Peer-review

Burce, Keisha L. 2025. “Harmony in Schools: The Role of Instrumental Ensembles in Fostering Student Belonging.” *Journal of High School Science* 9 (2): 1–22

I enjoyed reading the manuscript. The statistical treatment in particular was thorough with assumptions of normality and heteroskedasticity accounted for. I do have some questions and comments that need to be addressed and which I believe will improve the manuscript.
1.I am having trouble understanding this statement : “…….To ensure consistency, the items within the social exclusion subscale require reverse scoring as they are negatively worded. This ensures that responses in which lower values are positive, are reversed so that higher values indicate higher levels of belonging when summed to score.”. How were these values 'reversed" ? subtracted from 4? are the values in table 1 for the odd numbered questions already presented as reversed? for example, Q10, musicians actually scored 0.373 ? and non musicians actually scored 2.532 before reversing? Please clarify in the manuscript.
2.If point 1 is true and the scores to odd numbered questions are already presented in table 1 as reversed, then there should be a significant difference for Q10 (although you don’t present SD), however, your p value for Q10 is non significant at .149. Why?
3.Please present standard deviations in Table 1.
4.you state “…..The researcher selected the hypothesized mean of 3 based on a theoretical framework…..” Please explain further. I do not like the idea of a hypothesized mean? You are comparing musicians and non-musicians. Why bring a hypothesized mean into the mixture?
5.I suggest that you also calculate p values for musicians versus non-musicians for only social inclusion (odd numbered Q), then calculate p values for musicians versus non-musicians for only social exclusion (even numbered Q). If the latter is significant (in favor of musicians) then you can conclude that attributes of social exclusion are significantly mitigated by music classes whereas those of social inclusion are not. I don’t know what the implications of this are. I will leave that to you to find out, should you find a difference.
6.In addition, you could perform tests for signficance among musicians for social inclusion versus social exclusion questions. You could do the same for non-musicians for social inclusion versus social exclusion questions. Again, I don’t know what any difference will imply. I will leave that up to you to find out. The objective is to analyze your collected data thoroughly so as to be able to derive as many cogent conclusions (related to your title) as possible. Such analysis also adds depth to the manuscript.
7.Yet another statistic you could run is separate the sexes and run the same analyses for males and females separately. This will tell you if there is a sex-related difference in student belonging between musicians and non-musicians.
8.You could also run grade-specific musican and non-musician statistics……any other subsets you can think of.

“Harmony in Schools: The Role of Instrumental Ensembles in Fostering Student Belonging”
Manuscript Revisions

1. “I am having trouble understanding this statement : “…….To ensure consistency, the items within the social exclusion subscale require reverse scoring as they are negatively worded. This ensures that responses in which lower values are positive, are reversed so that higher values indicate higher levels of belonging when summed to score.”. How were these values 'reversed" ? subtracted from 4? are the values in table 1 for the odd numbered questions already

presented as reversed? for example, Q10, musicians actually scored 0.373 ? and non musicians actually scored 2.532 before reversing? Please clarify in the manuscript."

a. Where I addressed it: I addressed the reviewer's comments in the second paragraph under the "Results of the Independent Samples t-tests" subsection, under the "Results" section from pages 7-8.
b. How I addressed it: In my first draft of the manuscript, I did not explain how I reverse scored items from the social-exclusion subscale. Consequently, in my revised manuscript, I clarified explicitly how I reverse scored on page 22 and added another table before presenting my data, which was previously table 1, to help visualize the reverse scoring process. I explained the application of the reverse scoring formula on the social-exclusion subscale to establish that higher values indicated higher levels of belonging, regardless of the subscale. Therefore, if a student reported one on a social-exclusion item, I recorded it as a four. If they reported a two, I recorded a three. I also explicitly stated that I applied the reverse scoring formula before performing any statistical tests to ensure that my calculations remained consistent throughout the entire analysis process.

2. "If point 1 is true and the scores to odd numbered questions are already presented in table 1 as reversed, then there should be a significant difference for Q10 (although you don't present SD), however, your p value for Q10 is non significant at .149. Why?"
    a. Where I addressed it: I addressed the suggestion on Table 2, subsection under the "Results" section on page 8.
    b. How I addressed it: I addressed the suggestion on Table 2, subsection under the "Results" section on page 8. I consulted with my high school statistics teacher and performed the t-test again, finding that Q10 is non-significant and was suggested to present the standard deviation to explain for the statistic's non-significance.

3. "Please present standard deviations in Table 1."
    a. Where I addressed it: I addressed it on Table 2, under the "Results" section on page 8.
    b. How I addressed it: I addressed it on Table 2, under the "Results" section on page 8.

4. "you state "…..The researcher selected the hypothesized mean of 3 based on a theoretical framework….." Please explain further. I do not like the idea of a hypothesized mean? You are comparing musicians and non-musicians. Why bring a hypothesized mean into the mixture?"
    a. Where I addressed it: I addressed this concern with revisions throughout my "Results" Section and "Methodology" sections pages 6-10.
    b. How I addressed it: Upon receiving this feedback, I realized that calculating a hypothesized mean falls outside of the scope of my study and I deleted any references to it from pages 6-10. My manuscript now addresses one focus. The two samples of the study were musicians and non-musicians and thus, I focused on statistical analyses comparing those groups.

5. "I suggest that you also calculate p values for musicians versus non-musicians for only social inclusion (odd numbered Q), then calculate p values for musicians versus non-musicians for only social exclusion (even numbered Q). If the latter is significant (in favor of musicians) then you can conclude that attributes of social exclusion are significantly mitigated by music classes whereas those of social inclusion are not. I don't know what the implications of this are. I will leave that to you to find out, should you find a difference.

a. Where I addressed it: I addressed it on page 9 and added Table 3 “Independent Samples t-test Results for Social Acceptance and Exclusion Among Musicians and Non-Musicians”. I addressed the implications on page 12.
b. How I addressed it: How I addressed it: To address this, I aggregated the means and standard deviations for musicians and non-musicians for each subscale. I used the aggregated data instead of using the group means and standard deviations. I performed the independent samples t-test separately for social-exclusion and social-acceptance. These tests revealed no difference in responses to the social-acceptance subscale but a difference between musicians and non-musicians to the social-exclusion subscale. I addressed the implications on page 12.

6. “In addition, you could perform tests for significance among musicians for social inclusion versus social exclusion questions. You could do the same for non-musicians for social inclusion versus social exclusion questions. Again, I don’t know what any difference will imply. I will leave that up to you to find out. The objective is to analyze your collected data thoroughly so as to be able to derive as many cogent conclusions (related to your title) as possible. Such analysis also adds depth to the manuscript.”
    a. Where I addressed it: I addressed this recommendation in “Social-acceptance/Social-exclusion Analysis” under the “Results of the Independent Samples t-tests” subsection, under the “Results” section on page 9. I made revisions in the discussion accordingly on page 10-12.
    b. How I addressed it:To address this, I conducted the t-test twice to separately compare responses between musicians and nonmusicians to the social-acceptance and social-exclusion subscales. For musicians, I analyzed the difference between social-acceptance and social-exclusion questions. I did the same for non-musicians.

7. “Yet another statistic you could run is separate the sexes and run the same analyses for males and females separately. This will tell you if there is a sex-related difference in student belonging between musicians and non-musicians.”
    a. Where I addressed it: I addressed this in “Direction for Future Research” on page 13.
    b. How I addressed it: I was unable to address this because I did not collect data on biological sex. My survey only collected data on age, grade level, and music ensemble participation status. I mention that future research could collect this data.

8. “You could also run grade-specific musician and non-musician statistics……any other subsets you can think of.”
    a. Where I addressed this: I addressed this in “Grade Level Comparison Analysis” under the “Results of the Independent Samples t-tests” subsection, under the “Results” section from pages 9-10.
    b. How I addressed this: I performed another t-test this time using aggregated data separating freshmen, sophomores, juniors, and seniors for each test. In this t-test, I determined whether ensemble participation would influence belonging within a specific grade level. I found that there was no statistical difference in the responses of musicians and non-musicians in any grade levels. I did not perform a t-test focusing on age as it corresponds closely with grade level.

Thank you for addressing my comments. One minor revision is necessary.

Before I can accept this manuscript, I wil need you to present the results of the overall t-test between musicians and non-musicians (fill out the highlighted square before Table 2.) in the attached document.

Addressed: The results of the independent samples t-test suggested that while there were no significant overall differences between musicians and non-musicians across all the 10 questions ($p = 0.130$), the test did yield nuanced results when individual questions were compared. The p-values for eight out of the ten questions on the SBS resulted in values > 0.05).

Accepted