

Response letter (July 2022)

Dear editor and reviewers,

many thanks for taking the time and effort to read and discuss my draft. I am grateful for the comments made and think that the quality of the paper has clearly improved. Find below my detail responses to each question.

Editor

Dear Dr. Bittmann,

Thank you for submitting your work to Collabra: Psychology. I sent the paper to two experts who are well qualified to review this paper. I thank these individuals for their service to this journal and for their timely reviews. I independently read the paper and then consulted the comments from the two reviewers.

As you will read below, the Reviewers and I think this work has promise. I believe a revised manuscript would likely meet publication expectations at this outlet providing there are changes that address the questions/concerns raised by the Reviewers and in my letter. The Reviewers did an excellent job so please attend to each of their points either in the revision or in the letter of response. I will outline my reactions in this letter. I acknowledge that you might disagree with some (or all) of our points so feel free to pushback against any suggestions you believe will harm your work. Just describe your counterpoints in the letter.

My reactions in rough order of appearance in the paper.

1. The Introduction is well organized and efficient. However, I think more could be included about possible processes behind why certain variable are considered confounders. Indeed, I think a additional few sentences explaining the logic of the concerns about each specific variable will enhance the paper. In other words, explaining why it is important to control for specific variables and what that means for the inferences drawn from the analyses would enhance the paper.

My response: I agree that this is a sensible suggestions. I have added information on this in section 1.1 and explain in more detail why omitted control variables are problematic. See also section 2.2.6. where information has been added.

2. A few minor issues in the Introduction/Method section:
 - a. I think you mean scales on page 3 rather than Big Five items?
 - b. I think the discussion about confidence intervals on page 4 should specify that you mean

CIs on standardized coefficients.

c. It might be useful to spell out the implications (if any) of moving from 5,778 to 4,607 students).

My response: agreed, I have amended all 3 points.

3. It might be useful to note potential concerns with attenuated effect sizes when using short measures of the Big Five domains. Crede et al. (2012) is an almost classic citation now:

Credé, M., Harms, P., Niehorster, S., & Gaye-Valentine, A. (2012). An evaluation of the consequences of using short measures of the Big Five personality traits. *Journal of personality and social psychology*, 102(4), 874.

My response: Thanks for pointing that out, I have added the reference.

4. Would including path coefficients in a version of Figure 1 make sense? I know this is complicated because there are 5 personality trait domains, but I thought I might be worth considering so readers get a picture of the effect sizes in question and an understanding of which personality domains are most relevant to the current discussions.

My response: I have thought about this for a while and made some graphs. I think its possible in theory but its messy as there are many more boxes and arrows (one for each trait). The second issue is that coefficients and R2 values are then mixed together. In my opinion many people have already problems to really understand what the Figure entails (as they are so used to path coefs or SEM models that they automatically interpret the numbers shown as coefs, even despite my best efforts to label and describe the model). Overall, I think that such a model with all coefs would be difficult to understand and packed and I believe it makes more sense to rely on the additional tables for a detail understanding.

5. I think Reviewer #1 raises a good point about alternative models. What would happen if SES and cognitive ability were simply considered as correlated predictors of educational attainment rather than assuming SES as a statistical predictor of cognitive ability?

My response: I have commented on this question in detail below when answering the reviewer, please see there.

6. I think the point about the importance of maternal cognitive ability on page 17 being included or omitted from models needs to be explained a bit more for readers.

My response: I have added some more information on this on page 17.

7. I appreciate the Discussion of limitations in the paper, but it might be too brief. For example, it might be helpful to readers to spell out if there are plausible untested alternative models that could be considered in future research. Likewise, I think you might be able to flesh out concerns with secondary data analysis and with causal inference using observational data. This way readers are crystal clear about the strengths and limitations of the current work.

My response: clearly, there are always untested models as no paper can test “all” the models. Personally, given the quite focused research question of the replication, I do not think that there are many obvious models out there that were not tested (otherwise, why not just test them to enhance the paper?). Therefore, I am a but reluctant to give a more

“generic” outlook (sure, we know there is always more research ahead). Also given that I do test some more models as I follow the ideas of reviewer 1, I kept this addition brief. Regarding the other point of causality, I have added some information.

8. I think the suggestion from Reviewer #1 to specify models with interaction terms would be worthwhile. These just need to be indicated as exploratory analyses to readers and presented with appropriate caveats. This paper might be interesting to consider:

Damian, R. I., Su, R., Shanahan, M., Trautwein, U., & Roberts, B. W. (2015). Can personality traits and intelligence compensate for background disadvantage? Predicting status attainment in adulthood. *Journal of personality and social psychology*, 109(3), 473.

My response: agreed, see the additional analyses under section 3.3.

Reviewer 1

The purpose of the research reported in this manuscript was to further investigate relative influences of socioeconomic status (SES) and psychological traits (cognitive ability, personality) on educational attainment (EA). Using dominance analysis, the author shows that cognitive ability and conscientiousness each account for about twice the amount of variance in EA as SES does (roughly 5% vs. 2.5%). This is an important, newsworthy finding, and the study has much to recommend it, including validated measures of the psychological traits and a large-N (from Germany).

I have two major comments:

- (1) The present analyses consider main effects of the predictor variables on EA, but it would also be interesting to consider interactions. The interactions that seem particularly interesting are: SES x g, SES x C, and g x C predicting Grades. For example, does a high level of g compensate for a lower level of SES (an under-additive interaction)? or vice-versa: Does g amplify the effect of SES (an over-additive interaction)? I think it would be interesting to include these analyses under an “Additional Analyses” section.

My response: agreed, I have conducted the analyses as suggested. One interaction has been shown to be statistically significant and is presented in some more detail as an additional and interesting analysis. I really think that adds something to the paper!

- (2) It appears that the author has access to multiple indicators for each presumed latent construct, e.g., reasoning, speed, and mathematics reasoning for g. Why not include, in addition to the analyses presented already, latent-variable models?

My response: this is clearly possible but I am not sure what would be the main benefit given the research question. Undoubtedly, these factors of ability do have a positive influence on grades, which has been shown in thousands of studies before, so just proving this point is probably not relevant for the replication itself. The only real benefit would be if the explained variance behaves differently in latent models, but as my tests have shown this is not the case. So I see little to be

gained from presenting these extra models. I think it is more interesting to the reader to see the interaction models as suggested above so I have only included these in the revision.

And one other comment:

(3) Is it worth considering alternative models? The present analyses assume SES influences g (see the path model figure). But this model may not be correct: one could make an argument that the variables should just be correlated, or perhaps even that the path goes from g to SES based on gene-environment correlation, i.e., that the g variable reflects genetic variance, and that some of same genetic factors driving this contribute to SES in parents (bright parents are able to earn higher wages and go farther in school, and pass these genes along to their kids).

My response: I understand the point but empirically this is of little influence. First, regarding the main outcome variable, grades, there is no difference. In this model, as can also be seen in table A2, the measures of SES and the cognitive ability are independent variables that each have some influence in explaining grades. In this model it is irrelevant whether one assumes if SES predicts ability, ability predicts SES, both are correlated, or the both are completely independent. From a statistical point of view, these models are equivalent to explain grades. Therefore, the main results of the study are not affected by this decision. Second, one could say that the model where ability is the dependent variable, which is only shown in table A2, is useless and should not be considered further. Sure, if one believes that the two variables are independent or only correlated (but not causing each other), one could say this model is irrelevant. It was just my notion that this additional model might be interesting but I agree that other researchers might have a different opinion. Clearly, SES and ability are correlated but with my data I cannot recover any causal relationship. I have added some of these caveats in section 1.2.2. but I believe there is no need to change the models itself given the research question.

Reviewer 2

This is a strong paper. It takes as a starting point a different recently published paper, with a model that makes claims around the relative weights of personality traits and cognitive ability in predicting educational outcomes for a large representative sample of British students. Using a large sample of German students, it replicates the original model. However, it then goes further by employing a more sophisticated statistical design, and demonstrates that school placement within the German system is a mediating variable. This- as the title suggests - both replicates but also extends the original model. Detailed sophisticated statistical analyses, with a large representative sample, robust measurements, and a clear interpretation.

My response: thanks, this positive comment is very much appreciated.