PEER REVIEW HISTORY

ARTICLE DETAILS

TITLE (PROVISIONAL)	Study protocol for a randomized-controlled study on emotion regulation training for adolescents with major depression: the KONNI study
AUTHORS	Greimel, Ellen; Feldmann, Lisa; Piechaczek, Charlotte; Oort, Frans; Bartling, Juergen; Schulte-Rüther, Martin; Schulte-Körne, Gerd

VERSION 1 - REVIEW

REVIEWER	Nele A.J. De Witte Thomas More University of Applied Sciences, Belgium
REVIEW RETURNED	06-Jan-2020

GENERAL COMMENTS	There are indications that training emotion regulation could be beneficial for prevention and treatment of mental illness such as affective disorders. Currently, there are not enough RCTs, especially in youth, investigating the effects of the training of cognitive reappraisal. The design of the current study allows to assess the effect of cognitive reappraisal training (in addition to CBT) on depressive symptoms, stress, and affect in youth with major depression. Additionally, some neurobiological indices of interest for emotion regulation are included. I only have one remark regarding the primary outcome measures. The relevance of depressive symptoms, perceived stress and positive and negative affect as outcome measures is immediately clear. However, I was surprised to see rumination as the first primary outcome measure since the concept of rumination was not mentioned in the manuscript before. I agree that it would be interesting to assess ER (in daily life) after the training, but it is unclear why the authors chose to give rumination a prominent role (when it was not included in the other parts of the manuscript). If possible, it would also be an added value to have a second follow-up after a longer period of time to investigate effects in the long term. Generally, the manuscript is well written and describes the procedures and instruments clearly. I also believe the study is relevant to the field.

REVIEWER	Dr Daniel Bressington The Hong Kong Polytechnic University, Hong Kong SAR.
REVIEW RETURNED	12-Jan-2020

GENERAL COMMENTS	Thank you for submitting your manuscript to BMJ Open. The paper reports a protocol for an RCT of emotion regulation training for adolescents with major depression. The protocol is generally very well written and includes most pertinent information. Referencing sometimes mentions “but see”, for example: “[15, 16, but see 17]”. Similarly, “also see XX”. I am not familiar with this use of terminology and wonder why it is necessary to include “but see, also see etc.” if all references support the point being made. This may require revision to improve clarity. A statistician should review the analysis plan because I am not sure that ANOVAs are sufficient, particularly where there may be missing data or data do not meet the assumptions for ANOVA, perhaps GEE should be considered. The rationale for the study is clear and well supported with published work in the area. Please provide some more information about the recruitment procedure. It is stated that they are recruited form the from the waiting list, or outpatients. Who will do this, and will they be randomly selected or will all patients meeting the criteria be approached? Also, will potential participants'' capacity to provide informed consent be assessed and by whom? Please provide more information about how allocation concealment will be ensured, i.e. who will keep the allocation list and inform the researchers/participants about group allocation? How will inter-rater reliability of outcome measure raters (specifically the PANAS-C-SF and diagnostic interviews) be established? Brief details about the validity/reliability/psychometric properties of outcome measures could be provided in relation to the study population. Sample size calculation is well justified. Please provide brief information about how missing data will be managed/imputed etc. Please mention how patient safety data (i.e. adverse incidents) will be monitored and reported. Similarly, will any feasibility data be collected/reported? Please outline the inherent methodological limitations of this study (particularly the very short follow-up period, single blinding) and how these may potentially introduce bias.

REVIEWER	Atsuo Nakagawa Keio University School of Medicine Clinical and Translational Research Center Japan
REVIEW RETURNED	14-Feb-2020

GENERAL COMMENTS	Study protocol for a randomized-controlled study on emotion regulation training for adolescents with major depression: the KONNI study The present study protocol aims to test whether a task-based

	training in cognitive reappraisal is effective in adolescents with major depression. 1) Before conducting this RCT, did the authors conduct a feasibility study or any clinical observation to see whether cognitive reappraisal program (training session) benefits for the adolescent population? 2) The calculation of the sample size was driven by the previous study that studied for the adult population. In terms of conducting to the adolescent population, is there any modification regarding the program? 3) The authors indicate that “treatment fidelity concerning the training is assured by standardized oral and written instructions and by comprehensive training of the experimenters.” ・Was a treatment manual for this cognitive reappraisal program (training session) developed for this study? ・Who will be the experimenters? Clinical psychologists?

VERSION 1 – AUTHOR RESPONSE

Reviewer: 1
Reviewer Name: Nele A.J. De Witte
Institution and Country: Thomas More University of Applied Sciences, Belgium
Please state any competing interests or state ‘None declared’: None declared

There are indications that training emotion regulation could be beneficial for prevention and treatment of mental illness such as affective disorders. Currently, there are not enough RCTs, especially in youth, investigating the effects of the training of cognitive reappraisal. The design of the current study allows to assess the effect of cognitive reappraisal training (in addition to CBT) on depressive symptoms, stress, and affect in youth with major depression. Additionally, some neurobiological indices of interest for emotion regulation are included. I only have one remark regarding the primary outcome measures. The relevance of depressive symptoms, perceived stress and positive and negative affect as outcome measures is immediately clear. However, I was surprised to see rumination as the first primary outcome measure since the concept of rumination was not mentioned in the manuscript before. I agree that it would be interesting to assess ER (in daily life) after the training, but it is unclear why the authors chose to give rumination a prominent role (when it was not included in the other parts of the manuscript). If possible, it would also be an added value to have a second follow-up after a longer period of time to investigate effects in the long term. Generally, the manuscript is well written and describes the procedures and instruments clearly. I also believe the study is relevant to the field.

We the thank the Reviewer for the positive evaluation of our study protocol and your important suggestion to include information about rumination already in the introduction. We added the following to our introduction section:

p. 3: “Questionnaire studies have found that depressive symptoms are associated with less habitual use of CR and more use of maladaptive emotion regulation strategies such as rumination.[12-14] It has been proposed that this habitual pattern in depressed individuals might originate from deficits in the inhibition of negative material which is thought to enhance ruminative thoughts but hampering the effective reappraisal of negative information.[15]”

p. 4-5: “While these results are encouraging, it needs to be examined whether the findings can be extended to MD samples and whether a CR training results in reductions in depressive symptomatology, including ruminative thoughts. In this context, it has been proposed that CR training improves cognitive control abilities, including the ability to inhibit negative material. As impairments in the ability to inhibit negative information are thought to play a causal role in rumination, training the ability to reappraise negative information should thus reduce ruminative thoughts.[15]”

We completely agree with the reviewer that it would be of added value to have a second follow-up after a longer period of time to investigate effects in the long term. However, as this study has already been initiated (the first participant was enrolled in May 27th, 2019) and we only have sought ethical approval for the two-week follow-up, this is not possible. However, in response to your review, we highlight the short follow-up interval as a limitation of the study (p. 16):

“A limiting factor of the study is the short follow-up interval of two weeks. Thus, future studies should include a longer follow-up interval to also examine whether the effects of the training are long-lasting.”

Reviewer: 2
Reviewer Name: Dr Daniel Bressington
Institution and Country: The Hong Kong Polytechnic University, Hong Kong SAR.
Please state any competing interests or state ‘None declared’: None declared.

Thank you for submitting your manuscript to BMJ Open. The paper reports a protocol for an RCT of emotion regulation training for adolescents with major depression.
The protocol is generally very well written and includes most pertinent information.

We thank the Reviewer for the positive evaluation of our study protocol.

Referencing sometimes mentions “but see”, for example: “[15, 16, but see 17]”. Similarly, “also see XX”. I am not familiar with this use of terminology and wonder why it is necessary to include “but see, also see etc.” if all references support the point being made. This may require revision to improve clarity.

We now revised the referencing accordingly throughout the manuscript.

A statistician should review the analysis plan because I am not sure that ANOVAs are sufficient, particularly where there may be missing data or data do not meet the assumptions for ANOVA, perhaps GEE should be considered.

We thank the Reviewer for raising this issue. Frans J. Oort, who is part of the project team and also co-authored the current manuscript, is professor of Methods and Statistics at the University of Amsterdam.

Please note that on p. 14-15, we do not suggest a regular (repeated measures) ANOVAs, but mixed-model ANOVAs, also known as multilevel analysis, with observations 'nested' within participants (Murrar & Brauer, 2018). We assume that the distributions of our outcome variables are (approximately) normal. However, we will of course check this assumption, and if our data screening shows substantial deviations from normality, we may resort to GEEs instead.

To make this evident for the reader, we slightly changed the text as follows:

p. 14: Mixed modeling (also known as multilevel analysis, with observations “nested” within participants) has the advantage over regular repeated measures ANOVA that all available data can

be used, including data from incomplete cases, without using imputation techniques for missing data.[67]

The rationale for the study is clear and well supported with published work in the area.

We thank the Reviewer for positive evaluation of our study rationale.

Please provide some more information about the recruitment procedure. It is stated that they are recruited form the from the waiting list, or outpatients. Who will do this, and will they be randomly selected or will all patients meeting the criteria be approached? Also, will potential participants'' capacity to provide informed consent be assessed and by whom?

We now added more information about the recruitment procedure (p. 6-7):

“Adolescents and one parent/legal custodian (for participants <18 years) will be contacted by an experienced study nurse certified in Good Clinical Practice (GCP) and will be informed about the study details, including the fact that the allocation to the training groups will be made based on a predefined randomization list. All potential study participants and their parents/legal custodians will be approached by the study nurse unless it is known beforehand that the exclusion criteria are met (e.g., acute suicidality, gender dysphoria). If the clinicians of potential participants conclude that the capacity to provide informed consent/assent are not met (e.g., in case of insufficient German skills, cognitive disability or an acute crisis), participants and their parents/legal custodians will not be approached. In case of interest in participation, their written informed consent/assent will be collected.”

Please provide more information about how allocation concealment will be ensured, i.e. who will keep the allocation list and inform the researchers/participants about group allocation?

As suggested by the Reviewer, we now provide more information about the allocation concealment (p. 7):

„Access to the allocation list is limited to the principal investigator (E.G.) and her deputy (L.F.), who will inform the experimenters about the allocation of the participant shortly before the first training session (after the diagnostic session and the decision to include the participant in the study). Randomization stratifying for age (<15 years vs. ≥15 years of age) and sex will be performed by a statistician, who is neither involved in recruitment nor in testing of participants. The randomization will be performed with a 1:1 allocation. A follow-up session will take place two weeks after completion of the forth training session. After the follow-up, participants are unblinded regarding group allocation by one of the experimenters.”

How will inter-rater reliability of outcome measure raters (specifically the PANAS-C-SF and diagnostic interviews) be established?

We thank the Reviewer for raising this point. The PANAS-C-SF is filled out by the participants and will be analysed according to a predefined scheme (Ebesutani et al., 2012). Likewise, all other outcome measures will be analysed based solely on objective criteria/predefined schemes. Therefore, inter-rater reliability does not need to be established for these measures. We will perform a quality check for 5% of the entered data to examine whether the data was entered correctly.
Please note that the diagnostic interview (Kinder-DIPS) does not serve as outcome measure. It is applied to assess the diagnosis of MD and other psychiatric disorders and thus to check for inclusion and exclusion criteria (see p. 10). To assess inter-rater reliability based on Cohen´s kappa (k), 10% of

the Kinder-DIPS interviews will be rated by a second experimenter. We now address this point in the manuscript (p. 10):

“The interview will be administered by experienced, psychologically trained experimentersstaff. To assess inter-rater reliability based on Cohen´s kappa (k), 10% of the Kinder-DIPS interviews will be rated by two experimenters.”

Brief details about the validity/reliability/psychometric properties of outcome measures could be provided in relation to the study population.

We thank the Reviewer for bringing up the important question of validity and reliability of outcome measures in relation to the study population. We now added relevant information on outcome measures to Table 1 and also state in which population the psychometric properties were assessed. For allmost all primary outcome measures, psychometric data is available from (clinical) adolescent populations.

Sample size calculation is well justified.

We thank the Reviewer for the positive remark.

Please provide brief information about how missing data will be managed/imputed etc.

As outlined in our response to Reviewer 2 regarding the data analysis plan, we will conduct mixed-model ANOVAs, also known as multilevel analysis, with observations 'nested' within participants. One of the advantages of mixed-modeling (and GEE) is that there is no need to impute data, but that all available data can be used, including data from incomplete cases, without using imputation techniques for missing data (Murrar & Brauer, 2018). We highlight this issue in the analysis plan (p. 14):

„Mixed modeling (also known as multilevel analysis, with observations “nested” within participants) has the advantage over regular repeated measures ANOVA that all available data can be used, including data from incomplete cases, without using imputation techniques for missing data.[67]”

Please mention how patient safety data (i.e. adverse incidents) will be monitored and reported.

We thank the Reviewer for addressing this point. We refer to this issue as follows (p. 8):

“The concomitant treatment as usual is permitted during the ongoing study and information on the type of treatment during the study will be assessed. Moreover, patient safety data will be assessed by recording along with any spontaneously reported adverse effects.”

Similarly, will any feasibility data be collected/reported?

The collection and report of the feasibility data is now outlined in the manuscript (p. 8):

“To assess the feasibility of conducting a large-scale multi-center RCT on the effects of a CR training in addition to standard treatment, the following data will be collected and reported: participation and non-participation rate, drop-outs and reasons for drop-outs, training attendance rates and spontaneously reported adverse effects.”

Moreover, before starting the present RCT, we conducted two feasibility studies to test whether adolescents with and without major depression are able to down-regulate negative affective

responses to negative pictures via CR (Greimel, Piechaczek, Schulte-Rüther, Feldmann, & Schulte-Körne, 2020; Piechaczek et al., in prep.). In these studies, a single training session was conducted that was very similar to a CR training session as outlined in the present study protocol. Results of our studies demonstrated that both adolescents with and without major depression a) understood and complied with the task instructions, and b) were are able to diminish negative affective responses via CR.

We now refer to our work that has already been published (Greimel et al., 2020) and added the following sentence to the manuscript (p. 8):

“The CR training task is well-established and is adapted from previous studies,[22, 33-35] including a study from our group, in which we demonstrated that adolescents with MD understand and comply with task instructions, and are able to down-regulate negative affective responses to negative pictures via CR.[36] The training procedure was adapted from [30].”

Please outline the inherent methodological limitations of this study (particularly the very short follow-up period, single blinding) and how these may potentially introduce bias.

As suggested by the Reviewer, we added the following limitations to our discussion section (p. 16):

“A limiting factor of the study is the short follow-up interval of two weeks. Thus, future studies should include a longer follow-up interval to also examine whether the effects of the training are long-lasting. Another limitation is that the study is single-blinded (participant-blinded) concerning the allocation to the CR training vs. control training. This single-blinding procedure entails the risk that the experimenters will transfer their expectations to the participants. However, as the participants will perform a comprehensive practice training that is guided by the experimenter, double-blinding would not be feasible. Finally, it should be stated the present study does not include the ecological momentary assessment of outcome measures. Expanding upon the present study, it would be important to also apply experience sampling methods in future work to be able to draw comprehensive conclusions regarding transfer effects of the CR training to daily live.”

Reviewer: 3
Reviewer Name: Atsuo Nakagawa
Institution and Country: Keio University School of Medicine Clinical and Translational Research Center, Japan
Please state any competing interests or state ‘None declared’: None

Study protocol for a randomized-controlled study on emotion regulation training for adolescents with major depression: the KONNI study

The present study protocol aims to test whether a task-based training in cognitive reappraisal is effective in adolescents with major depression.

1) Before conducting this RCT, did the authors conduct a feasibility study or any clinical observation to see whether cognitive reappraisal program (training session) benefits for the adolescent population?
Before starting the present RCT, we conducted two feasibility studies to test whether adolescents with and without major depression are able to down-regulate negative affective responses to negative pictures via CR (Greimel et al., 2020; Piechaczek et al., in prep.). In these studies, a single training session was conducted that was very similar to a CR training session as outlined in the present study protocol. Results of our studies demonstrated that both adolescents with and without major

depression a) understood and complied with the task instructions, and b) were are able to diminish negative affective responses via CR.
We now refer to our work that has already been published (Greimel et al., 2020) and added the following sentence to the manuscript (p. 8):

“The CR training task is well-established and is adapted from previous studies,[22, 33-35] including a study from our group, in which we demonstrated that adolescents with MD understand and comply with task instructions, and are able to down-regulate negative affective responses to negative pictures via CR.[36] The training procedure was adapted from.[30]”

2) The calculation of the sample size was driven by the previous study that studied for the adult population. In terms of conducting to the adolescent population, is there any modification regarding the program?

We thank the Reviewer for raising this important point. The calculation of our sample size is based on two studies in adults (Siegle et al., 2014; Denny & Ochsner, 2014). The training procedure is closely adapted from the study by Denny & Ochsner (2014), who also applied four training sessions in CR and a control training. As outlined to the Reviewer´s previous point, we conducted two studies prior to the present RCT (Greimel et al., 2020; Piechaczek et al., in prep.) that confirmed the feasibility of the training task as such, albeit these studies were restricted to one session.
To adapt the training to adolescents, we made a couple of modifications. This includes that we use a rating scale for the assessment of affective responses which is often applied in studies with children and adolescents (SAM scale; Musser, Galloway-Long, Frick, & Nigg, 2013; Reichel et al., 2014). Furthermore, we only chose pictures which are appropriate for adolescents, e.g., excluding pictures of dead persons or pornographic images. Related to this point, we deliberately also included pictures from a stimulus set that is tailored to adolescents (Besançon Affective Picture Set-Adolescents BAPS-Ado; Szymanska et al., 2015). Another modification is that we monitor adherence to task instructions in our adolescent population. For this reason, participants of our study will fill in a questionnaire after each training session to indicate which strategies they used during the task.

We refer to these issues in the manuscript as follows:

p. 8: Following each picture, participants are instructed to indicate their affective response to the image on the portrait version of the nine-point self-assessment manikin scale for valance [SAM; 37, for the portrait version see 38, 39], which has been frequently applied in youth samples.[40,41]

p. 9: To ensure adherence to task instructions, participants will fill in a questionnaire after each training session to indicate which strategies they used during the task.

p. 9: Developmentally appropriate pictures (e.g., excluding pictures of dead persons or pornographic images) are taken from the International Affective Picture System,[IAPS; 43] Besançon Affective Picture Set-Adolescents,[BAPS-Ado; 44] and Besançon Affective Picture Set-Adults,[BAPS-Adult; 45] with the latter two sets being derived from the Besançon Attachment Pictures Set.

3) The authors indicate that “treatment fidelity concerning the training is assured by standardized oral and written instructions and by comprehensive training of the experimenters.”

・Was a treatment manual for this cognitive reappraisal program (training session) developed for this study?

We thank the Reviewer for addressing this point. The training procedure for instructing participants was adapted from Denny and Ochsner (2014) and comprehensibility of this adapted version was examined and confirmed in our previous studies (Greimel et al., 2020; Piechaczek et al., in prep.). We added more information on this issue to the following section in the manuscript (p. 8):

“The CR training task is well-established and is adapted from previous studies,[22, 33-35] including a study from our group, in which we demonstrated that adolescents with MD understand and comply with task instructions, and are able to down-regulate negative affective responses to negative pictures via CR.[36] The training procedure was adapted from [30].”

・Who will be the experimenters? Clinical psychologists?

The experimenters are either clinical psychologists or advanced and in-depth trained psychology students. The following information was added to the manuscript (p. 8):

„No other criteria for discontinuation are defined. Treatment fidelity concerning the training is assured by standardized oral and written instructions and by comprehensive training of the experimenters. The experimenters will be either clinical psychologists or advanced and in-depth trained psychology students.”

References

Denny, B. T., & Ochsner, K. N. (2014). Behavioral effects of longitudinal training in cognitive reappraisal. Emotion 14(2), 425-433.
Ebesutani, C., Regan, J., Smith, A., Reise, S., Higa-McMillan, C., & Chorpita, B. F. (2012). The 10-Item Positive and Negative Affect Schedule for Children, Child and Parent Shortened Versions: Application of Item Response Theory for More Efficient Assessment. Journal of Psychopathology and Behavioral Assessment, 34(2), 191-203.
Greimel, E., Piechaczek, C., Schulte-Rüther, M., Feldmann, L., & Schulte-Körne, G. (2020). The role of attentional deployment during distancing in adolescents with major depression. Behaviour research and therapy, 126:103554.
Murrar, S., & Brauer, M. (2018). Mixed model analysis of variance. In B. B. Frey (Ed.), The SAGE encyclopedia of educational research, measurement, and evaluation (pp. 1075-1078). Thousand Oaks: SAGE Publications, Inc.
Musser, E. D., Galloway-Long, H. S., Frick, P. J., & Nigg, J. T. (2013). Emotion regulation and heterogeneity in attention-deficit/hyperactivity disorder. Journal of the American Academy of Child & Adolescent Psychiatry, 52(2), 163-171.e162.
Piechaczek, C., Schröder, P.T., Feldmann, L., Schulte-Körne, G., & Greimel, E. (in preparation). The effects of attentional deployment on reinterpretation in depressed and healthy adolescents: evidence from an eye-tracking study.
Reichel, V. A., Schneider, N., Grünewald, B., Kienast, T., Pfeiffer, E., Lehmkuhl, U., & Korte, A. (2014). “Glass fairies” and “bone children”: Adolescents and young adults with anorexia nervosa show positive reactions towards extremely emaciated body pictures measured by the startle reflex paradigm. Psychophysiology, 51(2), 168-177.
Szymanska, M., Monnin, J., Noiret, N., Tio, G., Galdon, L., Laurent, E., . . . Vulliez-Coady, L. (2015). The Besancon affective picture set-adolescents (the BAPS-Ado): development and validation. Psychiatry research, 228(3), 576-584.

VERSION 2 – REVIEW

REVIEWER	Dr Daniel Bressington The Hong Kong Polytechnic University, Hong Kong, SAR China
REVIEW RETURNED	18-Apr-2020

GENERAL COMMENTS	Thank you for resubmitting your manuscript. All previous comments have been well-addressed.

REVIEWER	Atsuo Nakagawa Keio University School of Medicine, Clinical and Translational Research Center, Japan
REVIEW RETURNED	17-Apr-2020

GENERAL COMMENTS	BMJ Open 2019-03-6093.R1 Study protocol for a randomized-controlled study on emotion regulation training for adolescents with major depression: the KONNI study The present study protocol aims to test whether a task-based training in cognitive reappraisal is effective in adolescents with major depression. The revised manuscript is well written and describes the procedures and instruments clearly, and have clarified all the points I have raised in the initial manuscript. 1) Before conducting this RCT, did the authors conduct a feasibility study or any clinical observation to see whether cognitive reappraisal program (training session) benefits for the adolescent population? ➡I am happy that the authors have clarified this in the revised manuscript. 2) The calculation of the sample size was driven by the previous study that studied for the adult population. In terms of conducting to the adolescent population, is there any modification regarding the program? ➡I am happy that the authors have addressed these points in the revised manuscript. 3) The authors indicate that “treatment fidelity concerning the training is assured by standardized oral and written instructions and by comprehensive training of the experimenters.” ・Was a treatment manual for this cognitive reappraisal program (training session) developed for this study? ・Who will be the experimenters? Clinical psychologists? ➡I am happy that the authors have clarified these points in the revised manuscript.