PEER REVIEW HISTORY

BMJ Open publishes all reviews undertaken for accepted manuscripts. Reviewers are asked to complete a checklist review form (**http://bmjopen.bmj.com/site/about/resources/checklist.pdf**) and are provided with free text boxes to elaborate on their assessment. These free text comments are reproduced below.

ARTICLE DETAILS

TITLE (PROVISIONAL)	Education, job position, income or multidimensional indices? Associations between different socioeconomic status indicators and chronic low back pain in a German sample: a longitudinal field study
AUTHORS	Fliesser, Michael; De Witt Huberts, Jessie; Wippert, Pia-Maria

VERSION 1 – REVIEW

REVIEWER	Patrick Präg University of Oxford, UK
REVIEW RETURNED	26-Oct-2017

| GENERAL COMMENTS | This manuscript has two main problems, A) the treatment of the SES indicators in the analyses and B) the use of causal language despite being an observational study. Both should be easy to fix, yet require a major revision of the manuscript.

1) I don't think it's justified to treat the categorical SES variables as being linearly related with the outcomes. I don't see how the explained variance or the beta coefficients can be interpreted in any meaningful fashion when the SES indicators are treated like this. A more meaningful approach would be to enter the SES variables as sets of dummy indicators into the regression equation and to explore potential non-linearities.

2) The SES indicators should be described in much greater detail; as it stands, the section raises more questions than it answers. ISCO-08 does not have just nine categories, and I think with ISCO it's particularly questionable to treat them as being ordered. It seems that the Authors are using ISCED 1997, yet provide a reference to ISCED 2011. In the text, the Authors mention that education ranges from 0 to 5, in Table 1 they report it ranges from 1 to 5. How were the brackets for income chosen? Is it household income and if yes, is it equivalized (and how)? WS index: How can you assign seven values for education if you only have measured six (or five)? How are the nine job positions identified classified into seven values? Why does the WS index take on non-integer values (Table 1)?

3) I think the results could be more interesting if stratified by sex. I would assume that there is some hidden heterogeneity to be found, maybe even with respect to income.

4) As it is an observational study, I'd remove the causal language—starting with the word "influence" in the title. I would also expect an explanation of the endogenous relationships between the SES dimensions and low back pain and how this endogeneity would |
|---|---|

affect the results shown. Further, the relevance of mere prediction should be spelled out in the discussion or introduction section.

5) I don't understand the study design—what is the benefit of the longitudinal design and why is it relevant to control for baseline values?

6) I would ensure the manuscript is proofread by a native speaker—e.g. "sites" in line 11, "major" in line 28.

7) Results in the abstract are mildly confusing because job position seems to be very different from the other coefficients, as its sign is reversed.

8) Paragraph from line 41 to 58: I don't understand why the Authors consider working conditions to be psychosocial rather than material.

9) I find it confusing how the exlusion criteria in lines 64–65 are different from those stated in the abstract. It would also be interesting to know how many participants were excluded and for which reason.

10) I did not find the theoretical framework advertised in line 54 in the manuscript.

11) It's probably important to discuss the German setting for the study, as Germans' problems with lower back pain might be quite unique. See as a start: Raspe, Heiner, Angelika Hueppe, and Hannelore Neuhauser. 2008. "Back Pain, a Communicable Disease?" International Journal of Epidemiology 37(1):69-74. doi: 10.1093/ije/dym220

12) The Authors might want to consider to mention the potential of alternative approaches of gauging SES in the discussion of their findings. E.g. Baćak, Valerio. 2017. "Measuring Inequalities in Health from Survey Data Using Self-Assessed Social Class." Journal of Public Health. doi: 10.1093/pubmed/fdx036 or Präg, Patrick, Melinda Mills, and Rafael Wittek. 2016. "Subjective Socioeconomic Status and Health in Cross-National Comparison." Social Science and Medicine 149:84-92. doi: 10.1016/j.socscimed.2015.11.044.

REVIEWER	Siegfried Geyer Medizinische Soziologie/ Medical Sociology Unit Medizinische Hochschule Hannover/ Hannover Medical School Carl-Neuberg-Strasse 1 30625 Hannover Germany
REVIEW RETURNED	10-Nov-2017

GENERAL COMMENTS	Comment on Education, job position, income or multidimensional indices? The influence of different indicators of socioeconomic status on low back pain in a German sample The paper examined the relative strengths of effects of SES-indicators and an index measure on low back pain. It contributes to the literature on differential effects of indicators of socio-economic position. This is an interesting paper that nevertheless needs revision. They should do this along the following lines:

Abstract: Under "Results" it should be made clear that the index and SES-indicators are considered in comparison.

Reference 1 refers to another study on the topic of differential effects of SES-indicators on health-/ disease-related outcomes. Although Condiff et al.'s paper is an interesting recent study, it is more appropriate to cite a paper that gives an overview of the topic. Although not new, I suggest to replace it by one of the more comprehensive papers e.g. those by Braveman in Annual Review 2006 or in JAMA 2005.

P. 5: At the end of the intro you should formulate a clear research question or the direction of the study. Do the authors follow Hradil's conclusion as summarized at the end of the introduction, or are they formulating a generic hypothesis? From the text I cannot sort out where they were starting from, i.e. what should be examined. Based on the studies cited in the introduction, can it be assumed that occupational position might be assumed to have the strongest effects with LBP. This might be hypothesized because the structure of work is the least remote factor with potential effects on LBP?

P. 5: Under "Method" the diagnoses of the patients should be stated.

P. 6, lines 32f.: The German Working Group has not recommended the revised SES-index as described in ref. 23, but the precursor by Winkler et al. As this German index should not be known at international level, a remark on similar measures available in English should be of help.

P. 7: How I understood it, the authors have used multivariate OLS regression by including the index and the three SES- indicators simultaneously. From model 1 to the next one always another indicator was added. Is that correct? If so, tables 2 and 3 should display all models instead of only the last one with all variables. The authors should also present a summarizing correlation matrix for the three indicators and the index. I am not sure whether I understood correctly what had been done. The section "Statistical analysis" should be more detailed in order to clarify what had been done in the analyses.

I also have objections against the scaling of the SES- indicators: It is not appropriate to use education and occupation as continuous variables. The situation is again aggravated by income actually being scaled continuously, but in the survey it was divided into categories and re-used as continuous in linear regression. From a statistical point of view this is not appropriate. The authors should do their analyses anew by leaving the indicators (independent variables) as categories. Unfortunately I do not know whether SPSS is permitting this solution without problems. If not, the authors should construct dummy variables and repeat the analyses. From table 1 I cannot see if sample size is sufficient, but try it. Then the conclusions might eventually look a bit different.

P. 9, lines 27-29, discussion section: No associations had been hypothesized, and this is still to be done in the introduction (see the comment above)

P. 10: The problems with SES- indices are not arising out of effect sizes. Indices are affected by a theoretical and empirical deficit: In the case of effects having emerged as well as in their absence it is

	not clear what may account for the results obtained. They just cannot be interpreted properly. If there are effects, it is not clear which indicator is to be held accountable. If no effects are found for an index, it may have occurred that effects of a single indicator may be masked by counterbalanced effects caused by the other indicator(s). At the end of the discussion section the authors should also discuss the limitations of their database and their study Taken together this is an interesting paper with problems in the analyses that can nevertheless be resolved. Revisions are necessary, and before resubmission the authors should also check the language of the paper. This refers to the wording and (to a lesser extent) the grammar of the paper.

VERSION 1 – AUTHOR RESPONSE

Editor comment:
E:- Please include the study design in the title. This is the preferred format of the journal.

A: The study design is now included. The new title is: "Education, job position, income or multidimensional indices? Associations between different socioeconomic status indicators and intermitted low back pain in a German sample: a longitudinal field study"

E: - Please complete and include a STROBE checklist, ensuring that all points are included and state the page numbers where each item can be found. The checklist can be downloaded from here: http://www.strobe-statement.org/?id=available-checklists

A: A strobe statement is uploaded with the revision.

E: Please discuss the limitations of the study in the discussion section.

A: Limitations are now discussed (see Lines 236 to 245).

Reviewer(s)' Comments to Author:

Reviewer: 1
Reviewer Name: Patrick Präg
Institution and Country: University of Oxford, UK
Please state any competing interests or state 'None declared': None declared

Please leave your comments for the authors below

This manuscript has two main problems, A) the treatment of the SES indicators in the analyses and B) the use of causal language despite being an observational study. Both should be easy to fix, yet require a major revision of the manuscript.

R1: 1) I don't think it's justified to treat the categorical SES variables as being linearly related with the outcomes. I don't see how the explained variance or the beta coefficients can be interpreted in any meaningful fashion when the SES indicators are treated like this. A more meaningful approach would

be to enter the SES variables as sets of dummy indicators into the regression equation and to explore potential non-linearities.

A: We agree, that it improves the manuscript when treating the categorical SES variables (education, job position and income since it is grouped) as dummies. We therefore recalculated the models and rewrote the results and the discussion section.

R1: 2) The SES indicators should be described in much greater detail; as it stands, the section raises more questions than it answers. ISCO-08 does not have just nine categories, and I think with ISCO it's particularly questionable to treat them as being ordered. It seems that the Authors are using ISCED 1997, yet provide a reference to ISCED 2011. In the text, the Authors mention that education ranges from 0 to 5, in Table 1 they report it ranges from 1 to 5. How were the brackets for income chosen? Is it household income and if yes, is it equivalized (and how)? WS index: How can you assign seven values for education if you only have measured six (or five)? How are the nine job positions identified classified into seven values? Why does the WS index take on non-integer values (Table 1)?

A1: ISCO-08 has ten major groups, which we used, but since no armed forces (major group 0) were included, this resulted in nine groups. We agree that it is not useful to discuss this so early in the paper. We added the following sentences (Lines 84 to 86) (changes in red):
Job position was measured using the ten major categories from the International Standard Classification of Occupation (ISCO-08), combining jobs according to main tasks, skill level and specialisation [23].

A2: We used ISCED 2011 as described in UNESCO, 2012, 'International Standard Classification of Education ISCED 2011' and grouped tertiary education (because we did not differentiate tertiary education further in the survey). This scale ranges from 0 (only early childhood education) to 5 (tertiary education). The range of 1 to 5 in the table results from the fact that no people with only early childhood education participated in the study. Now we changed this table anyways, which should better explain what was done (Line 140).

A3: Brackets for income were chosen based on 'demographic standards' by the German Federal Statistical Office (DESTATIS), whereby the lowest and highest answer options were grouped. To make that clear we added the following sentence (Lines 87 to 89):
"Monthly net personal income was surveyed using 7 categories (less than 1,250€, 1,250€ to 1,749€, 1,750€ to 2,249€, 2,250€ to 2,999€, 3,000€ to 3,999€, 4,000€ to 4,999€, 5,000€ and more) based on the recommendations of the German Federal Statistical Institute and grouped at the extremes [24]"
Income was not household, but personal income. To make that clear we added the following (Line 87): "Monthly net personal income"

A4: We did not derive the WS-index education variable from ISCED, but from general and job specific qualification (which we also asked in the survey). To make that clear we added the following (Lines 90 to 96):
"This index, a revised version of the German Working Group for Social Epidemiology recommendations [25], is based on three dimensions: education (generated as combination of general and job specific educational level), job position and income. The composition is similar to those of international additive indices. For example Hollighead's "Index of Social Status" [26], where participants are scored between 1 and 7 for each of the single indicators."

A5: Non-integer numbers are taken on because the values, which are assigned for education, job position and income in the WS-index, are not necessarily integer numbers. Lower secondary education with no job specific educational degree for example scores a value of 2.8.

R1 3) I think the results could be more interesting if stratified by sex. I would assume that there is some hidden heterogeneity to be found, maybe even with respect to income.

A: We agree, that it is always plausible that SES interacts with sex and therefore it is reasonable to take sex into account. We also did calculate our models separate for men and women. The results for men remained very similar to the results of the complete group and we saw no effects for women. We believed doing so would only double the number of tables making the manuscript more confusing for little benefit. Also, because sex differences were not the focus of this manuscript (and would also require additions to the theory section), we decided to not report the separate models, but instead to control for it, adding sex as a control variable in the models.

R1: 4) As it is an observational study, I'd remove the causal language—starting with the word "influence" in the title. I would also expect an explanation of the endogenous relationships between the SES dimensions and low back pain and how this endogeneity would affect the results shown. Further, the relevance of mere prediction should be spelled out in the discussion or introduction section.

A 1: The study design (with the evaluation of SES on first measurement point and back pain at the second one was chosen to allow predicitve conclusions. We agree causal language should be used with great care and therefore have avoided it in the new version of the manuscript.

A2: Regarding the remark concerning endogeneity, we see no direct association between SES dimensions and LBP and assume there are mediating variables, as suggested in the manuscript. We are not sure, if we understood your comment about the endogenous relationship completely.

A3: The prediction, in our opinion, allows for the estimation of which SES dimension is most strongly connected with LBP and therefore should be the preferred target for prevention and intervention work. We added the following sentences in the discussion section (lines 247 to 249): "Nonetheless, this study showed especially job position as an important dimension concerning SES influence on intermittent LBP. Further research aiming to prevent and reduce intermittent LBP should therefore focus on conditions that may be influenced by job positions.
"

R1: 5) I don't understand the study design—what is the benefit of the longitudinal design and why is it relevant to control for baseline values?

A: The study design with two measurement points was chosen to allow predictive statements and to investigate new onsets of pain (because SES is investigated at the first measurement point and back pain at the second). For this reason, we also excluded all people having (strong) LBP at baseline. Age and sex at baseline was controlled for, as they have known associations with LBP. Additionally, we added the following sentence in the methods section (Lines 117 to 124):
"After descriptive statistics, four separate hierarchical regression analyses were conducted for each pain outcome using either education, job position, monthly personal net income or the multidimensional index as the respective predictors, while controlling for age and sex, two variables known for their high predictive value in the development of back pain [28]. Education, job position and income thereby were treated as dummy variables to reflect the categorical character of these variables. As most studies only use single indicators to represent SES, a separate model for each indicator was used here as this allowed for demonstration of how much the association between the different indicators and LBP may vary if the other indicators were not taken into account."

R1: 6) I would ensure the manuscript is proofread by a native speaker—e.g. "sites" in line 11, "major" in line 28.

A: The manuscript was carefully proofread by a native speaker. Changes were made all over the manuscript (indicated in red).

R1: 7) Results in the abstract are mildly confusing because job position seems to be very different from the other coefficients, as its sign is reversed.

A: Since the results changed, we have also changed the abstract. It now reads:
"Hierarchical linear regression of four different models for each dependent variable revealed job position to be the best predictor for LBP intensity followed by the multidimensional index, whereas education and income had no impact. For LBP disability, again job position proved the best predictor followed by the multidimensional index and education, while income had no significant association.

R1 8) Paragraph from line 41 to 58: I don't understand why the Authors consider working conditions to be psychosocial rather than material.

A: The terminus here is misleading. What we (and Chou and colleagues, who we cited) mean by working conditions working conditions such as dissatisfaction with the job or compensations for sick leaves. We changed the sentence as follow (Line 46):
"…. chronic LBP is strongly associated with social and psychological factors such as depressive symptoms [16], stress [17] and dissatisfaction with work organisation working conditions [18]."

R1 9) I find it confusing how the exlusion criteria in lines 64–65 are different from those stated in the abstract. It would also be interesting to know how many participants were excluded and for which reason.

A: Exclusion criteria in the abstract are ordered different than in the methods section. We changed that, so that now the order is the same. Reasons for exclusion are now included (Lines 63 to 71):
"This led to a primary sample size of N = 1071 participants. To ensure homogeneity and avoid bias, only participants were included, who at the time of study had an actual job position and answered all relevant questions needed for SES operationalisation. This reduced the sample size to N=654. Furthermore, because prediction of LBP was the focus of this study, participants already reporting serious chronic pain syndromes at baseline were excluded. This led to a sample of N = 367 participants. Finally, since numbers in the groups of primary educated and lower secondary educated people, agricultural workers, machine operators and elementary occupations was very small (under 10 people in each group), these groups were excluded from further analysis, leading to a final sample size of N=352. "

R1: 10) I did not find the theoretical framework advertised in line 54 in the manuscript.

A: We did not advertise a theoretical framework here, but we see that our formulations are misleading. We therefore tried to simplify this (Lines 49 to 56):
"Based on these findings, it can be assumed that SES indicators most closely associated with social and psychological factors and health behaviour exert stronger influence on LBP. Hradil et al., examining the influence of SES indicators on cardiovascular diseases, assumed job position to be most strongly connected with social and psychological factors, education with health behaviour and income with material factors [5]. Using these assumptions, we hypothesise that from the single indicators job position will be strongly associated with LBP, followed by education. Income, we believe, will have the weakest, if any, association. The multidimensional index, covering all possible pathways, should however yield the greatest predictive power."

R1: 11) It's probably important to discuss the German setting for the study, as Germans' problems with lower back pain might be quite unique. See as a start: Raspe, Heiner, Angelika Hueppe, and Hannelore Neuhauser. 2008. "Back Pain, a Communicable Disease?" International Journal of Epidemiology 37(1):69-74. doi: 10.1093/ije/dym220

A: We don't see Germany to be unique with regards to back pain (at least now that the differences between the former East and West Germany has dissipated). See for example, Breivik and colleagues (2006, Survey of chronic pain in Europe: Prevalence, impact on daily life, and treatment, European Journal of Pain, 10, 287-333), who show that Germany is, compared to other European countries, quite average in nearly every dimension associated with (chronic back) pain. We agree, nevertheless, that the setting may be important. For that reason we included the following statement in the limitations (Line XX to XX):
"Furthermore, the study was conducted in four study centres in different federal states of Germany. Although Germany has comparable numbers to other European countries regarding prevalence and severity rates of pain [29], country specific differences may lead to other results in other countries.
"
R1 12) The Authors might want to consider to mention the potential of alternative approaches of gauging SES in the discussion of their findings. E.g. Baćak, Valerio. 2017. "Measuring Inequalities in Health from Survey Data Using Self-Assessed Social Class." Journal of Public Health. doi: 10.1093/pubmed/fdx036 or Präg, Patrick, Melinda Mills, and Rafael Wittek. 2016. "Subjective Socioeconomic Status and Health in Cross-National Comparison." Social Science and Medicine 149:84-92. doi: 10.1016/j.socscimed.2015.11.044.

A: We agree and added the following paragraph in the limitations section (lines 244 to 247):
"Additionally, not all approaches of gauging SES were taken into account. Newer approaches, for example self-assessed socioeconomic status [30, 31] or neighbourhood indicators [3], could improve future studies, but as of yet, no statement can be made as to whether these approaches would have influenced results.

Reviewer: 2
Reviewer Name: Siegfried Geyer
Institution and Country: Medizinische Soziologie/ Medical Sociology Unit, Medizinische Hochschule Hannover/ Hannover Medical School, Carl-Neuberg-Strasse 1, 30625 Hannover, Germany
Please state any competing interests or state 'None declared': There are no competing interests/ None declared

Please leave your comments for the authors below
Comment on Education, job position, income or multidimensional indices? The influence of different indicators of socioeconomic status on low back pain in a German sample

The paper examined the relative strengths of effects of SES- indicators and an index measure on low back pain. It contributes to the literature on differential effects of indicators of socio-economic position. This is an interesting paper that nevertheless needs revision. They should do this along the following lines:

R2: Abstract: Under "Results" it should be made clear that the index and SES-indicators are considered in comparison.

A: Since results has been new calculated, we revised the abstract. It now reads:
"Hierarchical linear regression of four different models for each dependent variable revealed job position to be the best predictor for LBP intensity followed by the multidimensional index, whereas

8

education and income had no impact. For LBP disability, again job position proved the best predictor followed by the multidimensional index and education, while income had no significant association.

R2: Reference 1 refers to another study on the topic of differential effects of SES-indicators on health- / disease-related outcomes. Although Condiff et al.'s paper is an interesting recent study, it is more appropriate to cite a paper that gives an overview of the topic. Although not new, I suggest to replace it by one of the more comprehensive papers e.g. those by Braveman in Annual Review 2006 or in JAMA 2005.

A: We agree and changed the reference to the suggested paper of Braveman in JAMA, 2005 (Line 3).

R2: At the end of the intro you should formulate a clear research question or the direction of the study. Do the authors follow Hradil's conclusion as summarized at the end of the introduction, or are they formulating a generic hypothesis? From the text I cannot sort out where they were starting from, i.e. what should be examined. Based on the studies cited in the introduction, can it be assumed that occupational position might be assumed to have the strongest effects with LBP. This might be hypothesized because the structure of work is the least remote factor with potential effects on LBP?

A: We tried to formulate hypothesis at the end of the introduction section, but obviously were not very successful. Therefore we tried to rephrase the paragraph (Lines 53 to 56):
Using these assumptions, we hypothesise that from the single indicators job position will be strongly associated with LBP, followed by education. Income, we believe, will have the weakest, if any, association. The multidimensional index, covering all possible pathways, should however yield the greatest predictive power.

R2: P. 5: Under "Method" the diagnoses of the patients should be stated.

A: Since we excluded people with severe pain (see Methods, exclusion criteria), there are no people with a diagnosis in the sample.

R2: P. 6, lines 32f.: The German Working Group has not recommended the revised SES-index as described in ref. 23, but the precursor by Winkler et al. As this German index should not be known at international level, a remark on similar measures available in English should be of help.

A: We agree and changed the sentence (Lines 90 to 96):
"This index, a revised version of the German Working Group for Social Epidemiology recommendations [25], is based on three dimensions: education (generated as combination of general and job specific educational level), job position and income. The composition is similar to those of international additive indices. For example Hollighead's "Index of Social Status" [26]. Participants are scored between 1 and 7 for each of the single indicators. The total of these three values then determines the participant's WS-index score [25].."

R2: P. 7: How I understood it, the authors have used multivariate OLS regression by including the index and the three SES- indicators simultaneously. From model 1 to the next one always another indicator was added. Is that correct? If so, tables 2 and 3 should display all models instead of only the last one with all variables. The authors should also present a summarizing correlation matrix for the three indicators and the index. I am not sure whether I understood correctly what had been done. The section "Statistical analysis" should be more detailed in order to clarify what had been done in the analyses.

A: We calculated not one OLS regression, but four for each dependent variable. The reason why we decided to use separate regression models, was because this allows us to demonstrate how widely

the predicted associations between SES and low back pain may vary if researchers only use one SES indicator and do not take into account other SES indicators as is often the case in practice (Fliesser et al., 2016). With this approach, we hope to further increase the awareness of the importance when selecting SES indicators in research, especially if the intention is to use only one indicator. Furthermore, as we now treated education, job position and income as dummies, a model with all these variables included, the number of predictors might be too many for the sample size. To make that clearer, we changed the manuscript as followed (Lines 117 to Lines 125):
"After descriptive statistics, four separate hierarchical regression analyses were conducted for each pain outcome using either education, job position, monthly personal net income or the multidimensional index as the respective predictors, while controlling for age and sex, two variables known for their high predictive value in the development of back pain [28]. Education, job position and income thereby were treated as dummy variables to reflect the categorical character of these variables. As most studies only use single indicators to represent SES, a separate model for each indicator was used here as this allowed for demonstration of how much the association between the different indicators and LBP may vary if the other indicators were not taken into account."

R2: I also have objections against the scaling of the SES- indicators: It is not appropriate to use education and occupation as continuous variables. The situation is again aggravated by income actually being scaled continuously, but in the survey it was divided into categories and re-used as continuous in linear regression. From a statistical point of view this is not appropriate. The authors should do their analyses anew by leaving the indicators (independent variables) as categories. Unfortunately I do not know whether SPSS is permitting this solution without problems. If not, the authors should construct dummy variables and repeat the analyses. From table 1 I cannot see if sample size is sufficient, but try it. Then the conclusions might eventually look a bit different.

A: We agree and recalculated the models as you suggested (see especially results section).

R2: P. 9, lines 27-29, discussion section: No associations had been hypothesized, and this is still to be done in the introduction (see the comment above)

A: We have tried to make our hypothesis clearer. See answer to comment above.

R2: P. 10: The problems with SES- indices are not arising out of effect sizes. Indices are affected by a theoretical and empirical deficit: In the case of effects having emerged as well as in their absence it is not clear what may account for the results obtained. They just cannot be interpreted properly. If there are effects, it is not clear which indicator is to be held accountable. If no effects are found for an index, it may have occurred that effects of a single indicator may be masked by counterbalanced effects caused by the other indicator(s).

A: We fully agree (and tried to point out this problem in the discussion section). We added the following sentence (Lines 211 to 213):
"The overlapping of indicators and mediating factors could be the problem here, leading to a shared influence on back pain for all indicators, or the effect of one indicator being counterbalanced by another."

R2: At the end of the discussion section the authors should also discuss the limitations of their database and their study

A: We agree, please see Lines 238 to 247:
The results presented in this paper are afflicted by some limitations, namely the small group sizes of primary and lower secondary educated people, agricultural workers, machine operators and elementary occupations, leading to the exclusion of these groups. A repetition of the research design

with a larger sample would make the results more reliable. Furthermore, the study was conducted in four study centres in different federal states of Germany. Although Germany has comparable numbers to other European countries regarding prevalence and severity rates of pain [29], country specific differences may lead to other results in other countries. Additionally, not all approaches of gauging SES were taken into account. Newer approaches, for example self-assessed socioeconomic status [30, 31] or neighbourhood indicators [3], could improve future studies, but as of yet, no statement can be made as to whether these approaches would have influenced results.

R2: Taken together this is an interesting paper with problems in the analyses that can nevertheless be resolved. Revisions are necessary, and before resubmission the authors should also check the language of the paper. This refers to the wording and (to a lesser extent) the grammar of the paper.

A: The revised version of the paper was carefully proofread by an English native speaker.

VERSION 2 – REVIEW

REVIEWER	Patrick Präg University of Oxford, UK
REVIEW RETURNED	19-Dec-2017

| GENERAL COMMENTS | A) I think there are still major language problems that make it impossible to publish the manuscript in its current form. Both Reviewers admonished the poor language in the first round of reviews already. A few examples from the first two pages of the revised manuscript: What is "strong, but different associations" meant to mean (page 2, line 30)? Or "Results could be limited" (p. 2, l. 45)--limited in what way? On page 3, lines 7-9, the Authors state: "The justification for using a specific SES indicator is often not adequately justified in articles [3], and they are even used interchangeably [4]." Here, the first part of the sentence is tautological, and the "they" in the second part should be further specified. Are the Authors sure it is called "intermitted" low back pain (title)? A Google search leads to virtually no prior usage of that term.

B) Next to the language problem, there are a number of points from my previous review that have not been addressed:

B1) I asked to remove the unwarranted causal language from the manuscript in the previous round. While the Authors made some changes to the manuscript to that effect, there are still instances of unwarranted causal language use in the manuscript, e.g. p. 2, l. 25: "education and income had no impact."

B2) I stated that it seems that the Authors are using ISCED 1997 for classifying education, yet provide a reference to ISCED 2011. This has not been changed.

B3) The Authors now state that the income they use is personal rather than household income. This is an unusual choice, as household income is a more salient predictor, which should be acknowledged in the limitations. Further, household income is a more important predictor of women's life situation than it is for men's, an important heterogeneity that is hidden in the analyses pooling women and men.

B4) This ties in with my comment from the last round review: "I think |

the results could be more interesting if stratified by sex. I would assume that there is some hidden heterogeneity to be found, maybe even with respect to income." which was left unaddressed.

B5) Regarding the WS index, I asked: How can you assign seven values for education if you only have measured six (or five)? How are the nine job positions identified classified into seven values? These questions were not addressed in the revised manuscript, and have become more important now that we learn that many groups of lower socioeconomic status are omitted from the analyses.

C) New points that arise in the revised manuscript:

C1) In the analyses reported in Tables 3 and 4, I would refrain from reporting the beta values of dummy predictors. For dummy variables, the standard deviation is a function of the mean p, namely the square root of $p * (1 - p)$, making the betas, which are standardized by the standard deviation, difficult to interpret.

C2) In Tables 3 and 4, I would suggest to report the reference categories in the third column.

C3) The Authors correctly write in the limitations section: "The results presented in this paper are afflicted by some limitations, namely the small group sizes of primary and lower secondary educated people, agricultural workers, machine operators and elementary occupations, leading to the exclusion of these groups." They follow with "A repetition of the research design with a larger sample would make the results more reliable." I would like to point out that reliability is usually not an issue with small samples. The issue the Authors should be acknowledging is that they have a biased sample, which leads to biased results. I would speculate that their findings underestimate the predictive value of education and occupation.

C4) Relatedly, the bottom category of income is rather broad for net personal income, stretching from the poor into the middle class, which will also lead to an artificially low income coefficient.

C5) As stated above, the use of personal rather than household income should be mentioned as a limitation.

REVIEWER	Siegfried Geyer
	Medical Sociology Unit
	Hannover Medical School
	Carl-Neuberg-Strasse 1
	30625 Hannover
	Germany
REVIEW RETURNED	10-Jan-2018

| GENERAL COMMENTS | The authors have adequately responded to the comments, and the paper is substantially improved, and it is presenting some very interesting findings. No more substantive revisions are necessary. Although I am not a native speaker, I found some smaller linguistic mistakes while reading through the text. The authors should go through the text again and amend them. |

Reviewer: 1

Question A) I think there are still major language problems that make it impossible to publish the manuscript in its current form. Both Reviewers admonished the poor language in the first round of reviews already. A few examples from the first two pages of the revised manuscript: What is "strong, but different associations" meant to mean (page 2, line 30)? Or "Results could be limited" (p. 2, l. 45)-- limited in what way? On page 3, lines 7-9, the Authors state: "The justification for using a specific SES indicator is often not adequately justified in articles [3], and they are even used interchangeably [4]." Here, the first part of the sentence is tautological, and the "they" in the second part should be further specified. Are the Authors sure it is called "intermitted" low back pain (title)? A Google search leads to virtually no prior usage of that term.

Answer A: A native speaker already proofread the last version of the manuscript, but obviously wasn't able to correct all mistakes. This time another native speaker proofread the manuscript. Changes were made all over the manuscript. We hope this erased all language problems. The title is also adjusted to use more common terms.

Question B) Next to the language problem, there are a number of points from my previous review that have not been addressed:

Answer B): These points have been addressed, not in the manuscript, but in the response letter we submitted with the revised manuscript. It seems that unfortunately Reviewer 1 didn't receive this response letter.

Question B1) I asked to remove the unwarranted causal language from the manuscript in the previous round. While the Authors made some changes to the manuscript to that effect, there are still instances of unwarranted causal language use in the manuscript, e.g. p. 2, l. 25: "education and income had no impact."

Answer B1) We again revised the manuscript carefully and (hopefully) now changed all sentences with causal language. Changes were made all over the manuscript (indicated in red).

B2) I stated that it seems that the Authors are using ISCED 1997 for classifying education, yet provide a reference to ISCED 2011. This has not been changed.

Answer B2: As stated in the last revision letter we used ISCED 2011 as described in UNESCO, 2012, 'International Standard Classification of Education ISCED 2011' and grouped tertiary education (because we did not differentiate tertiary education further in the survey).

Question B3) The Authors now state that the income they use is personal rather than household income. This is an unusual choice, as household income is a more salient predictor, which should be acknowledged in the limitations. Further, household income is a more important predictor of women's life situation than it is for men's, an important heterogeneity that is hidden in the analyses pooling women and men.

Answer B3: We agree and added the following in the limitations (Lines 231 to 233): "In this study, income was evaluated as personal net income, but perhaps the material situation of some would be better reflected by household income, which may hide some heterogeneity.."

Question B4) This ties in with my comment from the last round review: "I think the results could be more interesting if stratified by sex. I would assume that there is some hidden heterogeneity to be found, maybe even with respect to income." which was left unaddressed.

Answer B4: As already stated in the last revision letter we agree, that it is always plausible that SES interacts with sex and therefore it is reasonable to take sex into account. We also did calculate our models separate for men and women (see tables attached to this response letter). The results for men remained very similar to the results of the complete group and we saw nearly no effects for women (only for the WS-index in the association with pain disability). Therefore we believe distinguishing between men and women would only make the manuscript more complex (by doubling the tables) for little benefit. Also, because sex differences were not the focus of this manuscript (and would also require additions to the theory section), we decided to not report the separate models, but instead to control for it, adding sex as a control variable in the models.

Question B5) Regarding the WS index, I asked: How can you assign seven values for education if you only have measured six (or five)? How are the nine job positions identified classified into seven values? These questions were not addressed in the revised manuscript, and have become more important now that we learn that many groups of lower socioeconomic status are omitted from the analyses.

Answer B5: As already answered in the last revision: We did not derive the WS-index education variable from ISCED, but from general and job specific qualification (which we also asked in the survey). Also job position for the index is a combination of job position and qualification. To make that clear we added the following (Lines 88 to 91):
"This index, a revised version of the German Working Group for Social Epidemiology recommendations [25], is based on three dimensions: education (a combination of general and job specific educational level obtained together with ISCED), job position (a combination of position and qualification) and income.."

C) New points that arise in the revised manuscript:

Question C1) In the analyses reported in Tables 3 and 4, I would refrain from reporting the beta values of dummy predictors. For dummy variables, the standard deviation is a function of the mean p, namely the square root of p * (1 - p), making the betas, which are standardized by the standard deviation, difficult to interpret.

Answer C1: We deleted the column reporting beta values (Lines 152-153 and 169-170).

Question C2) In Tables 3 and 4, I would suggest to report the reference categories in the third column.

Answer C2: We agree and changed the tables accordingly (Lines 152-153 and 169-170).

Question C3) The Authors correctly write in the limitations section: "The results presented in this paper are afflicted by some limitations, namely the small group sizes of primary and lower secondary educated people, agricultural workers, machine operators and elementary occupations, leading to the exclusion of these groups." They follow with "A repetition of the research design with a larger sample would make the results more reliable." I would like to point out that reliability is usually not an issue with small samples. The issue the Authors should be acknowledging is that they have a biased sample, which leads to biased results. I would speculate that their findings underestimate the predictive value of education and occupation.

Answer C3: We agree and amended the following in the limitations (Lines 218 to 223): "The results presented in this paper are afflicted by some limitations, namely the small group sizes of primary and lower secondary educated people, agricultural workers, machine operators and elementary occupations, which led to the exclusion of these groups. This reduces the explanatory power of the results and may produce an underestimation in the predicted association of education and job position. A replication study including more people from lower education and job groups would assure more conclusive and generalizable results"

Question C4) Relatedly, the bottom category of income is rather broad for net personal income, stretching from the poor into the middle class, which will also lead to an artificially low income coefficient.
Answer C4: We agree and added the following in the limitations section (Lines 223-225): "Furthermore, the lowest income bracket was relatively broad. We speculate income may have had a stronger association, had this category been split up.."

Question C5) As stated above, the use of personal rather than household income should be mentioned as a limitation.

Answer C5: We agree. See answer B3.

Reviewer: 2
Reviewer Name: Siegfried Geyer
Institution and Country: Medical Sociology Unit, Hannover Medical School, Carl-Neuberg-Strasse 1, 30625 Hannover, Germany
Please state any competing interests or state 'None declared': None declared

Please leave your comments for the authors below
The authors have adequately responded to the comments, and the paper is substantially improved, and it is presenting some very interesting findings. No more substantive revisions are necessary. Although I am not a native speaker, I found some smaller linguistic mistakes while reading through the text. The authors should go through the text again and amend them.

Answer: A native speaker carefully proofread the manuscript and made necessary changes.

Table 1: Four different hierarchical regression models predicting influence of different operationalisations of SES on CPG pain intensity score (higher values more pain), controlled for age. ΔR^2 indicates model improvement after application of SES indicator: **Women only N=257):**

Model	ΔR²	Variable	B	SE B	Beta	p
Model 1: **Education** **Reference group: Tertiary Education**	0.01	Upper Secondary Education	3.31	2.5	0.10	0.19
		Post-secondary non-tertiary education	-1.38	2.8	-0.04	0.62
Model 2 job position reference group:	0.05	Professionals	-3.0	2.8	-0.1	0.30
		Technicians	-2.5	3.0	-0.08	0.40*

Model	ΔR²	Variable	B	SE B	Beta	p
Manager		Clerical Support Workers	2.1	3.8	0.05	0.59
		Service and Sales Workers	6.7	4.1	0.13	0.11
		Craft Workers	13.1	8.9	0.10	0.14
Model 3: Income Reference group: more than 5000€	0.03	Under 1250	-4.5	6.7	-0.11	0.51
		1250-1749	-6.8	6.3	-0.19	0.27
		1750-2249	-5.2	6.3	-0.13	0.40
		2250-2999	-6.6	6.5	-0.14	0.31
		3000-3999	-9.9	6.3	-0.25	0.12
		4000-4999	-0.5	6.9	-0.01	0.94
Model 4: WS-Index	0.01		-0.67	0.39	-0.12	0.09

*p < 0.05; **p < 0.01

Table 2: Four different hierarchical regression models predicting influence of different operationalisations of SES on CPG pain disability score (higher values more pain), controlled for age. ΔR^2 indicates model improvement after application of SES indicator: **Women only N=257):**

Model	ΔR²	Variable	B	SE B	Beta	p
Model 1: **Education** **Reference group: Tertiary Education**	0.01	Upper Secondary Education	2.4	2.2	0.08	0.29
		Post-secondary non-tertiary education	0.09	2.5	0.00	0.97
Model 2 job position reference group: Manager	0.05	Professionals	-3.0	2.5	-0.10	0.25
		Technicians	1.0	2.7	0.04	0.72
		Clerical Support Workers	3.6	3.5	0.09	0.30
		Service and Sales Workers	4.2	3.7	0.09	0.26
		Craft Workers	17.1	8.0	0.15	0.04*

Model 3: Income Reference group: more than 5000€	0.02	Under 1250	0.0	6.0	0.0	1.0
		1250-1749	-2.0	5.6	-0.06	0.72
		1750-2249	-0.6	5.7	-0.02	0.92
		2250-2999	-5.8	5.9	-0.14	0.32
		3000-3999	-4.7	5.6	-0.13	-0.84
		4000-4999	0.90	6.2	0.02	0.88
Model 4: WS-Index	0.02*		-0.73	-0.34	-0.15	0.04*

*$p < 0.05$; **$p < 0.01$