

Peer Review

Review of: "Does Teacher Feedback Potentiate a Positive Feedback Loop That Underlies the “Achievement Gap”? The Role of Race and Gender in the Interrelations Between Self-Perceptions, Teacher Feedback, and Academic Performance at West Point"

Anna Maria Mayer¹

1. Catholic University of Eichstätt-Ingolstadt, Eichstätt, Germany

This study is an interesting article on longitudinal interactions between supervisor judgements, self-beliefs, and GPA. It considers race and gender stereotypes in a US-based predominantly White setting – the military. Generally, it offers important insights into teacher feedback and long-term effects on students' outcomes and careers. A short explanation for the overall rating: It is a solid article that meets most standards I would expect. However, there are still some details that warrant attention. Most importantly, I miss a clearly formulated research aim with hypotheses. The information is somewhat implicitly mentioned at different points in the manuscript, but I strongly suggest including those explicitly in the present study section. Additionally, as the article is US-centred, it should be mentioned. Some background information on the US military might be common knowledge in the US, but not in other parts of the world. Below, I provide more detailed comments and questions for each section.

++Title++

When reading your items for instructor ratings, I wondered whether the term Feedback in your title is adequate. I would have expected written or verbal detailed feedback and not two ratings on a scale. Consider changing the title to not promise more than can be expected from the article. Additionally, “West Point” is a non-explained term in the text. As a non-US researcher, I have no clue what this could

mean, and I would suggest deleting this information from your title, except if it really contributes to your research aim. If it does so, introduce West Point in your text.

++Abstract++

General comment across the text: An association across years is not really a dynamic process. Dynamic processes usually focus on variability in the short term, while your research focuses more on mean-level changes. As such, I would suggest changing the term. Moreover, your first and second sentences are content-wise very similar. Please consider shortening and including your research questions/hypotheses instead.

Minor comments: Include a more detailed sample description (minimum: age with SD, gender, race). You could consider shortening the methodological details (e.g., “which accounts for between-person average levels of each variable,...”) and elaborating more on the results instead.

I am unsure whether keywords are needed, but they surely would be helpful

++Introduction++

The beginning of the introduction is well-written and delivers the importance of the study well! One comment upfront: Your citation is not coherent. Sometimes you cite with numbers and sometimes with surnames. This should be coherent.

When you write about the military for the first time, it would be very useful for non-US readers to know that you are writing for the US context. For example, as a European, I would not agree that the military is a meritocracy with better race relations. Furthermore, I cannot judge whether White officers are extremely, somewhat, or adequately represented at the officer level.

You introduce stereotype threat and discrimination experience as explanatory mechanisms. Was this assessed in your study and could be included as a measure for your statistical analysis?

You describe how USMA’s instructor evaluation is comparable to feedback on p. 4. First, consider moving this section to “The specifics of the USMA environment” section to avoid having similar information at different points in the manuscript. Second, in this description, I am not entirely sure what the USMA feedback entails and what it does not. Please reread the section and provide more reader guidance to ensure understandability – even for non-US researchers.

Move the descriptive statistics to the results or sample description section. They are not part of your theory review.

I also noted that some information is duplicated in your introduction. You could consider removing redundant sentences or sections.

++The present study++

Again, longitudinal mean-level changes should not be considered dynamic.

In this section, I would expect your research questions and hypotheses. They seem to be missing completely and should be added. If none were present when writing the paper, please indicate your research as exploratory. This should then be mentioned in the limitations. Generally, this section includes information on items and statistical analysis. Both pieces of information should be moved to their respective sections (Measured variables & Statistical Analysis). Currently, it is unclear why LMM and RI-CLPM were used. They seem to be two analytical strategies for the same aim. The reason for this choice might become clear by adding the study's hypotheses. If both strategies follow the same aim, consider moving the less informative one into the supplement.

++Participants+++

Information on model estimation should be moved to the "Statistical Analysis" section. Please add more sample characteristics, if possible. These would include: Country of study, SES of students. Refer to Table 1 in this section.

I would assume that rater characteristics would influence the rating of different student groups differently. Is there information on the raters available that could be included in the analysis? If not, can you describe them on a general level to contextualize the setting?

++ Measures++

This section should only include measures. Move the paragraph about the rating system to the setting section. Please also provide an item for empathy and tact. Regarding GPA, provide the minimum and maximum of the scale. As a non-US person, I do not know whether the maximum of 4.33 is good or bad and therefore cannot verify your interpretation of your statistical analysis.

You could consider shortening this passage, along the lines of: "GPA was assessed from X to X. It included academic performance only. Military and physical GPA are not considered because. ...". Your measures section is unexpectedly lengthy, which, however, is not necessarily an issue.

++ Analytical Approach+++

The analytical approach is thoroughly described. In this section, you mention hypotheses for the first time. These should be added in the present study section. By doing so, it is also easier to understand why both analytical procedures are necessary (i.e., why are average differences of interest if you include them in your RI-CLPM as well? What is the added value here?).

++ Results++

In this section, you should report statistics for all effects. I assume these are partially mentioned in Table 1. If that is the case, you could reference it here and move Table 1 to this section.

++ Discussion++

As you mentioned discrimination experience or stereotype threat, I was wondering whether you could include this in your RI-CLPM to explain the longitudinal associations? The same applies to the race of the instructor. If that is available data, it would add an interesting insight. This is not a must; I was just wondering whether the data is accessible and why you did not examine it, since stereotype threat is an explanatory factor according to your introduction. It would contribute to the interpretation of your results on p. 22: (“...We cannot say whether the processes underlying poorer scores given equivalent entry tests and stronger internalization of teacher feedback for Black students were predominantly student-teacher interactions, implicit beliefs, explicit racism, structural/institutional elements, ...”).

In your limitations, you mention that instructor demographics could not be evaluated. Why was that the case (i.e., conscious choice or context restriction)? If it is restricted by the context, it would be helpful to describe the general demographics of instructors when you describe the setting to contextualize the teacher-student interaction.

++ Overall++

Minor detail: Model selection could be moved into the appendix, as it does not address any research hypotheses directly. If I am mistaken, the specific hypotheses should be mentioned.

Minor details: If you are a single author, it is fine to write “my” primary aim, instead of “our”. The reference to a group is misleading.

Declarations

Potential competing interests: No potential competing interests to declare.