Personality and Life-Satisfaction: A Hierarchical, Multi-Rater Perspective

Dear Dr. Schimmack,

Decision: Reject after Peer Review

I have now received two reviews of your manuscript "Personality and Life-Satisfaction: A Hierarchical, Multi-Rater Perspective", which you submitted for consideration by the Journal of Personality and Social Psychology: Personality Processes and Individual Differences. The reviews that I obtained were provided by experts in areas related to your work, and I found their comments to be detailed, thoughtful, and constructive. As you will see, both reviewers thought that the broad area of research was interesting and important, and each identified some strengths of the methods. However, these reviewers also identified some limitations of the work, and these concerns are serious enough to prevent me from accepting this paper for publication in JPSP. Below I describe the reasons for this decision.

I wish to stress that both reviewers and I found some promising potential in this research. For example, Reviewer 1 noted: "I think there may well be merit in the author's intentions." Similarly, Reviewer 2 wrote: "I think using multi-rater assessment and novel modelling approaches to better understand personality-well-being relationships at domain and facet level is important. So I do encourage the author to pursue the topic further. I also appreciated the use of the OSF for data and analyses." Despite their enthusiasm, both reviewers recommended rejection, and after reading their reviews and rereading your article, my decision aligned with their recommendations. Nevertheless, I wish to emphasize that your methodological (multi-rater) and modeling (hierarchical CFAs) approaches are indeed novel and creative. Despite these strengths, the manuscript has some potentially serious limitations, which both reviewers outline in their respective reviews.

The reviewers' comments are thorough and complete. I will not reiterate each of them here. Nevertheless, like Reviewer 1, I felt that readers would benefit from a more detailed description of each sample, the items used, and modeling choices for each analysis. As Reviewer 1 puts it, "The author may have good reasons to stand by their claims, but they need to articulate a much stronger case and actually engage with what was done in the work they criticize — and, of course, they need to actually describe their own measures." I also generally agree with Reviewer 2's assessment that "the current paper is still quite rough," including lacking any text under the "Conclusion" heading.

Regarding Reviewer 1's comments, he wrote me after submitting his review to note that he may have neglected or misinterpreted a couple of issues in his review. He wrote:

"…from another reference I found, I think I was wrong about the items they used to assess the two Big Five facets (they did not describe what they used so I tried to reverse engineer this).

"I think they actually used "real" personality items:

"Second, these facets were also used in Dobewall et al. (2013) study. Thus, we were able to replicate and extend their findings using the same personality constructs. Facets were measured with three-item scales. The Cheerfulness items were: ‘I tend to be a cheerful and high-spirited person’, ‘I tend to radiate joy’, and ‘I tend to have a lot of fun’. The Depressiveness items were: ‘I tend to feel blue’, ‘I tend to feel discouraged’, and ‘I tend to give up when things go wrong’.. So, that part of my criticism is probably invalid. You may want to indicate this in your decision. This does little to change my overall assessment, though."

To be sure, I think this paper advances some potentially important measurement contributions, but falls short of making a substantial theoretical contribution. In other words, I think the broad question you are investigating is important. Although I think the Mõttus et al. (2024) could be methodologically or data-analytically improved somewhat along some of the lines you’ve suggested, it’s unclear to readers how this re-analysis of sorts advances theory related to personality psychology beyond Mõttus et al.’s (2024) contribution (or other similar papers on this topic).

Please note that the number of submissions to JPSP has increased quite dramatically over the past few years, and space limitations dictate that only a small percentage of these can be published here. This means that many very good papers that have no fatal flaw cannot be accepted. I wish you well with this interesting work, and thank you for considering JPSP as an outlet for your research.

Please contact me if you have any questions, concerns, or suggestions about this decision letter or the broader editorial process at JPSP: PPID.

Best wishes,

Reviewers' comments:

Reviewer #1: I started to read the submission with great interest, but quickly realized that it reads more like an (early) sketch of a paper than a final submission in which the author fully describes the problem to be addressed, while also comprehensively addressing the relevant existing research; carefully describes the data, analyses and results; and fully discusses the findings — again, adequately addressing the relevant existing literature.

I also quickly realized that I knew who the author was because they were using the data I had provided, and we had several discussions, both by email and on social media, about their interpretations of my work and their intentions to re-analyse our data. However, I was not aware that the re-analyses would be submitted to a journal.

Most importantly, while I think it might be worthwhile to re-analyse already published results to either identify errors or gain new insights, and it is definitely worthwhile to present new data on an important research problem, it is not clear to me what exactly the author is trying to achieve in this particular version of the paper.

As I understand it, the author re-analyses a subset of our data and presents some new data, both to "reconcile" our findings with "the evidence in previous studies" and to model domain satisfaction as a mediator between personality traits and (general) life satisfaction. I will address the latter aim briefly below (it may be worthwhile, but should be explained in much more detail to make a clear case for it), concentrating here on the former.

In their reanalysis of the data from Mõttus et al. (2024), the author takes a subset of our variables and data (for example, they do not address data from Russian speakers) and reanalyzes them differently.

But why is this new type of analysis better, or why and how does it provide sufficient additional value? It may be, or provide, but explaining this would require describing in detail what was done in Mõttus et al. (2024) in the first place. Failure to do so is one of the major limitations of this paper. Mõttus et al. (2024) included a wide range of analyses, supplemented by robustness analyses, that included modelling personality traits and satisfaction in different ways, as well as including 10-year longitudinal analyses with results consistent with the cross-sectional analyses. None of this is described. Why? Nor does the author discuss that the associations with general life satisfaction were replicated when general life satisfaction was replaced by the sum of eight different domain satisfactions (and these items were not presented among the personality items, which the author criticises).

Nor is it at all clear why the author models personality traits in this particular way. Why use (only) these particular items in each Big Five domain and facet, from among the nearly 200 available items? And which items were actually modelled (to find this out, the reader will have to explore the Mplus scripts at OSF)? Why were these Big Five different/better than those modelled in Mõttus et al. (2024)? None of this is described.

Importantly, the data of Mõttus et al. (2024) are highly misdescribed. For example, personality traits were assessed with more than 100 items, and eight domains of satisfaction were assessed, not three. And why only consider three in this re-analysis? If it was because the three had both self and informant ratings, then this leads to another problem — the author chose/declared the wrong domain satisfactions because these three were not the ones with both self and informant reports. The author may have had good reasons for making certain choices, but they should have been explained and the data should not have been so obviously misdescribed — it just gives the impression of sloppy work.

As for presenting new data, this may well provide more incremental value. However, these additional data are not fully described, nor discussed in relation to that of Mõttus et al. (2024; size, comprehensiveness). How many participants in total and in each group? The Big Five were only assessed three items (which is a major yet undiscussed limitation, especially given author's conclusion that personality traits' associations with life satisfaction are faceted): but which three items? Next, only two facets were assessed, again each with three items. But which items? This is not described. Moreover, exploring other studies using these data (e.g., Schimmack & Kim, 2020; reveals that the three items may not have been designed to assess personality traits but mood/(dis)pleasure, coming from the positive affect and negative affect scales describe in an earlier study (Schimmack, Diener, & Oishi, 2002; as follows: "The questionnaire relies on three items to assess pleasure (pleasant, good, positive) and three items to assess displeasure (unpleasant, bad, negative). Ratings were made on 4-point intensity scale (not at all, slightly, moderately, strongly)." If these were really the items used to assess the Big Five facets, I don't think it is appropriate to make the reader simply accept this at face value, not explaining or even referring to the scales.

[The author claims that, unlike in Mõttus et al. (2024), personality domains were assessed with "well-established" (p. 13) items. Unless explained, this is a strong and likely misleading statement. First, no serious psychometrician would say that three items provide an optimal way of assessing broad Big Five domains; in Mõttus et al., (2024) each domain was assessed much more comprehensively with numerous items. Second, the author should explain why the items used in Mõttus et al. (2024) were less-established, given the rationale and extensive evidence provided for selecting these particular items from existing item pools (Henry & Mõttus' work referred to in Mõttus et al., 2024). Finally and ironically, the author does not even describe which items they used to assess personality traits in the novel data and possibly used items not designed to assess personality traits but mood. The author may have good reasons to stand by their claims, but they need to articulate a much stronger case and actually engage with what was done in the work they criticize — and, of course, they need to actually describe their own measures.]

Nor the models nor the results are appropriately described. For example, which items defined which traits? What were the model fit and path estimates? Instead, the reader is expected to explore the Mplus output files at OSF. This is not fair to the reader, and prevents most readers from being able to assess the findings; even if one is familiar with the Mplu syntax, this is a tedious task. This level of presentation simply does not allow me to assess the meaning of the findings. If the author can't be bothered, why should the reader be?

What does the "Other R2" refer to in Table 2?

In integrating the findings across studies, the author gives equal weights to findings based on 20,000 participants and a few hundred participants. This does not seem appropriate, unless there is a reason to down-play more powerful data that the author leaves unarticulated.

In the Discussion, the findings are interpreted in a fairly random way, with only aspects of the results consistent with the author's preferred narrative being discussed. For example, in arguing that facets (or mood scales, really) added little incremental predictive value, I am not sure what part of their results the author is discussing. It certainly does not seem consistent with Table 2, to the extent that its content is clear. I have no idea where the conclusion that personality traits explain abut 50% of life satisfaction variance comes from.

Also, as far as I can see, the author's results (based on their preferred model, not on how we analysed the data in multiple ways, as described in detail in Mõttus et al. 2024 main text and extensive supplemental material) based on a) students and b) our large Estonia-speaking sample converge on how much variance in life satisfaction personality traits explain - 64% to 66%. How does this lead to the 50% conclusion?

The prediction strength in Mõttus et al. (2024) that the author takes issue with was obtained with a different combination of traits (domains and two facets) that the author (explicitly at least; there seems to be more going on according to the Mplus scripts) prefers to use. Our approach was oriented in maximising out-of-sample prediction, not parsimony; we carefully articulate the reasons for taking this approach, which the author does not engage with.

Additional remarks about the Introduction:

There are more facet-level studies on the links between personality traits and life satisfactions than the author describes in the introduction (p3). For example, numerous studies have used the Big Five Inventory, also finding that only some of, say, extraversion facets are linked with well-being.

Having discussed extraversion and neuroticism, the author dismisses the other Big Five domains' correlations with life satisfaction as "less relevant". Why dismiss, say, conscientiousness, while its meta-analytic correlation is not substantially lower (0.27) than that of extraversion (0.32)? Is this because the author prefers the conclusion that personality traits' relations with satisfaction are through "affective experiences" while the dismissed results don't fit this narrative?

Variance not due to "heritable and stable personality factors" is not only due to "life-circumstances" (p. 5); much of this unaccounted variance is due to a) random measurement error and b) stable assessment biases.

It is factually incorrect that Mõttus et al. (2024) assessed life satisfaction with one item (we used three items, validated against other measures including the Diener's scale, but they also ran robustness analyses with one item). Likewise, I am not sure what it means that "personality was assessed with single items" (p. 5). Personality is usually assessed with single items that researchers then choose to aggregate in certain ways — as did we — and sometimes choose to also leave un-aggregated for additional granularity — as did we. If the author intends to suggest that the items used in this study do not assess, or allow assessing, the same Big Five domains than other Big Five assessments, including the Costa and McCrae's NEO-PI model, then it would be accurate to discuss this claim in the light of the empirical findings presented in Mõttus et al. (2024; correlations with the NEO-PI scales) and Anni et al. (2024; correlations with other Big Five scales). These findings suggest otherwise.

The author makes the case for hypothesizing stronger correlations among students than among their parents by extrapolating cross-cultural comparisons and stronger social policy among Western nations. I really cannot follow this logic. Was this an a priori (before seeing the data) or a posteriori (after seeing the data) hypothesis?

I found the logic for assuming domain satisfactions to be (at least partial) mediators between personality traits and general life satisfaction not fully explained. There are strong causal claims packed into this statistical hypothesis, and it is not universally accepted/supported, so it is unfair to make the reader simply accept this at face value. Also, it would seem appropriate that the relevant discussion in Mõttus et al. (2024) is also referred to; there, we chose not to model this mediation, articulating the reasons for not doing this. That is, I don't oppose the analyses, but simply expect better-articulated rationale for them (and of course, full details of the models and results).

In conclusion, I think there may well be merit in the author's intentions, but the quality of this work needs to improve dramatically to be publishable.

Reviewer #2: The current study analyses several datasets to examine the correlation between personality and life satisfaction. The paper touches on themes like (a) implications of multi-informant assessment, (b) domains versus facets, and (c) modelling approaches.
In general, the topic is very important. I think using multi-rater assessment and novel modelling approaches to better understand personality-well-being relationships at domain and facet level is important. So I do encourage the author to pursue the topic further. I also appreciated the use of the OSF for data and analyses.

However, the current paper is still quite rough. For instance, (a) the introduction could do with more nuance and could be reorganised for clarity; (b) the method is unclear and incomplete, (c) the results are unclear, (d) a section labelled "conclusion" has no content underneath. In short, I encourage you to develop everything further. But I think this is a very long way off the standards required in JPSP. But I imagine that with considerable polishing, it might be suitable for a different journal.

ASSORTED POINTS
1. The abstract says that there are five datasets; it would be good to have samples sizes for each in the abstract
2. You don't have a section labelled "Method". You should. In general, you should conform to APA style.

3. You have a heading called "Datasets" and then have information about Study 1, and then a subheading for "Study 2", which includes information about Study 3. But the abstract says that there are five datasets. The method needs to be substantially reorganised for clarity.

4. For each dataset, make sure that it is clear whether this a new dataset that you have collected or whether you re-analysing a previously published dataset.

5. Regarding "The key results are so strong that these biases are practically irrelevant." I think I know where you are going with this point, but you need to express your idea with more precision.

6. The description of the Study 1 dataset does not concisely state how the variables were measured. It might be possible to discern, but it could be expressed more clear.

7. Your analytic approach and your descriptions of your datasets should be disentangled.

8. Study 3 should include sample sizes.

9. Table 1: It is not clear from the title what the values in the table represent. Probably standardized betas. You need to expand abbreviations. You should include openness as a predictor even if it is non-significant. Is it self-report or not?

10. Regarding " Table 1 shows the results for the Big Five factors." What results? What analysis? What data?

11. Regarding " The result for Extraversion is consistent with the theory that affect is the key mediator of personality effects on life-satisfaction because cheerfulness is the disposition to have more positive affect." I think this might be overselling the point. There are a range of mechanisms through which personality is theorised to influence well-being. Of course, on some level pat of having well-being is the experience of more positive affect and less negative affect, but dispositional personality traits are not the full story. Furthermore, there are plenty of theories of how personality traits might have more instrumental effects separate from dispotiaionl affect tendencies (e.g., conscientiousness leading to health behaviours, meaning, purpose; extraversion leading to positive experiences; agreeableness leading to less interpersonal conflict, etc.)

12. Regarding " Anxiety and anger are negative affects, but they do not directly predict life-satisfaction. One possible explanation for this could be that these negative affects are functional and transitory during the pursuit of goals." It depends what you mean. They are correlated with life satisfaction, but I guess you are saying that they do not incrementally predict life satisfaction (or at least much) over and above depression (or global neuroticism).

13. You may want to think more about and integreate discussion of variables theories about how personality influences life experiences and their appraisal, set-point theories, top-down and bottom-up theories of life satisfaction.

14. In general, I think you should include openness. Ultimately, it is an empirical question as to whether it predicts life satisfaction. The benefit of the Big Five is that it seeks to model all the major factors of personality so estimates of variance explained and so on have a certain theoretical meaning.

15. I think you should be careful in how you think about what is domain and what is facet prediction of life satisfaction. Every facet includes variance shared with the domain (and potentially shared with other facets), and unique variance, and different modelling approaches divide up the shared variance in different ways: Different ways of residualising or dividing up that shared variance has influence answers you get. There is also a theoretical element to all this in how you think about personality causally influencing life satisfaction.

16. You have a section labelled conclusion but there is no conclusion.

REFERENCES