Dear authors.

First of all, I would like to thank you for the proposal you have prepared and sent, as it seems to me to be of great interest and relevance. Of course, it is necessary to praise the effort and work done.

In addition, I would like to make some contributions of little significance, but which can be easily dealt with and which are the following:

> Authors' Response: Thank you for the review for taking the time to provide a constructive review. Our team thinks that all contributions from an external lens are significant! We have responded to all recommendations below. Thank you again!

- 2.1. Study Participants and Study Design: Why did the students in the sample have to be involved in a sports season? I think this is something interesting that should be explained; with regard to the prior sending of the questionnaire, so that it could be reviewed by the families, I have the feeling that this could interfere with the results obtained, since it cannot be guaranteed that the families have not 'worked' with their children on the questionnaires. Why is this done? For what purpose?

> Authors' Response: The students had to be in a current season as to limit the number of previous athletes. We thought that with active student-athletes, we could make a stronger case for the validation process. We certainly agree that parents/guardians could have 'worked' with their child before the data collection. We had to send all materials home as part of the school district research IRB process. We have added a statement in the manuscript to make this clear. "The school district's research office mandated the process of sending materials home."

- 2.2. Instrumentation: Why do the number of possible responses vary in the Likert scales? Would it not be more appropriate to try to have the same number of possible responses when dealing with the data as a whole (physical literacy); When talking about the non-inclusion of the score on the fitness level, this is repetitive (rows 96-99).

> Authors' Response: The questionnaire (PLAY*self*) was used exactly as it was validated in other populations to allow for direct comparisons. That said, we do agree that using the same response options is a more favorable process. We removed the initial reference to the physical literacy item to allow for a comment by another reviewer to be maintained.

- 2.3. Data collection: I do not quite understand the reason why there are two data collection moments for students who are participating in sports from the same starting point. The longer time of practice could affect the results...

> Authors' Response: From rereading this section, we understand the confusion. Participants only completed the data collection once. We, the research team, collected

data over the course of two sports seasons. We have added language to clarify this distinction.

- 2.4. Data analysis: Why were the sections recoded with these scores, 600, 900 and 1200?

  Author's Response: First, the Likert item was recorded. Then, each of the subsections was summed individually: 600 for the environmental, 1,200 for the self-description, and 900 for the relative rankings of literacy.