**Addressing comments from Reviewer 1**

*The title doesn't reflect the overall outcomes of the study*

We changed the title to “Commonly-used subjective effort scales do not predict directly-measured physical workloads and fatigue in Hispanic farmworkers”.

*The material is interesting but the overall manuscript not up to the level of publication*

Thank you for your feedback. We made changes according to your suggestions.

*What is the study biggest contribution?*

the end of the “relevance to industry”, we added that “The results of this study informed ergonomics practitioners that subjective effort surveys like Borg scales used in exercises among healthy young men might not be able to accurately quantify effort or fatigue among the Hispanic farmworker population.”

*Introduction is very brief*

*Relevant literature should be presented more deeply in order to support the research problem*

Our initial introduction has 1,259 words in total. We understand that you mean by only the context of Hispanic migrant farmworkers in North America, which has only 165 words, but excluding the existing methods in ergonomic assessments and the objectives of the study. We eliminated sub-headings in the introduction to focus more on the context by adding more details about the study population. The new introduction covers 454 words.

*Suppose to be put under Methodology*

We divided and moved the section “1.2 direct measurement and subjective self-reported questions in ergonomic assessment” to the different parts in the new sections “2.2 ergonomic exposure parameter selection”.

*Aim of the study is not clear*

*Why this study was conducted*

*Why only certain target group was selected*

We revised the section “1.3 objectives of this study” to “determine whether the subjective measures of the exertion of the overall body and at the local body parts could be used to predict the workload and fatigue directly measured with sensors.” This is the reason why this study was conducted. To answer why only certain target group was selected, it is mentioned in the section “2.1 participants” that directly follows the introduction.

*Why only 24 Hispanic male?*

In the section “2.1 participants”, we added that “Participants were selected to represent the major characteristics of Hispanic male farmworkers in an industrialized orchard in the Washington State as well as in the United States in general. The number of participants were determined to yield statistical powers and data saturation among this population.” We also explained later in the results section that “Despite our effort to include both genders in the participant recruitment, there were no recruit female volunteers to participate in the study.”

*More precision is necessary regarding the sampling strategy and access to the target population*

In the section “2.1 participants”, we also added that “The research team members include male and female local researchers of Latin American origin, who understand the cultural dynamics within the population. The orchard owners and managers were initially contacted for permission for conducting a research at their site. Then all participants were recruited on the day prior to the first day of data collection.”

*Justification is needed why this measurement is important in this study*

We moved part of the text on the existing techniques (previously in the section 1.2) in the introduction to justify the importance of each measurement.

*Why the level was set at 0.95?*

We clarified that we have “95% confidence to reject the null hypothesis or 5% probability of making an error of rejecting the null hypothesis while it was true”.

*Need extra explanation on this statistical analysis*

After the phrase “All the statistical analysis was conducted using R programming language”, we explained more that “Particularly, the ‘lm’ function was used to run a linear regression to find correlation coefficients of the relationship between parameters and the ‘anova’ function was used for extracting the significant levels of the effect. All the graphics were made using the ‘ggplot2’ package.” We are willing to provide the codes we use upon request or a link to GitHub repository if it would be an interest for audience.

*Better to start with the results on sociodemographic characteristics of the respondents*

Thank you for your suggestion. We added this into the section 3.1 that “Twenty-four Hispanic male apple pickers participated in the study.” and that “The participants’ ages were on average 28.4 years (range 18-47 years). Their experience as farmworkers harvesting tree fruits in the United States were on average 3.4 years (range 1-14 years).”

*This p-value is for what?*

For your question at the text “Based on the Shapiro-Wilk test, initially % HRR data was not normally distributed (p = 0.013).” we changed the text to “The calculated % HRR was not normally distributed according to the p-value of 0.013 in the Shapiro-Wilk test for normality.”

*Not clear on the p-value*

For your comment at the text “Overall, the % HRRs among the Ladder group were higher than 234 the % HRR among the Ground and Platform workers (Tukey HSD p-value = 0.0001 and 0.009, respectively).” we changed the text in the parenthesis to “(Tukey HSD p-value = 0.0001 for the comparison between the Ladder and Ground workers, and 0.009 for the comparison between the Ladder and Platform workers)”

*Need explanation on why the analysis of your data in the way that you do, or why your statistical analysis are appropriate This should be clarified*

For the Figure 11, the section “3.5 association between muscle fatigue and subjective local discomfort”, we explained more that “there was no correlation between the slope in the equation (iv), i.e., the increase or decrease of EMG mean power frequency over time, and the Borg CR10 difference between the beginning and the end of the work shift.”

*There is a lot of results but discussions very minimal*

Originally, there are 960 words in the discussion as compared to 979 words in the results. Nevertheless, we explained more in the section “4.1 interpretations and implications from negative or no correlation”. In the second paragraph, we added that “This finding on stronger correlations suggested that the Omni RPE with the pictures of apple harvesters may be more useful than the Borg RPE with only verbal anchors on the scale from 6 to 20, which might not be comprehensible for farmworkers.” In the third paragraph, we added that “the Omni RPE with the pictures of apple harvesters carrying an apple bag in different phases of tiredness (Figure 1) may still not be robust and should be improved.” In the forth paragraph, we added that “Thus, the use of Borg CR-10 to assess local body fatigue is not recommended for this population.”

*A stronger discussion of implications for future research and potential intervention work is needed*

We added to the section “4.2 comparisons to previous studies and directions for future work” as a recommendation for future work that “Alternatively, we propose a combination of farmworkers carrying apple bag and faces representing emotions in the Omni RPE.”

*Do you anticipate all the limitations before conduct this research?*

Yes, we anticipated these limitations as the nature of field research. We added to the section “4.3 study limitations” that “Above all, these limitations could be anticipated prior to the study. On one hand, part of the systematic bias during the data collection were not able to fully eliminated due to the nature of the fieldwork where researchers were present and indirectly influencing how the farmworkers behave, and where quiet sitting for resting heart rate measurement was infeasible. On the other hand, the data loss due to EMG connection was mitigated by the development of the algorithms to remove errors and retain only meaningful signals.”

*This is just repeating back the results*

*A final conclusion should be added in order to answer the initial research questions. Otherwise, the reader is left with an incomplete idea of the article*

We revised the final conclusion with summary that “All things considered, for the Hispanic farmworkers population, direct measures of ergonomic exposures could not be replaced by subjective measures. If only the subjective measures are possible in the field assessment, the results will have to be interpreted with cautions. If necessary, Omni RPE, i.e., the scales accompanied with pictures, would be a better option than Borg RPE or Borg CR-10 which only have verbal anchors and were developed during different activities performed by different population.”

*Why leave it blank?*

We moved the information on acknowledgement and the participants to complete these areas: funding, institutional review board statement, informed consent statement. We also added that we did not have any conflicts of interest. Meanwhile, the author contributions were specified in the submission page.