

Reviewer Comments-2:-

While the manuscript is written clearly and easy-to-follow, and the authors are applauded for having the start to an interesting project, my recommendation is to “major revision” this manuscript.

This paper explores the relationship between spanking and externalizing behavior at varying ages, as moderated by “race” and “(child) gender.” Investigating potential “cultural” and “contextual” differences in salient correlates of problematic behaviors in childhood, like the current study attempted to do, represents a very important and ever-timely topic. However, this domain is also incredibly nuanced and complex, which thus necessitates great thoughtfulness regarding the importance of ensuring the scholarship conducted is of the highest quality possible.

This research is well-aligned with the scope of The Journal of Mental Health and Social Behaviour.

The manuscript is well-written and explores an interesting topic, with hardly any hiccups (e.g., APA-style formatting issues, typos) (see Minor Comments, below). The Introduction and Discussion sections, specifically, are thoughtfully presented and (seemingly) theoretically supported – although, of course, there are always concerns about whether Authors “cherry picked” studies to tell “their story.” That said, concerns arose about whether the literature review was biased at all, given the inherent biases of the social science literature itself (i.e., systemic racism and oppression within scientific reporting and publishing), it is difficult to ensure that “all” voices / perspectives are represented and avoid committing “confirmation (selection) bias.”

Despite these strengths, I have major concerns regarding the sufficiency of the research design, data employed, and, consequently, the methodological rigor of the study itself. While these comments are not exhaustive, efforts were made to provide specificity for guidance. Please see “Comments regarding Appropriateness” (textbox above) and “Major Comments #1-4” (below).

Therefore, although the overall presentation (extant literature; supporting theoretical framework) is strong, the design of the study and the research methodology employed, as well as the presentation of the data/analyses, do not meet the level of rigor and quality needed to thoroughly investigate such a nuanced, complex topic, like “cultural” racial* differences in parenting.

Major comments:

1. Operationalization / Conceptualization Issues

i. Race

1. Unclear how “race” was measured beyond “parent-report.”

2. Binary (Black vs. White) categorization oversimplifies the lived experiences and systemic inequities faced by Black families.

ii. Spanking

1. Single-item measure with no exploration of context or nuance - undermines the depth of the investigation.

2. To better understand the data distributions, reporting additional statistics (beyond mean and standard deviation) would be helpful

b. Limited Scope: “Cultural” and “Contextual” Factors

i. Weak support for framing the study as examining "cultural" differences

1. On page 9/10, paragraph 5 - the authors acknowledge that within-group comparisons (e.g., variability among Black households) might provide more meaningful insights than cross-racial comparisons.

ii. Framing the study as "racial" differences may be more accurate than describing it as "cultural" (although "race" may also represent an overly simplified "conceptualization" of a complex, nuanced "construct")

c. Ambiguity: "Gender" Measure

i. Unclear how "gender" is measured. Is it assumed? Was it actually "biological sex" being assessed? The distinction is critical, given that "gender" and "biological sex" are not the same thing.

2. Research Design/Methodological Weaknesses

a. Simplistic Design

i. Overall, the design/analyses are too simplistic for such a nuanced topic

b. Limited / No Contextualization:

i. Black and White families are compared without accounting for systemic inequities that may shape parenting practices and child outcomes - this "apples-to-oranges" comparison risks perpetuating biased narratives.

3. Statistical Concerns

a. No Corrections (for Multiple Analyses)

i. No corrections for family-wise error rate (FWER) or false discovery rate (FDR) were applied, increasing the risk for Type I/II errors

1. Corrected p-values should be calculated and reported

a. Would the findings still be significant at the .01 level?

b. Insufficient Statistical Reporting