Author's response to reviews

Title: Occupational stressors and its organizational and individual correlates: A nationwide study of Norwegian ambulance personnel

Authors:

Tom TS Sterud (tom.sterud@medisin.uio.no)
Erlend EH Hem (erlend.hem@medisin.uio.no)
Oivind OE Ekeberg (oivind.ekeberg@uus.no)
Bjorn BL Lau (bla@stami.no)

Version: 2 Date: 21 December 2007

Author's response to reviews: see over
Dear Editors,

Manuscript: 1165858761548718
Title: Occupational stressors and its organizational and individual correlates: A nationwide study of Norwegian ambulance personnel

Thank you for your e-mail of 11 December 2007. We hereby submit the revised manuscript.

The manuscript has been revised thoroughly, and in accord with the valuable comments from the reviewers and the editor. The changes are summed up below.

Response to reviewer 1

There are one or two aspects of the paper which require either correction or clarification, either because of the term is unclear or the phrases are epidemiological/statistical in nature and would require explanation for a wider audience.

For example, in the abstract, line 2, is the word “empirical” correct?

Response
The sentence in the abstract line 2 “there is little empirical evidence for this”, has been rewritten into: “there is little research to support this notion”

On page 6, para 2, the phrase “appraisal of stressors” could be reviewed.

Response:
We agree. The phrase is deleted. The formulation is now, p. 4 paragraph 3, lines 4-5: “through modifying the experience of stress severity associated with the stressors”

The term “neuroticism” could I think be replaced with a better term.

Response
We have now described in greater detail the different terminology used when referring to personality traits, p 11 and 12 under “individual characteristic”. An alternative term to neuroticism might be vulnerability, thus we now write on p. 11: “Personality was measured by 27 items from the Basic Character Inventory (BCI), which is based on an original questionnaire constructed by Lazare, Klerman, and Armor [25], and modified by Torgersen [26]. BCI is based on the ‘big three’ personality dimensions. The BCI – vulnerability scale (α = .74) measures the neuroticism dimension (e.g. I’m very touchy about criticism), the BCI – intensity scale (α = .72) assesses extroversion/introversion (e.g. ‘Many people consider me a lively person’), and the BCI – control scale (α = .66)
(e.g. ‘Everything I do must be precise and accurate’) describes the degree of compulsiveness/obsessiveness. Here, the terms neuroticism, extroversion and control were used, respectively. Each dimension is based on nine questions with a dichotomous response (0 = does not apply, 1 = applies), allowing each dimension a range of scores between 0 (low) and 9 (high).” However, we would prefer to keep the terminology which is used in the paper as we believe that the terms neuroticism, extroversion and control are more commonly known and used in the literature.

**On page 10, the phrase “tree factors with eigenvalues” would require explanation to register with the wider potential readership.**

Response

We have added one sentence to explain the concept of eigenvalues, p 10 under “Ambulance specific stressors”, line 11: “the variances extracted by the factors”

**Finally, in the Conclusion, it would be helpful to have a list of recommendations regarding the next steps, both in terms of further work required in this area and how the findings of this study might be used to benefit the personnel involved.**

Response

We have tried to highlight some additional key findings and suggest some possible implications in the discussion section, page 18 paragraph 2, lines 13-17:”Overall, however, there were weak associations between personality and reported stressors frequency and severity. The personality traits control and introversion were not significant in the model. The fact that many of these individual moderators were not significant suggests that work related factors might be stressful in themselves, and may very well be more easily addressed at an organizational level”, and suggestion for future studies in the conclusions section, p. 20: “Future research may learn more about how ambulance personnel deal with potentially stressful incidents by a greater focus on approaches yielding in depth explorations of ambulance personnel facing stressful conditions over time and across occasions, in the context of their aspirations, beliefs and strategies of coping.”
Response to reviewer 2

1. In the sample and procedure sections of the Method, 3 different figures are given regarding sample size. One final sample size is needed. That is, which, if any of the figures 1180, 1005 or 1286 actually represent the valid sample? With a failure to include statistics that contain degrees of freedom, the task to identify the actual sample size is made more difficult. Further, on page 7, are the authors suggesting they deleted (listwise) all cases from the data set that had even 1 missing value?

Response
We agree with the reviewer that the reporting of different sample sizes was a bit confusing. N = 1286 refers to the total number of respondents. Among these N = 1180 reported to be operational ambulance personnel. Finally, n = 1005 refers to the sample used in the analyses. We have now rewritten the procedure and sampling parts into one section in order to present the numbers in a logical order (from 1286 → n = 1180 → n = 1005 (final sample in the regression analyses). The actual sample size (n = 1005) is now also reported in the headings of table 4 and table 5. Moreover, cases were not deleted listwise if they had 1 missing value, but missing values on one variable. This is now stated more clearly, p. 9 paragraph 2, lines 1-2: “After listwise deletion of respondents with missing values on one or more variables”

2. Please include the results of the significant independent samples t-test alluded to in the sample section with relation to age and gender.

Response
We now report the t value p. 9 paragraph 3, line 4: t = 5.96

3. In the procedure section the authors mention differences in demographic factors such as education and work role. Were any tests conducted to address the potential differences on variables of interest due to these categorical distinctions?

Response:
We did not originally include the education variable. However, the effect of education level was tested, this is now reported under “sample”, p. 8 paragraph 1, lines 5-8: “There were no significant differences in stressor level among respondents with education at high school level or lower compared to respondents with education at high school level (Multivariate Analysis of variance (MANOVA): F = 0.8, degrees of freedom = 14, p = 0.64).” Only operational ambulance personnel were included in this study.

4. On page 12 a series of paired sample t-tests are conducted and the reader is directed to Table 3. Table 3 only presents means and SDs for the items and extracted factors. The authors need to include the results of the t-tests. Was any adjustment made for multiple comparisons to ward off the heightened potential for type 1 error?
Response

Table 3 presents Means and SDs, and the sample size is known. Thus, the t-test may be replicated: Formulae: \( t = \frac{(m_E - m_C)}{\sqrt{\frac{(s^2_E + s^2_C)}{n}}} \). In order not to make the table too complex and comprehensive we find it most adequate to refer to p-values >.01 only.

We agree that corrections (e.g., Bonferroni post-hoc test) are important and meaningful when two or more groups are compared with regard to many different outcome variables, because multiple comparisons increase the likelihood of detecting differences by chance. However, comparing stressor A to Stressor B, stressors A to C etc. within the same sample using a paired t-test is a different kind of comparison, and the comparison of stressor A to B, does not influence the likelihood of detecting a difference between stressor A and stressor C within the same sample. Hence, Bonferroni (or similar) adjustments for multiple comparisons are too strict for the type of comparisons made in table 3.

5. Using multiple linear regression for this study is not the most efficient or effective statistic to employ. It appears that the whole regression was conducted and then organisational factors were removed to see what was left. At least, the analyses should include hierarchical regression where groups of variables (e.g., individual differences, organisational) are added in steps to observe any changes in \( r \) squared.

Response:

It is not clear to the authors what the reviewer alludes to when stating that multiple linear regression analyses is not the most effective or efficient statistic to employ, and then suggesting that hierarchical regressions should be used. In our point of view, whether one chooses to enter all variables in one or more blocks, it is all the same still a multiple linear regression analysis. In the present study we were interested in the unique variance explained by each block of variables as additional information to the amount of variance explained by individual variables. We would argue that the chosen approach (i.e., removing or entering groups of variables) is the most appropriate way to estimate the amount of unique explained variance, because this allows us to partial out the amount of shared variance. If, however, the blocks of variables were entered in steps, changes in \( r \) squared for each block of variables (as an expression of unique variance) would depend on the order in which the block were entered. On the other hand, however, if the purpose is to estimate the relative influence of confounding variables on the estimated relationships of interest, then a block-wise enter of the variables is the preferable procedure. However, this question was not addressed in the present article, and would demand a table with three times as many columns as the present table. If we were to present estimates of the three blocks for each of the outcome variables, this would be a very comprehensive figure. Therefore we would prefer to keep the original table.

6. It is essential in my view that the issue of statistically significant results as a function of the sample size is addressed. Effect sizes need to be included, adjustments for multiple comparisons would be a positive addition to the statistics and a discussion about which, if any, relationships are psychologically significant would add to the paper.
Response

We agree that statistically significant results are directly related to sample size. Given the relatively large sample size in the present study we have therefore throughout the text focused upon independent variables that were consistently related to higher level of stressor severity on all dimensions in order not to overstate marginal statistically significant findings. Throughout the text we have taken great care to pinpoint when this was not the case, e.g., p. 13 paragraph 1, lines 4-7: “We did not find any strong relationship between formalized colleague support and lowered stress. However, people who answered ‘No, but would be nice if there was’ reported significantly higher stressor severity on three out of seven stressors”.

We are aware that it is common to report “effect sizes” when reporting results from Univariate analyses of variance (Anova). However, when applying linear regression unstandardized or standardized beta parameters is the usual way of reporting “effect sizes” or parameter estimates. If, however, one wants to know the effect sizes (as they are commonly reported in ANOVA), they are equal to the square of the reported standardized betas in the present article. Hence, the reader could easily calculate “effect sizes” if for some reason that is preferable.

---

Minor Essential Revisions (such as missing labels on figures, or the wrong use of a term, which the author can be trusted to correct)

1. It would be advantageous to label or make explicit that the intersecting figures in Table 2 (correlation matrix) are in fact the alpha coefficients. The inclusion of gender in this table is confusing as not only is it a categorical variable but the table suggests a SD 3 times larger than the mean. I know a mention is made in the notes but removal would be clearer.

Response

This information was reported in the text on p.12 under results, paragraph 1, lines 2-3: “Alpha coefficients are reported in the diagonal for index scores.” This information is now also added under “Note” in Table 2. Regarding gender, 76.8 was not intended to refer to the SD of gender, but to the proportion of men in the sample. This is now made clear by adding “%”, see Table 2 variable 15: “23.2% 76.8%”.

2. Top of page 15 the word ‘ranged’ may best be changed for ‘rated’. Two lines further down a full stop is missing after references 4, 30, 31 and the words ‘meaningful’ and ‘motivating’ are not clear in the context of the discussion.

Response.

We agree. The term ranged has been replaced by rated. A full stop has been inserted. Moreover we have added a few lines to further motivate the use of the words ‘meaningful’ and ‘motivating’, p. 16 under, paragraph 1, lines 2-7: “Ambulance specific operational demands may be an expected part of the job and one of the main reasons why these people chose the ambulance occupation in the first hand. However, although
ambulance personnel may consider ambulance specific operational demands to be the most meaningful and the most motivating stressors, the high severity stressors may nevertheless be risk factors for post traumatic stress symptoms.”

3. On page 24 there are a number of comas instead of full stops between figures (e.g., 0,54 and 0,51 rather than 0.54 and 0.51).

Response
Thank you. The use of commas was not intentional and has been replaced by decimal points.

Discretionary Revisions (which the author can choose to ignore)

1. Identifying a sampling strategy would be beneficial (i.e., criterion based).

Response:
It is not clear to the authors what the reviewer is referring to (i.e. criterion based sampling strategy (?)). As stated in the method section, our aim was to include all eligible ambulance personnel in Norway. This procedure, we believe, is described quite detailed in the methods section.

2. Although not vital, at the end of page 8 when reference to the MANOVA is made, I believe it is customary to provide the F-test details to demonstrate the lack of significance.

Response
We have now provided more detailed information on the F-test, p. 9, paragraph 1, lines 6-7:” (F = 1.3 degrees of freedom = 28.0, p = 0.14).

3. Table 1 is not very clear to read and the issue of complex loadings is not really discussed (e.g., take care of seriously injured and dying patients and uncertainty about what you will meet…)

Response
We agree. In Table 1 the factor loadings of the items considered to constitute the respective factors are now emphasized in bold. In addition we now comment on the complex factor loadings of the two items, p. 11 paragraph 1: “It should be noted that the items “take care of seriously injured and dying patients” and “uncertainty about what you will meet” loaded approximately equally high on two of the factors (i.e. ‘serious operational tasks’ and ‘non-emergency tasks’), however, based on an evaluation of the content of these items, they were included in the serious operational task index (Table 1).
4. At the bottom of page 11 an r value of .32 is presented; please clarify the point of this.

Response
The r values reported referred to Pearson’s correlations, this is now made explicit, p. 12, under results, paragraph 1 lines 4-5:” The mean correlation between frequency and severity indexes was Pearson’s r = .32 for the general occupational stressors, and Pearson’s r = .16 for the ambulance specific stressors.”

5. Age is often confounded with length of service in emergency service populations. Those who find the work has become too much for them tend to self-select out just as this group self-selects in to the service.

Response
Age was highly correlated with working period in the services (Pearson’s r = 0.53, p < .01). Hence, the negligible effect of working period in the services was adjusted for when adjusting for age in the multiple regression analyses. We agree that self-selection is important, but not easily addressed in a cross sectional study, and we did not have the possibility to address this issue in the present study.

6. At the top of page 18 please clarify that this is the largest investigation in Norway.

Response
We have now moderated this sentence to (p. 19 under strengths and limitations): “The strengths of this study are that it is one of the largest investigations of ambulance personnel conducted”. See Sterud et al 2006 for a systematic review of the ambulance literature.