

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 thoroughly assess our manuscript. In the following we address your and each reviewer's concerns point-by-point. We hope that we could adequately address all the issues raised. In this revision, we have also updated the figures, they are now created with Inkscape rather than RStudio, which gives a more elegant result. The content of the figures did not change.

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: Thank you for pointing this out. We changed the term "teacher trainees" to "student teachers" to make clear that the participants in the Grass et al. sample were still attending university. We also emphasized that the replication was done in a teacher sample in the respective parts of the manuscript.

Thank you for suggesting a more elaborate consideration of sample differences. We have added the following paragraph to the discussion:

Overall, the differences between Grass et al. (2018) and the present analysis suggest that burnout is being mediated by different factors in student teachers than in teachers. In a classroom context, controlling oneself and one's emotional expression might be more relevant in dealing with the workload than in a university context, where a person is responsible for themselves first and foremost. This could even imply a different interpretation of the concept of personal efficacy in these two samples – student teachers might focus on how well they can follow lectures, complete work on time, and schedule their tasks efficiently, while teachers might focus on how well they can manage their students and whether they get good grades. Therefore, the understanding of personal efficacy could change from one's own success to one's students' success depending on the career stage.

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 the 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: Thank you for this suggestion. Our intention was not to focus on the significant mediations of Grass et al. (2018), but to deepen the understanding of the role of NFC in burnout. Therefore, we focussed on RPE, because it is likely the most relevant subscale of the MBI scale in the context of NFC (as Grass et al. also argued). While Grass et al. (2018) did not find significant correlations between NFC and the subscales DP and EE, Naderi et al. (2018, http://saber.ucv.ve/ojs/index.php/rev_lh/article/view/15958) showed that in students, NFC was significantly correlated with DP and EE, but less so than

with RPE. It has been shown that NFC is positively associated with higher self-efficacy (Rooij et al. (2017), <https://doi.org/10.1080/13596748.2017.1381301>, and Tan et al. (2020), <https://doi.org/10.1007/s12528-019-09239-6>), so we also expected NFC to be more strongly associated with the RPE subscale than with DP or EE. And if the focus in burnout studies with teacher samples has been on the DP and EE subscales, it is all the more reason to also focus on the RPE subscale, especially because NFC was equally correlated with all MBI subscales in our teacher sample. We had preregistered the exploratory analysis of the Demand-Resource-Ratio model with RPE in place of the MBI score, but this model did not have better fit indices than the initial model, which might be attributed to the evenly distributed associations between NFC and the subscales. Furthermore, when constructing the exploratory model it became clear that NFC was indeed associated with RPE and EE via DRF and DTH, respectively, but the DP subscale did not have any meaningful paths within the model. Since we have not preregistered any analysis of the association between NFC and DP or EE, and we do not have any theory-driven hypotheses regarding these associations, we would like to refrain from including these analyses in the paper in order not to report analyses that are less well-reasoned and not pre-registered. A change that we will gladly implement is the more specific distinction between the terms “burnout” and “RPE subscale” in the manuscript, thank you for this advice.

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:

Burnout has become more prevalent, mainly in social jobs, and there is evidence that certain personality traits promote or protect against burnout. Only recently, studies have focused on investment traits like Need for Cognition (NFC), the stable intrinsic motivation to seek out and enjoy effortful cognitive activities.

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

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

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: Although one can assume that Covid burden and years of experience are associated, there is a variety of factors that may also play a role. Firstly, years of experience cannot be equated with age, as some teachers are lateral entry employees. Secondly, age itself is not the only risk factor for Covid, but one of a long list of risk factors according to the Center for Disease Control and Prevention, such as disability, asthma, cancer, chronic heart disease, diabetes, HIV, weak immune function, obesity, mental disorders, pregnancy, physical inactivity, etc. Since the item does not specify the kind of risk factor, there is no need to correct for any confound between Covid burden and years of experience.

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: Thank you for addressing this. We do not consider “being busy” to be the same as “not being bored”, because while an employee might have a high workload, the tasks in question might not be interesting or stimulating at all. In the case of teachers it might be grading essays on the same topic they have taught year after year, or staying at work until the late evening for parent-teacher conferences. We looked into the literature on boredom in teachers, which is unfortunately quite sparse, and found that boredom was one of the main reasons for leaving the job (Tye & O’Brien 2022, <https://doi.org/10.1177/003172170208400108>), teachers were just as bored during class as their students (Tam et al. 2019, <https://doi.org/10.1111/bjep.12309>), and teachers have a strong feeling of mission and responsibility but lack the opportunities to implement these priorities (Tuisk 2007, <https://doi.org/10.2478/v10099-009-0011-8>). Therefore, we cannot claim that boredom is less prevalent in the teaching profession. As the distinction between “being busy” and “not being bored” is reflected in the positive correlation between DTH and DTL, we have added a more in-depth discussion of this correlation in the manuscript (see also our response to the fifth comment of Reviewer 3):

However, the positive association of *demands too high* and *demands too low* in the correlation analysis suggests that boredom and burnout are not mutually exclusive. A teacher can have a very high workload, which is reflected in an elevated *demands too high* score, but the work in question might not be intellectually stimulating at all or might not appeal to their ideals and goals, which increases their *demands too low* score. Both of these variables were correlated with higher MBI scores, but in this model, high demands appear to have contributed more to explaining variance in MBI scores than low demands, suggesting that the perceived workload is more important in the burnout context than how stimulating the work is.

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: Thank you for this suggestion. We have changed the numbering of the supplementary material into a S2.1, S2.2, ... notation.

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 Table 2. In our initial submission, there was an error in the code creating the table, so the diagonal with Cronbach's α and MacDonald's ω was missing.

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: Thank you for this assessment. It is true that the theoretical framework shifts within the manuscript. This is motivated by the three different aims of the paper. As the investigation of demands and resources in the context of Need for Cognition is very new, a smaller confirmatory model and a larger exploratory model likely contribute more to developing this approach than a highly complex and long paper with a theoretical framework that is still in its infancy. The overarching framework in this manuscript is the role of Need for Cognition in burnout or burnout subscales, and depending on which specific aim the reader is interested in, they may skim over the others. We have added subheadings to the introduction and expanded the last paragraph before the methods section to better convey these three aims:

To sum up, this study used one questionnaire data set to investigate three aims: Firstly, replicating findings of mediation between NFC and reduced personal efficacy in a teacher sample to identify possible differences between teachers and student teachers. Secondly, investigating the impact of different demand-resource-ratios on the relationship between NFC and burnout to see whether teachers with high NFC tend to overestimate their own resources or underestimate the current demands, respectively. And lastly, exploring the impact of other variables in the burnout context, such as the perceived burden by the pandemic, to create a starting point for future research.

Regarding the covariation of DTH, DTL, and DRF, please see our response to your comment about pages 15-16 below.

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: Thank you for pointing this out. We deleted the reference to Grass et al. because referencing it in full would have exceeded the word limit for the abstract. In fact, the limit of 150 words is so small that we cannot include all variable names, so we rewrote the abstract to better explain the three aims of the study. It now reads as follows:

Burnout has become more prevalent, mainly in social jobs, and there is evidence that certain personality traits promote or protect against burnout. Only recently, studies have focused on investment traits like Need for Cognition (NFC), the stable intrinsic motivation to seek out and enjoy effortful cognitive activities. This study had three aims: First, replicating findings of NFC and the burnout subscale reduced personal efficacy in student teachers in a sample of N = 180 teachers, second, investigating the role of perceived ratios of demands and resources in the context of NFC and burnout, and finally, creating an exploratory model as a starting point for future research. The results indicated that unlike the student sample, the teachers' association of NFC and reduced personal efficacy was mediated by self-control but not reappraisal. Teachers with higher NFC and self-control also had lower burnout because they experienced their resources as fitting to the demands.

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: We have added the correlations of NFC with neuroticism, self-esteem, and locus of control in the manuscript:

NFC on the other hand shows opposite associations with those variables ($r = -.19$ with passive coping (Grass et al., 2018), $r = -.19$ — $.33$ with neuroticism (Fleischhauer et al., 2019), $r = .42$ with self-esteem (Osberg, 1987), $r = .32$ with internal locus of control (Fletcher et al. 1986)), suggesting that people high in NFC are less prone to experience burnout.

In the study done by Fleischhauer et al. (2019, <https://doi.org/10.3389/fpsyg.2019.00420>, Table 2), NFC could provide incremental validity in predicting the MBI score and its subscales over and above the Five Factor model. The adults in this study were a population-based sample of over 4,100 participants from different work backgrounds.

Thank you for suggesting to provide more information on the characteristics of the Grass et al. sample. We have added the following information to the manuscript:

Grass et al. (2018) investigated such a mediation in a sample of 167 students, who were studying to become teachers in primary, secondary, special needs, or vocational schools, and who had been studying for $M \pm SD = 5.13 \pm 2.63$ semesters.

We wanted to replicate the findings of Grass et al. (2018) in a sample of teachers to investigate the different mechanisms that influence the development of burnout in different life/career stages within the same profession. As burnout mostly affects social jobs and theories of person-environment fit have been discussed for quite some time, it begs the question whether the structure of social jobs causes burnout or whether the type of employee in social jobs is more prone to burnout. Studying the same profession in different stages can shed light on this issue.

RC: Page 4.

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

AR: We did not have specific hypotheses about the relations of the demand-resource-ratios to each other. Initially it would stand to reason that they are not substantially correlated at all, since each participant might have a unique level of demands and resources, so they should score high in one of the variables and rather low in the other two. However, as they correlation table (Table 2) shows, this is not the case. DTH and DTL are positively correlated, and both are negatively correlated with DRF. Although initially counterintuitive, these associations are quite likely in practise: A teacher can have a very high workload in a strict time frame (high DTH) but the work itself might not be interesting or cognitively stimulating at all (high DTL), such as teaching the same material year after year, not being able to influence the syllabus, teaching at a much lower standard than what the teacher learned at university, keeping notorious troublemakers under control, having parent-teacher conferences until the late evening, etc. So a surplus of demands and a surplus of resources can co-exist, but not together with a demand-resource-fit. We have added the following section to the discussion:

However, the positive association of *demands too high* and *demands too low* in the correlation analysis suggests that boredom and burnout are not mutually exclusive. A teacher can have a very high workload, which is reflected in an elevated *demands too high* score, but the work in question might not be intellectually stimulating at all or might not appeal to their ideals and goals, which increases their *demands too low* score. Both of these variables were correlated with higher MBI scores, but in this model, high demands appear to have contributed more to explaining variance in MBI scores than low demands, suggesting that the perceived workload is more important in the burnout context than how stimulating the work is.

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: We are aware that the double-blind review process somewhat interferes with Open Science issues. We learned how OSF projects can be anonymized only after submission of our manuscript. We apologize that we did not give the reviewer the opportunity to check our preregistration/data/code. Of course, we share our data and code openly on OSF and also provide a license for re-use (CC-4.0-BY). The anonymized link is: https://osf.io/36ep9/?view_only=5169a1624e6d40c193bdc96121266ea9.

RC: Page 6.

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

AR: Thank you for identifying this lack of clarity. The original data collection was done in the context of a thesis with different hypotheses than this study. Therefore, the sample size calculation of the original data collection reads as follows:

The initial study was aiming to recruit a sample of N = 287 teachers as determined using the R package pwr (Champely, 2020), assuming a small to medium correlation of $r = .20$ between the measures of interest and targeting at a power of $1 - \beta = .80$ at $\alpha = .01$.

And the post-hoc power analysis of the current study is presented at the end of the same paragraph:

Using the smallest standardized indirect effect from the mediation in Grass et al. (2018) in a post hoc power analysis with G*Power (Faul et al., 2007, 2009) yielded a power of $1 - \beta = .85$ for linear multiple regression, given $\alpha = .05$, $N = 180$, and $f^2 = .05$.

This is the alpha-level we had preregistered and have adhered to in the manuscript.

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: We revised the respective sentence that now reads:

“... the Cronbach’s Alpha of the three demand-resource-ratios in this sample were *acceptable*.”

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: Parcelling items for SEMs has been shown to increase goodness-of-fit, reduce bias, and alleviate the effects of non-normally distributed data (e.g. Bandalos (2009, https://doi.org/10.1207/S15328007SEM0901_5), Hall et al. (1999, <https://doi.org/10.1177/109442819923002>), Matsanuga (2008, <https://doi.org/10.1080/19312450802458935>)). To benefit from these effects and reduce the number of parameters, especially given the sample size of $N = 180$, we wanted to use item parcelling for the Need for Cognition Scale. Additionally, because the MBI was also modelled in a way similar to item parcels (i.e. the three subscales were used as indicators), modelling the NFC scale with item parcels gives in a more balanced model.

We used the standard bootstrapping algorithm implemented in *lavaan*. To make our procedure more transparent, we now also provide the seed and the random number generator.

Following Grass et al. (2018), bootstrapped confidence with $N = 2,000$ replicates and the seed 13 were computed with the Bollen-Stine bootstrapping of *lavaan()* to account for deviations from normality.

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: We do not assume that NFC completely accounts for their covariation. We have computed an exploratory model which includes the covariation of the three factors as you suggested in the comment below, and it was not superior to the confirmatory model regarding fit measures (please see our response to your comment about pages 15-16).

RC: Page 10.

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

AR: That is right, thank you for pointing it out. We have added the fit measures of the replication of Grass et al. (2018) and of the demand-resource-ratio model without the outlier in the supplementary material below Table S2.2 and Table S2.3, respectively.

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

AR: The *lavaan*-function *fitMeasures()* calls the function *lav_fit_measures()*, which uses the robust RMSEA formula by Broseau-Liard, Savalei, and Li (2012, doi: 10.1080/00273171.2012.715252, equation 8). It corrects

for nonnormality of the data, which might be the reason why $RMSEA = \sqrt{(\chi^2/df)/(n-1)}$ gives a slightly different result.

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: That was an oversight on our part, we have added the degrees of freedom to the Chi Square test statistic in the manuscript. We have also added the analysis code for the demand-resource-ratio model with covariance between DTH, DTL, and DRF to our *Analysis.R* script. The results indicate that the improvement in fit measures is very marginal: From CFI = .775, SRMR = .173, and RMSEA = .131 (model without covariance) to CFI = .831, SRMR = .091, and RMSEA = .114 (model with covariance). None of the fit measures indicate good fit. Therefore, we would like to refrain from including this model as yet another exploratory/supplementary analysis.

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: We did start the exploratory model with the intention of using a full model, but lavaan was not able to compute it. For transparency, we have added the intended code of the full model into our *Analysis.R* script:

```
fullysat_model <- ' # measurement model NFC =~ nfc1 + nfc2 + nfc3 + nfc4 DTH =~ dth1 + dth2 +  
dth3 DRF =~ drf1 + drf2 + drf3 DTL =~ dtl1 + dtl2 + dtl3 MBI =~ ee + dp + rpe SC =~ scs COV =~  
covb YST =~ years
```

structural model

```
DTH ~ hnfc * NFC + hsc * SC + hcov * COV + hyst * YST  
DTL ~ lnfc * NFC + lsc * SC + lcov * COV + lyst * YST  
DRF ~ fnfc * NFC + fsc * SC + fcov * COV + fyyst * YST  
MBI ~ mdth * DTH + mdtl * DTL + mdrf * DRF
```

```
ee ~~ rpe  
ee ~~ dp  
dp ~~ rpe
```

```
NFC ~~ YST  
NFC ~~ SC  
NFC ~~ COV
```

```

SC ~~ COV
SC ~~ YST
YST ~~ COV

DTH ~~ DTL
DTH ~~ DRF
DRF ~~ DTL

dth1 ~~ dth2
dth1 ~~ dth3
dth2 ~~ dth3

drf1 ~~ drf2
drf1 ~~ drf3
drf2 ~~ drf3

dtl1 ~~ dtl2
dtl1 ~~ dtl3
dtl2 ~~ dtl3

# Indirect effects Indirect1 := hnfc * mdth + hsc * mdth + hcov * mdth + hyst * mdth Indirect2 := lnfc * mdl + lsc * mdl + lcv * mdl + lyst * mdl Indirect3 := fnfc * mdrf + fsc * mdrf + fcov * mdrf + fyst * mdrf

# contrast (if significant, the effects differ) Contrast := Indirect1 - Indirect2 - Indirect3

# total effect Total := Indirect1 + Indirect2 + Indirect3 ' However, this code yielded multiple error messages:
lavaan was not able to compute standard errors as the model might not be identified, it could not invert
the information matrix, and estimated some latent variables with negative variances. In short, the model
could not be computed, so we used the modindices() function to modify it step by step, starting with the
modifications with the biggest impact on model fit, but implementing only those that were contentually
meaningful. We did not document this process, which would have been a good choice in hindsight. We
adapted the explanation in the results section as follows:

```

We started with a full model with all variables except for the ERQ and its subscales as they had hardly any correlations with the other variables. This model could not be computed by *lavaan*, so we modified the structure using the *modindices()*-function in order to increase the goodness-of-fit indices within the framework of contentually meaningful variable relationships.

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: Thank you for this suggestion. The three-part-structure was exactly what we intended with this manuscript; a replication in a different sample, an analysis based on our own previous research, and an exploratory analysis to serve as a starting point for future research. (As we have provided an answer to this issue in our response

to your first comment, we would kindly refer to it here.)