

Author response to reviews of

Manuscript HPO-22-0057

NFC and burnout in teachers - A replication and extension study

Josephine Zerna, Nicole Engelmann, Anja Strobel & Alexander Strobel
submitted to *Health Psychology Open*

RC: Reviewer Comment AR: Author Response ☐ Manuscript text

Dear Dr. McParland,

thank you very much for taking the time to consider our manuscript for publication at *Health Psychology Open*. We would like to thank the reviewers for their efforts to thoughtfully assess our manuscript. In the following we address your and each reviewers' concerns point-by-point. We hope that we could adequately address all the issues raised.

1. Reviewer #1

RC: Difference between the samples of Grass et al. (2018) and this study

- The authors should clearly mark the difference between their teacher sample and Grass et al.'s (2018) student teacher sample. This includes designating Grass et al.'s sample as "student teachers" (on p.3 and p.20), not "teacher trainees". This is more precise (the latter could easily be confounded with 2nd teacher education phase, called "Referendariat") and corresponds to Grass et al.'s own use. Additionally, "replication/replicated" should be qualified as "replication/replicated in a teacher sample" throughout the manuscript. This would also mark the novelty of this paper.

- Furthermore, the authors should consider and discuss implications of the different professional situations of the two samples. This includes considering question like "Are there difference in how personal efficacy is perceived among students compared with in-service teachers?", "What are the implications for burnout, need for cognition and other key concepts/variables with reference to the difference between academic learning and working in a school?" Literature dealing with these questions should be included.

AR: This is our response.

Assignment: JZ

Work in progress

RC: Focus on the burnout characteristic reduced personal efficacy (RPE), cf. Table 3

I don't see the point of focusing only on the significant correlations and mediations from Grass et al. (2018). Non-significant results are also findings. In Grass et al. (2018), they analyzed the other two characteristics emotional exhaustion and depersonalization, too. There, they focused on RPE in further mediation analysis apparently because it was more strongly related to NFC and Self-control. Obviously,

correlations and mediations are different in this study (which might be due to investigating a different population). This justifies analyses of all assessed variables. Therefore, analyses of associations of the two other burnout characteristics (emotional exhaustion and depersonalization) should be included.

Furthermore, RPE can not be the most important burnout characteristic. On the contrary, it frequently has the lowest correlation with burnout (MBI score) among the three subscales. Additionally, it is the subscale that receives less attention in research on teachers' health, in the sense that if only certain burnout subscales are investigated, then those are usually emotional exhaustion and depersonalization (e.g.,

- Bermejo-Toro, L., Prieto-Ursúa, M., & Hernández, V. (2016). Towards a model of teacher well-being: Personal and job resources involved in teacher burnout and engagement. **Educational Psychology, 36**(3), 481–501. <https://doi.org/10.1080/01443410.2015.1005006> - Dicke, T., Parker, P. D., Marsh, H. W., Kunter, M., Schmeck, A., & Leutner, D. (2014). Self-efficacy in classroom management, classroom disturbances, and emotional exhaustion: A moderated mediation analysis of teacher candidates. **Journal of Educational Psychology, 106**(2), 569–583. <https://doi.org/10.1037/a0035504>)

This means that further analyses of the two other subscales would increase relevance of this paper in the scientific field.

When using the term “burnout”, the manuscript should carefully distinguish between burnout characteristic RPE and burnout as the commonly used concept including EE and DP. Mostly, this is accomplished successfully, but not in the discussion (p.19: “we replicated findings of mediators between Need for Cognition and burnout” when only mediations between NFC and RPE were investigated), and conclusion (first two sentences p.25), for example.

AR: This is our response.

This is a section quoted from the revised manuscript to illustrate the change.

Assignment: JZ

Almost there

RC: Abstract: “little is known about the personality traits that promote or protect against burnout”

- This phrase is unfortunate. The readers might doubt whether the authors are aware of, for example, the more than a thousand times cited meta-analyses that exist on this topic and the literature they themselves cite later that elucidates this issue. If promoting and protecting in the narrowed sense of having shown predictive validity in longitudinal studies is meant, then, the presented study does not add insights because it is cross-sectional. Please clarify what was meant.

AR: We acknowledge that this sentence was unfortunate. We changed it accordingly and now write in the abstract:

While there is evidence that certain personality traits promote or protect against burnout, the role of dispositional cognitive motivation has been investigated only recently.

Assignment: AS

Done

RC: p.17 typo: the lavaan function should be spelled modindices(), not modincides()

AR: Thank you for noticing this typo. We have corrected it.

Assignment: **Last Author**

Done

RC: p.19/23 somewhere it should be conceded that the positive relation Covid burden and years of experience is influenced or confounded by item number 2 of the Covid burden questionnaire (“Are you in a Covid-19 risk group?”). Persons that do not belong to the risk group of old people (50+) can by default not have the highest years of experience.

AR: This is our response.

Assignment: **JZ**

Work in progress

RC: p. 21 “The mediator that did not reach significance was the perception of own resources exceeding the job demands. As this latent variable was conceptualized as boredom at work, we could not confirm the positive association of boredom and burnout found by Reijseger et al. (2013.)”

- The authors might want to deepen the discussion with the possible explanation that boredom is less prevalent in the teaching profession (including a suitable reference) than in professions where you can “watch the paint dry” as Reijseger et al. describe. For some readers it might be counterintuitive that teachers experience boredom in the form of resources exceeding demands at all. Teaching appears to be an activity that could always address more individual needs and a class of app. 20 students seems to always have the potential to keep their teachers busy.

AR: This is our response.

Assignment: **JZ**

Work in progress

RC: supplementary material S2

For me, the numbering of supplementary material headings, table and figure numbers was confusing. I would find a system that links supplemental topic and table/figure numbering with each other much more intuitive and easier to navigate, e.g., heading S2, table S2a, table S2b,... or heading S2, table S2.1, table S2.2,...

AR: This is our response.

Assignment: **JZ**

Work in progress

2. Reviewer #2

RC: Thank you so much for providing me with an opportunity to review this manuscript. I enjoyed reading it. In brief, this study replicates the findings of mediation between NFC and burnout, investigates the

effect of different demand resource ratios on the relationship between NFC and burnout, and explores the impact of other variables such as perceived burden by the pandemic.

The introduction is well-written and follows a logical order in explaining the literature and the gap to explore the other variables mediating the relationship between NFC and burnout. The method seems to be clear and concise enough to explain the sample, procedure, and data analysis strategies. However, I was wondering if the authors have checked the reliability of the last scale used to assess the Covid-19 pandemic burden.

The results have also been presented in detail and clearly according to the appropriate methods of analysis and have been properly discussed.

The only weak point of the study is the use of self-reports, which limits taking causal conclusions. However, this point along with others has been presented in the limitation of the study and nothing can be done now. Overall, I found the study interesting with significant results to share and enjoy good quality. Therefore, I suggest acceptance.

AR: We like to thank the reviewer very much for this positive assessment of our manuscript. In our revision, we now provide a reliability estimate (in terms of Cronbach's α) of the scale used to assess Covid-19 burden, see p. XX.

Assignment: JZ

3. Reviewer #3

RC: ... the goals of the study and the theoretical framework it uses switch multiple times as the manuscript progresses. Notably, the different models presented on the same dataset are clearly different (in terms of structure, predictors, and outcome variables) and are not clearly related to each other. So the theoretical focus needs to be improved to increase readability and clarity. Specifically, the different models need to be related to other in a theoretical and empirical manner. Also, the main model in figure 2 did not fit well possibly because it assumed conditional independence between three main DTH, DTL, and DRF factors that are expected to covary even after controlling for NFC.

AR: This is our response.

Assignment: JZ

Work in progress

RC: Page 1.

The abstract needs to be rewritten as it does not clearly present the variables of interest and their relations ("we analysed..." fails to delineate the outcome variabel, predictor, mediator). Also the Grass et al. study should be fully references in the abstract.

AR: [address frist sentence] We deleted the reference to Grass et al. because referencing it in full would have exceeded the word limit for the abstract.

RC: Page 2.

- "small to large" implies that the effect sizes vary. - "always in conflict" often?

AR: We rewrote the respective sentences as the phrasing obviously was misleading. The respective sentences now read:

“with these associations being in a small to medium range.”

“... because the worker is often in conflict ...”

RC: Page 3. How is NFC related to other constructs like neuroticism, self-esteem, and locus of control? And would NFC provide incremental validity in predicting burnout?

“adults” in which context?

please provide more information on the Grass et al. study in terms of sample, context, and sample size. Why was this study the focus of a replication?

AR: This is our response.

Assignment: **JZ**

4. Comment: see Table 2 in Fleischhauer et al. (2019)

RC: Page 4.

how are demands too high, demands too low, and demand-resources fit related to each other?

AR: This is our response.

RC: Page 5.

Are data also shared openly for further research?

Note that I couldn't assess the preregistration, but that the OSF offers a way to anonymise preregistrations for peer review.

AR: This is our response.

RC: Page 6.

Which nominal Alpha level is used? .01 or .05? Please explicate all deviations from the preregistration.

AR: This is our response.

RC: Page 7.

Note that reliabilities are not characteristic of scales as they are also a function of variation of true scores in the samples.

AR: This is our response.

RC: Page 8.

Why did the authors choose (preregister?) to use parcels instead of the five and seven-point items scores as input in the SEMs? The mean and variance corrected WLS estimator in Lavaan should be able to handle the raw item data well as long as the model is not too large.

How were bootstrapped fit measures computed?

AR: This is our response.

RC: Page 9.

How are the three demand factors related to each other? shouldn't they be allowed to covary in the SEMs? why do the authors assume that NFC completely accounts for their covariation?

AR: This is our response.

RC: Page 10.

Unfortunately, the appendix does not present fit measures of the model without the outlier.

AR: This is our response.

RC: Page 13.

How is the RSMEA computed? My computation with $RMSEA = \sqrt{((chisq-df)/(df(n-1)))}$ gives a value around .14 instead of .000

RC: Pages 15-16.

please provide the DF of the model alongside the Chi-square. The model doesn't fit very well and it is not clear that non-normality accounts for this as the authors are using robust chi-square. Please consider potential model revisions of the model in figure 2 using modification indices and clearly report these as an exploratory analysis. My expectation is that much of the misfit is created by additional covariances between the DTH, DTL, and DRF factors.

AR: This is our response.

RC: Page 18.

Here the authors do present another exploratory model that did consider ad hoc revisions using the modification indices, but they also added additional variables to the model instead of choosing to revise the original model in Figure 2. Actually, in the new model the authors changed the entire outcome variable from burnout to a study of EE and RPE (but without DP). For this reason, the text and study lose much of their focus: wouldn't it be better to focus on an overarching model that incorporates all relevant prior expectation and/or compares different (nested) models to pit ideas against each other?

AR: This is our response.

RC: Pages 20-24.

In their discussion, the authors also focus on different aspects, namely the replication of Grass et al, the demand-resource ratio model, and the exploratory analyses with personal efficacy and covid burden. A more general framework would probably work much better in interpreting the results and relating them to the wider literature.

AR: This is our response.

5. References