PEER REVIEW HISTORY

ARTICLE DETAILS

TITLE (PROVISIONAL)	Adverse Childhood Experiences (ACEs) and Later-life Depression: Perceived Social Support as a Potential Protective Factor
AUTHORS	Cheong, E Von; Sinnott, Carol; Dahly, Darren; Kearney, Patricia

VERSION 1 - REVIEW

REVIEWER	Jochen Hardt Uni Mainz, Germany
REVIEW RETURNED	17-Nov-2016

GENERAL COMMENTS	Title: Adverse Childhood Experiences (ACEs) and Later-life Depression: Perceived Social Support as a Potential Protective Factor Corresponding author: Cheong E Von The paper reports associations between three categories of childhood adversities (ACE) and later life depression in a large Irish sample, moderated by social support. It is an interesting paper, clearly worth to become published. I have, however, some concerns. 1) On page 3, what this study adds, and on page 4, results, as well as in the discussion, the authors focus strongly on the interaction between ACE's in general and social support. Data show that it is significant only for abuse. I think this is misleading and would recommend to change it. 2) On page 3, lines33ff, the authors report the well known fact that it is perceived social support what helps, not the available one. The authors conclude that enhancing the perception of social support may be similarly important than providing social support itself. I would not think so. Various aspects need to be considered to explain this fact, among them measurement issues. 3) Data analysis: First, the authors put the 10 ACE's into three categories and perform all analyses based on them. Doing so hides the impact of the individual ACE's. I would rather suggest to perform a regression of depression on the 10 individual ACE's including their interactions with social support. Then performing a backward selection, with removing non significant interaction terms first, then removing non significant main effects. If none single one reaches significance, then a score may be tested. 4) The CESD score was dichotomized, the social support scale grouped into three categories. Previous reviewers criticized the latter, I am rather concerned about the former. There is a rationale to categorize a moderator variable, as interactions become easier to

	detect then. Utilizing a trichotomization has my sympathy, but goes on cost of power compared to a dichotomization. Dichotomizing a response variable additionally leads to a loss of power. There is a bulk of literature showing this, and there are not much advantages in dichotomization, here. 5) Page 11, line17ff: t-test, anova and X2-test are no descriptive statistics. 6) Page 11, line36ff: The analyses which were performed do not correspond to the DAG displayed. Cox and Wermuth (Cox, D. R. & Wermuth, N. (1996). Multivariate dependencies: models, analysis and interpretation. London, Chapman & Hall) developed a method to test such a DAG containing interactions. 7) Page 11, line 50ff: The description how missing data were handled is not clear to me. How would a case be treated having one item missing in the CESD? Additionally, imputing fractional polynomials would not be the first choice if interactions are of primary interest. Conditional imputation (on the three levels of social support), including the interaction terms "just as another variable(JAV)" or passive imputation would be better choices, see van Buuren, S. (2012). Flexible imputation of missing data. Boca Raton, CRC Press (Chapman & Hall), p133ff. 8) Page12, lines 30ff: the sample size is not exactly clear to me. 2047 filled out the baseline questionnaire, 1926 the ACE's. 16.1% had a CESD>cut off, if this are 302, I come to a total sample of 1876? If I add up the data for depression in table 1, I come to (309+61+64+1153+125+81=) 1793. Additionally, it would be good to include the numbers of the social support categories into table 1. Otherwise table 1 is great! 9) Table 2 does not show any significant interaction between ACE and social support. In the appendix, there are some significant ones, but if I interprete the data right, all of them are due to abuse. Additionally, testing the same variables in various ways (any ACE, total score and sub scores) requires a Bonferroni correction. 10) Discussion: For me the most interesting result is that there are (almost) no buffering effects of social support for neglect and household dysfunction. I ask myself why? One possible explanation would be that the Oslo social support scale does not capture such an effect, but when I look onto the items I would not think so. Another possible explanation would be that the various childhood adversities are different to cope with. This would mean that we should stronger focus on the non-abuse aspects, in particular on neglect. Maybe such thoughts would enrich the discussion. Finally, for me it would be interesting to see if the various aspects of abuse (emotional, physical or sexual) are similarly moderated by social support or if there are differences, too.

REVIEWER	Soonhee Roh University of South Dakota USA
REVIEW RETURNED	18-Nov-2016

GENERAL COMMENTS	Title: Adverse Childhood Experiences (ACEs) and Later-life

	Depression: Perceived Social Support as a Potential Protective Factor The paper deals with an important topic. Overall it is a well-written, interesting, and timely appropriate paper. The purpose of the paper is clear. I only have the following three comments: 1. This article uses the stress sensitisation theory that suggests that childhood adversity reduces an individual's threshold for developing depressive reactions towards stressful events, causing one to have depressive reactions towards current mild stressors or greater reactivity towards severe stressful events. Introduction was very good and need more strong literature review focused on older adults rather than young women example. 2) Need more strong discussions about the significance of the study. 3) Need to elaborate implications of practices based on the findings much more.

VERSION 1 – AUTHOR RESPONSE

Response to Reviewer 1.

The paper reports associations between three categories of childhood adversities (ACE) and later life depression in a large Irish sample, moderated by social support. It is an interesting paper, clearly worth to become published. I have, however, some concerns.

1) On page 3, what this study adds, and on page 4, results, as well as in the discussion, the authors focus strongly on the interaction between ACE's in general and social support. Data show that it is significant only for abuse. I think this is misleading and would recommend to change it.

We agree that the interaction is very clear for those reporting abuse, and much less so for those reporting neglect or household dysfunction. We have edited the text as suggested, and moderated our final recommendations, as follows:

Page 2, What the study adds:

This interaction was primarily driven by individuals reporting ACEs reflecting abuse, and was much less pronounced, and perhaps non-existent, in those reporting neglect or household dysfunction.

Page 3, Abstract, results:

This pattern of results was similar when exposures were based on ACE subtype and ACE scores, though the interaction was clearly strongest among those reporting abuse.

Page 19, Discussion, first paragraph:

Importantly, we found that the deleterious impact of ACEs was typically limited to those individuals who also reported poor and moderate PSS. However, the statistical evidence for this interaction was only strong among those reporting abuse.

Page 20, 3rd paragraph:

Consistent with the stress buffering model(Cohen, 1985), the relationship between ACE (overall, subtype or ACE score) and depression later in life were found to vary according to level of perceived social support, though this interaction was clearly strongest amongst those reporting abuse.

Further, to better reflect Professor Hardt's point, we have added the p-values from the LRT chi2 tests for the interaction terms in each model to the respective tables. However, it's worth noting that recoding of perceived social support (PSS) to more parsimoniously capture this interaction (and reduce the df for the LRT tests) would likely result in "significant" interactions. Further, linear models tend to be less well powered for interactions than main effects. We have this persisted in including them in the reported models and would like to leave it for the reader to make their own judgements.

2) On page 3, lines33ff, the authors report the well known fact that it is perceived social support what helps, not the available one. The authors conclude that enhancing the perception of social support may be similarly important than providing social support itself. I would not think so. Various aspects need to be considered to explain this fact, among them measurement issues.

We agree that social support is complex as it encompasses both perceived and received support. We have added additional detail on this in the background:

Introduction, pages 5 and 6:

The term social support encompasses perceived and received support. It is suggested that perceived support is best understood as an individual difference variable, with evidence that those who report that others will provide them with aid when they are in need (perceived social support) are protected from the pathogenic effects of life stress (Cohen, 2004). Studies on perceived social support have consistently shown it to be associated with reduced stress and improved physical and mental health (Haber et al 2007), and that perception of available social support was found to be a better buffer of psychological distress than actual availability of social support in some studies (Evans SE et al 2013, Cohen S, 1985).

We agree that enhancing support needs to address various aspects including perceptions, and that we haven't directly measured received support, and so can't separate them. We have edited "What this study adds" and the conclusions to reflect this:

Page 2, What the study adds:

These results suggest that interventions that aim to prevent poor mental health outcomes among survivors of childhood adversity, particularly abuse, might benefit from focusing on strengthening social support, or possibly perceptions of social support.

Pages 3 and 4, Abstract, conclusions:

Interventions that enhance social support, or possibly perceptions of social support, may help reduce the burden of depression in older populations with ACE exposure, particularly in those reporting abuse.

3) Data analysis: First, the authors put the 10 ACE's into three categories and perform all analyses based on them. Doing so hides the impact of the individual ACE's. I would rather suggest to perform a regression of depression on the 10 individual ACE's including their interactions with social support. Then performing a backward selection, with removing non-significant interaction terms first, then

removing non significant main effects. If none single one reaches significance, then a score may be tested.

We agree that the associations between the 10 individual ACEs, perceived social support, and depressive symptoms are important. We estimated those 10 models and summarized the results in the main text, and given the full set of estimated ORs and 95% CIs for each model in the supplemental material. However, we don't agree in principle on this step-wise approach to model selection, because we think that small cell sizes will lead to more volatile estimates that will disproportionately affect the nature of the "final" model that is reported.

Page 17, results, paragraph 3
Regarding individual ACEs, those reflecting abuse (Supplemental Table 6) and neglect (Supplemental Table 7) tended to be more strongly associated with depressive symptoms than those reflecting the various forms of household dysfunction (Supplemental Table 8). Further, the effect modification by PSS was most clearly demonstrated for the estimated effects of abuse.

4) The CESD score was dichotomized, the social support scale grouped into three categories. Previous reviewers criticized the latter, I am rather concerned about the former. There is a rationale to categorize a moderator variable, as interactions become easier to detect then. Utilizing a trichotomization has my sympathy, but goes on cost of power compared to a dichotomization. Dichotomizing a response variable additionally leads to a loss of power. There is a bulk of literature showing this, and there are not much advantages in dichotomization, here.

We of course agree that categorization must result in a loss of power. However, the CES-D is a construct designed to facilitate categorization of people, in contrast to a variable like blood pressure for example which is inherently continuous characteristic in nature. Moreover, the CESD we used the standard cutoff to dichotomize the scores (16 or greater) that are known to identify individuals at risk for clinical depression, with good sensitivity and specificity and high internal consistency (Lewinsohn, Seeley, Roberts, & Allen, 1997). Thus we have chosen to retain the models featuring the categorized variables in the main analysis. However, we have added a model that uses both the continuous CES-D and PSS scores in the supplemental information, and have noted that it qualitatively agrees with the other reported models.

5) Page 11, line17ff: t-test, anova and X2-test are no descriptive statistics.

Thank you, this has been corrected by editing the subheading.

6) Page 11, line36ff: The analyses which were performed do not correspond to the DAG displayed. Cox and Wermuth (Cox, D. R. & Wermuth, N. (1996). Multivariate dependencies: models, analysis and interpretation. London, Chapman & Hall) developed a method to test such a DAG containing interactions.

Our use of DAGs was in the non-parametric sense promoted by Greenland, Robins, and Pearl (2001) as a heuristic tool for identifying potential confounders of a key exposure-outcome association, rather than a graph intended to specify a complete statistical model describing a set of associations. We would maintain that our use of a DAG improved our own thinking about the hypothesized causal relationships amongst the variables, but we also don't want to overstate its value. Ultimately, what is at question is "Did we control for the right set of variables to satisfactorily address confounding" and we have added a statement in the weaknesses to reflect our thinking here.

Page 22, discussion, paragraph 2

Lastly, and most importantly, this is an observational study where both the exposure and outcomes will certainly share causes. We have tried to adjust for this through the careful selection and adjustment for confounders, but these in turn will be measured with some error and will certainly not represent an optimal set of covariates to adjust for, so it is important that these results are viewed as part of a larger and still developing body of research.

7) Page 11, line 50ff: The description how missing data were handled is not clear to me. How would a case be treated having one item missing in the CESD? Additionally, imputing fractional polynomials would not be the first choice if interactions are of primary interest. Conditional imputation (on the three levels of social support), including the interaction terms "just as another variable(JAV)" or passive imputation would be better choices, see van Buuren, S. (2012). Flexible imputation of missing data. Boca Raton, CRC Press (Chapman & Hall), p133ff.

Regarding the CESD, we only have data on the final scores. Instead of passive imputation, we have taken the “transform then impute approach” demonstrated in Von Hippel (2009), or synonymously the “Just another variable” approach described by White IR, Royston P, Wood AM (2011). The inclusion of fractional polynomials, which only affects the continuous covariate age, is included in the imputation model to ensure that it is richer (or at least as rich) as the final analysis model. On this point, we neglected to mention that the interaction between ACE and PSS was included in the imputation model, as it should be, and we have now corrected this oversight and edited the text to clarify the process we used.

Page 12, paragraph 1:

Missing data were handled using multiple imputation, so that all participants who completed the baseline questionnaire were included in the analytical sample, even if they were missing values for one or more variables. For each estimated statistical model, thirty imputed datasets were created, after a burn-in of 30 replications using predicted mean matching (Van Buuren S, 2012). Each imputation model included all variables used in a given statistical model, allowed for non-linear relationships using restricted cubic splines with 3 knots, and included the key interaction of interest (ACE X PSS). The statistical model of interest was them estimated in each imputed dataset, and parameter estimates were combined using Rubin’s rules (Rubin DB, 1987). A complete case sensitivity analysis was also performed for comparison.

8) Page12, lines 30ff: the sample size is not exactly clear to me. 2047 filled out the baseline questionnaire, 1926 the ACE's. 16.1% had a CESD>cut off, if this are 302, I come to a total sample of 1876? If I add up the data for depression in table 1, I come to (309+61+64+1153+125+81=) 1793. Additionally, it would be good to include the numbers of the social support categories into table 1. Otherwise table 1 is great!

If you add the 1793 with values for depressive symptoms to the 133 of those missing these data, you wind up at the 1926 who completed the ACE questionnaire. So the analytical sample is still the 2047, with MI used to account for the missing responses for ACEs, CESD, and other covariates. The MI estimates, assuming they are correctly modelled, are valid under MAR assumptions, but there is certainly a non-negligible chance data on ACE or CESD are NMAR, and we have noted this in the discussion.

Page 22, paragraph 1:

There was a non-negligible amount of missing data, which is not uncommon for such studies. We have used multiple imputation, rather than case-wise deletion, to both improve the efficiency of

analyses (by retaining more observations in the analysis) and to reduce chances of bias. Multiple imputation, assuming the model was correctly specified, is unbiased given an assumption that data were missing at random, conditional on other variables accounted for in the model (i.e. the MAR assumption). This is a more defensible position than the assumption that data were missing completely at random (MCAR) required for valid estimates using case-wise deletion. However, we cannot rule out the possibility that missing data, particularly for ACEs and the CESD were missing not-at-random (MNAR).

The counts for PSS were in the table under “Socioeconomic Factors”. They are now separated out.

9) Table 2 does not show any significant interaction between ACE and social support. In the appendix, there are some significant ones, but if I interpreted the data right, all of them are due to abuse. Additionally, testing the same variables in various ways (any ACE, total score and sub scores) requires a Bonferroni correction.

As noted above, we agree with this conclusion and have edited the text accordingly.

Regarding the multiple testing issue, we would argue that we aren’t testing, but are estimating parameters. In a clinical trial context where a regulator will ultimately be forced to make a black or white decision regarding the approval of a new treatment, it would be critically important to account for multiple endpoints/tests (though Bonferroni tends to be quite severe even in these cases, and these days is often replaced by control of the false discovery rate). Applying a Bonferroni or similar adjustment also ignores the fact that the multiple outcomes in this case are strongly correlated (by definition). Further, if we were choosing to report some subset of parameters based on their p-values, then accounting for the multiplicity would be important. However, we aren’t doing either of these things, but are instead reporting the full sets of model estimates, the corresponding 95%CIs or exact p-values, and leaving it to the reader to make their own judgements regarding the strength of the evidence. For what it is worth, this is an approach promoted by the major epidemiological journals (especially Epidemiology and the IJE), and is consistent with the recent statement on p-values issued by the American Statistical Association (Wasserstein RL and Lazar NA, 2016). We have added the following text to the methods to help clarify our position:

Page 12, paragraph 2.

All parameter estimates are reported with 95% confidence intervals and/or exact p-values. While we have estimated a fairly large number of parameters, we have not selectively reported any of these, nor made any other decisions based on statistical significance testing. This is consistent with current practice in major epidemiological journals, particularly with observational study designs, and recent guidance from the American Statistical Association (Wasserstein RL and Lazar NA, 2016).

10) Discussion: For me the most interesting result is that there are (almost) no buffering effects of social support for neglect and household dysfunction. I ask myself why? One possible explanation would be that the Oslo social support scale does not capture such an effect, but when I look onto the items I would not think so. Another possible explanation would be that the various childhood adversities are different to cope with. This would mean that we should stronger focus on the non-abuse aspects, in particular on neglect. Maybe such thoughts would enrich the discussion. Finally, for me it would be interesting to see if the various aspects of abuse (emotional, physical or sexual) are similarly moderated by social support or if there are differences, too.

As noted above, we have included the interactions between PSS and the different subtypes and specific ACEs. It is of course a challenge to interpret these interactions, especially in relation to each other, given the small cell sizes and subsequent volatility of the estimates. We have added text to this

effect and recommended that this could be looked at in larger studies.

Page 21, last paragraph:

We have reported models for each of ten individual ACEs. However, given the relatively small number of participants experiencing any one specific ACE, the respective parameter estimates will be volatile. While those results qualitatively conformed with the models for any ACE and ACE subtypes, a larger study would be needed to further examine the impact of the individual ACEs.

Reviewer: 2

Reviewer Name
Soonhee Roh

Institution and Country
University of South Dakota
USA

Please state any competing interests or state 'None declared':
None declared

Please leave your comments for the authors below
Title: Adverse Childhood Experiences (ACEs) and Later-life Depression: Perceived Social Support as a Potential Protective Factor

The paper deals with an important topic. Overall it is a well-written, interesting, and timely appropriate paper. The purpose of the paper is clear. I only have the following three comments:

1. This article uses the stress sensitisation theory that suggests that childhood adversity reduces an individual's threshold for developing depressive reactions towards stressful events, causing one to have depressive reactions towards current mild stressors or greater reactivity towards severe stressful events. Introduction was very good and need more strong literature review focused on older adults rather than young women example.

Thank you, we have added the following to the introduction, page 5, paragraph 1:

The relationship persists into older adulthood; data from the Health and Retirement Study, a U.S. population-based study of adults age 50+, showed that in accordance with the stress sensitization theory, childhood trauma (especially physical abuse) amplifies the effect of stresses in adulthood on depressive symptoms (Arpowang, 2016).

2) Need more strong discussions about the significance of the study.

Thank you, we have added the following to the discussion, page 23, paragraph 2:

There is an increasing literature showing the positive effect of interventions that increase perceptions of social support in patients with terminal disease or end stage kidney disease (Cohen 2004). By showing the buffering effect of PSS on the ACE- mental health relationship, our findings highlight the potential for interventions targeting PSS to reduce the likelihood of depression in patients who experienced childhood adversity.

3) Need to elaborate implications of practices based on the findings much more.

Based on the comments from reviewer 1, we have hopefully strengthened both the significance of the study and clarified its implications for practice.

VERSION 2 – REVIEW

REVIEWER	Jochen Hardt Medizinische Psychologie und Medizinische Soziologie, Klinik für Psychosomatische Medizin und Psychotherapie, Universitätsmedizin der Johannes Gutenberg-Universität Mainz, Germany
REVIEW RETURNED	27-Mar-2017

GENERAL COMMENTS	well done, authors! No further objections. There is just one minor comment on an opinion you expressed. I leave it to you to modify it or not. I'd think that the Bonferroni correction applies in cases when parameters are estimated with p values - as done here. The point is that Bonferroni correction itself is too strongly rejecting effects, letting the beta-errors rise. Others, like Holm's modification of the method have different problems. I'd recommend to simply say that no corrections were performed