PEER REVIEW HISTORY

ARTICLE DETAILS

TITLE (PROVISIONAL)	Development and validation of an emotional picture set of self-injury (EPSI) for borderline personality disorder: protocol for a validation study
AUTHORS	Bachmann, Katharina; Schulze, Marcel; Sörös, Peter; Schmahl, Christian; Philipsen, Alexandra

VERSION 1 - REVIEW

REVIEWER	David Eddie Massachusetts General Hospital, Harvard Medical School, USA
REVIEW RETURNED	01-Nov-2018

| GENERAL COMMENTS | The authors are to be commended for proposing this important work. Such image sets are much needed by the field. There are numerous strengths to this study design. It utilizes well validated approaches and incorporates a diverse battery of measures. I have some comments and suggestions noted below. I hope these are useful.

The "This Study" section speaks to participants viewing images of before, during and after NSSI, but the methods suggest images will show, 1) objects, 2) just before NSSI, and 3) during NSSI. Please clarify which is true. If it is the former, how do you plan to ethically attain images of post-NSSI? If I may, based on my experience studying, diagnosing, and treating BPD, I would suggest using a combination of 1) objects, 2) just before NSSI, 3) during NSSI, and 4) after NSSI.

Consider rewording for clarity: "Those findings suggest that measuring emotion dysregulation in BPD might only arise in contexts that are psychologically challenging (Sloan et al., 2010; Suvak et al., 2012)."

"Statistical and Diagnostic Manual of Mental Disorders" should be Diagnostic and Statistical Manual of Mental Disorders

You may want to check out the following paper that was just released as it seems very relevant to the proposed research: Eddie, D., Bates, M. E., Vaschillo, E. G., Lehrer, P., Retkwa, M., & Miuccio, M. (2018). Rest, Reactivity, and Recovery: A Psychophysiological Assessment of Borderline Personality Disorder. Frontiers in Psychiatry, 9, 505. doi:10.3389/fpsyt.2018.00505
(SKID I,II) should be (SCID I,II).

Should "To assess the extend" be "To assess the extent"? |

The example images in the figures are very small (they probably got messed up during the conversion to PDF). Consider increasing their size.

Is the plan to have a rater in the room recording people's arousal ratings?

How and when will participants rate their self-reference with the images? Will this just be done at the end of the image set, or will participants provide rating for each image. If rating for each image, is the plan for participants to provide both an arousal rating and self-reference rating in between viewing each image? My recommendation would be to avoid this. Usually, in similar paradigms, we have shown participants an image set twice; once to get their SAM arousal ratings, and once to get their SAM valence ratings.

Figure 3 might be improved by showing the picture exposure paradigm in more detail so it's clear to the reader that participants will see each image for 5 sec and then have 5 sec of blank screen while the rate the image.

Consider also look at heart rate as a measure. It will provide important information about sympathetic arousal not captured by HF HRV.

Note, if you have 45 images in the NSSI image set, and 5sec on and 5sec off, you should have just under 5mins of ECG recording. This is a long enough recording to allow you to look at LF HRV in addition to HF HRV.

Consider including a very brief summary of how you plan to post-process the ECG recordings and handle HRV outliers. Even in lab-based paradigms, ECG recordings can be quite noisy so it's critical recordings are checked and if needed amended using HRV extraction software. Also, outliers can really affect HRV results so worth mentioning how you plan to check for outliers.

On line 353, I believe 'BPS' should be 'BPD'.

REVIEWER	Sofie Westling Lund University, Sweden
REVIEW RETURNED	21-Nov-2018

GENERAL COMMENTS	Ð Firstly – this is an important study, with an interesting approach and high likelihood to have an impact on future research. Ð No dates are included Ð Exclusion criteria somewhat insufficiently described: I assume that apart from life-time NSSI, also attempted suicide and perhaps also suicide ideation are exclusion criteria for the depressed and healthy controls. Perhaps state that any outcome above zero on ISAS or SHQ should be exclusion criterion for both non-BPD groups?

Reviewer#1

1. The "This Study" section speaks to participants viewing images of before, during and after NSSI, but the methods suggest images will show, 1) objects, 2) just before NSSI, and 3) during NSSI. Please clarify which is true. If it is the former, how do you plan to ethically attain images of post-NSSI? If I may, based on my experience studying, diagnosing, and treating BPD, I would suggest using a combination of 1) objects, 2) just before NSSI, 3) during NSSI, and 4) after NSSI.

Reply: We thank the reviewer for the comment. We realized that the term could be misleading; therefore, we changed it to 'during NSSI'. We thank the reviewer for his suggestion to also assess emotional response after NSSI. While we agree that assessing emotional response after NSSI would be very interesting we have some ethical and methodological concerns. Such as, that the experiment would take too long, we set a limit of 45 min to avoid overburdening of patients. Also, post-NSSI pictures might be even more distressing. However, we will keep this suggestion in mind for future trails.

2. Consider rewording for clarity: "Those findings suggest that measuring emotion dysregulation in BPD might only arise in contexts that are psychologically challenging (Sloan et al., 2010; Suvak et al., 2012)."

Reply

Thank you for your comment, we reworded the sentence to: "Those findings suggest that measuring emotional responses that are characteristic for BPD only make sense in contexts that are psychologically challenging"

3. "Statistical and Diagnostic Manual of Mental Disorders" should be Diagnostic and Statistical Manual of Mental Disorders

Reply: We corrected this in the manuscript.

4. You may want to check out the following paper that was just released as it seems very relevant to the proposed research: Eddie, D., Bates, M. E., Vaschillo, E. G., Lehrer, P., Retkwa, M., & Miuccio, M. (2018). Rest, Reactivity, and Recovery: A Psychophysiological Assessment of Borderline Personality Disorder. Frontiers in Psychiatry, 9, 505. doi:10.3389/fpsyt.2018.00505

Reply: Thank you very much for highlighting this interesting study. The results are very interesting and relevant to our work. Therefore we added the study to our references.

See line 100-101:

Emotional vulnerability in individuals with BPD is characterized by high sensitivity to emotional stimuli, unusual emotional intensity and a slow return to emotional baseline (emotions are long- lasting). In addition, the identification, expression, and inhibition of emotions are impaired [1-3].

See line 123-127:

Emotional stimuli that activate specific, self-relevant information seem to arouse a more intense emotional reaction than more general emotional stimuli [3 4]. Therefore, to elicit a distinctive and BPD-specific emotional response the stimulus material has to have a high relevance for persons with BPD and has to trigger sensitivities distinct for BPD [3 5].

5. (SKID I,II) should be (SCID I,II).

Reply: We corrected this.

6. Should "To assess the extend" be "To assess the extent"?

Reply:

We corrected this.

7. The example images in the figures are very small (they probably got messed up during the conversion to PDF). Consider increasing their size.

Reply:

Thank you very much for your comment. We followed your recommendation and increased the size of the pictures.

8. Is the plan to have a rater in the room recording people's arousal ratings?

Reply: It is planned to evaluate people's arousal in the following way: behavioral SAM ratings from the participant and their emotional face activation tracked by the facereader. The next step will be to calculate the correlation between the behavioral ratings and physiological face-activation in order to find out group-wise arousal levels.

9. How and when will participants rate their self-reference with the images? Will this just be done at the end of the image set, or will participants provide rating for each image. If rating for each image, is the plan for participants to provide both an arousal rating and self-reference rating in between viewing each image? My recommendation would be to avoid this. Usually, in similar paradigms, we have shown participants an image set twice; once to get their SAM arousal ratings, and once to get their SAM valence ratings.

Reply: We plan to have all three SAM-ratings directly after the images. However, having your point in mind, we decided to present the ratings pseudorandomized across patients. Consequently, participant A starts with the sequence of e.g., image – SAM arousal – SAM valence – SAM dominance. Participant B would start then for instance with the sequence image – SAM valence – SAM arousal – SAM dominance.

We added the sentence:' Image presentation and SAM rating-screens will be pseudorandomized across all categories.'

Furthermore we added an additional rating self-reference rating of the EPSI pictures on the end of the experiment (see line 272-274).

10. Figure 3 might be improved by showing the picture exposure paradigm in more detail so it's clear to the reader that participants will see each image for 5 sec and then have 5 sec of blank screen while the rate the image.

Reply: We have made a figure and extended Figure 3 by a more detailed paradigm overview.

11. Consider also look at heart rate as a measure. It will provide important information about sympathetic arousal not captured by HF HRV.

Reply: We agree to have a further look at heart rate as a measure, since Brosschot et al. (2003) could show that the physiological responses last longer after negative events (they also use a combined approach of SAM-ratings and HR).

We added the sentence: 'Further, heart rate reactivity will be calculated.'

12. Note, if you have 45 images in the NSSI image set, and 5sec on and 5sec off, you should have just under 5mins of ECG recording. This is a long enough recording to allow you to look at LF HRV in addition to HF HRV.

Reply: We agree, however, since we are also interested in heart rates (see point 11) for the different image categories, we believe that measuring for the whole stimulation time is appropriate.

13. Consider including a very brief summary of how you plan to post-process the ECG recordings and handle HRV outliers. Even in lab-based paradigms, ECG recordings can be quite noisy so it's critical recordings are checked and if needed amended using HRV extraction software. Also, outliers can really affect HRV results so worth mentioning how you plan to check for outliers.

Reply: We thank the reviewer for highlighting this important point. We included the software we will use and described how we will take care of the artifacts.

We added the statement:

'Processing of the HRV-data will be done with the Kubios-HRV software. A threshold based artefact correction algorithm, as it is implemented in the Kubios-software will be done. To separate ectopic and misplaced beats from the normal sinus rhythm, the automatic artefact correction algorithm will be used.'

Literature

Brosschot, J. F., & Thayer, J. F. (2003). Heart rate response is longer after negative emotions than after positive emotions. International journal of psychophysiology, 50(3), 181-187.

Reviewer#2

1. No dates are included

Reply:

We included table 1 with dates on the study flow.

2. Exclusion criteria somewhat insufficiently described: I assume that apart from life-time NSSI, also attempted suicide and perhaps also suicide ideation are exclusion criteria for the depressed and healthy controls. Perhaps state that any outcome above zero on ISAS or SHQ should be exclusion criterion for both non-BPD groups?

Reply:

We thank the reviewer for this important comment and added the suggested additional exclusion criteria.

The paragraph now reads as follows:

"The patients in the depressed control group need to have a depressive episode (depressive symptoms for at least > 2weeks). Depressive patients who also met diagnostic criteria for a psychotic disorder will be excluded. The control groups will be matched to the BPS group for age and sex. The healthy control group has not to exhibit a current psychiatric disorder or history of self-injury. Additional exclusion criteria for both control groups are attempted suicide or current suicide ideation. The control groups will be matched to the BPD group for age and sex."

In addition, we added the following to the diagnostic procedure section (line 233-235):

"To record the history and methods of self-injury, the Inventory of Statements about Self-Injury (ISAS) and the Self-Harm Behavior Questionnaire (SHQ) will be applied 37 38. Any outcome above zero on the ISAS or SHQ will be exclusion criteria."

VERSION 2 – REVIEW

REVIEWER	David Eddie
	Massachusetts General Hospital, Harvard Medical School
REVIEW RETURNED	21-Jan-2019

GENERAL COMMENTS	Thank you for carefully addressing the reviewer comments in this revision. I have no further comments to add.