences in handedness between the lonely and non-lonely group and add handedness as a covariate in an exploratory model.

Changes to the manuscript:

- Page 12, Line 441ff: “The repetition effect in Study 1 was lateralised to the right hemisphere (see Figure 1c). To assess the impact of differences in brain lateralisation, we collected handedness information to assess the impact of handedness differences on the results of Study 2.”

11. What was the rationale for only considering the first and the fifth trials, and not the complete five trials?

Response: The mean amplitude of the ERP within the channel region and time window identified through spatiotemporal clustering decreases with the number of repetitions (see also response 4 to reviewer 1 and Figures 1 & 2 above). This decrease is not linear. We considered modelling the decrease and comparing the slope parameter of that function between the lonely and non-lonely group. However, we found that this approach was quite sensitive to noise. Therefore, we decided to pursue a more robust analytic strategy that would still allow us to test our hypotheses. We selected the 5th repetition as the highest repetition with the most trials in our paradigm.

Changes to the manuscript:

- Page 13, Line 463-467: “The mean amplitude in response to happy and angry faces decreased with the number of repetitions. To characterise the effect of repetition, we focused on the contrast between the first repetition and the highest repetition with the greatest number of trials, namely the 5th repetition.”

12. What was the rationale for setting the alpha level by 0.02 and not 0.05?

This alpha level is recommended for registered reports and is a requirement for some journals, including *Cortex* (Reference: [Guidelines for Authors](#))

13. The exploratory analyses are not very clear. It is read: “[...], we will conduct additional exploratory analyses that control for social anxiety and depression. If the conclusions change when these variables are included as predictors, [...]”. What are these “additional exploratory analyses”? Regressions? Which type of regression? And if a regression is calculated, why do the authors need to calculate a mediation analyses?

Response: Thank you for bringing this to our attention. We included a more detailed description of the exploratory analyses in the revised manuscript.

Changes to the manuscript:

- Page 12, Line 433-439: “For these control analyses, we employed an analysis of covariance (ANCOVA) model with ERP amplitude as the dependent variable, lonely versus non-lonely as the independent variable, and continuous social anxiety, depression, social isolation, and perceived stress scores as covariates. Significant effects of the covariates on ERP amplitudes were followed up with mediation analyses within each group (lonely and non-lonely). These analyses assessed the direct and indirect effects of continuous loneliness scores on ERP amplitudes, including each covariate separately as potential mediators.”