**RE: manuscript revision and resubmission for IJERPH-1971752**

**Commonly-used subjective effort scales may not predict directly-measured physical workloads and fatigue in Hispanic farmworkers**

**Point-to-point replies to the Editors:**

The authors applied both objective and subjective instruments on a rather small group of Hispanic farm workers in the United States. In the methods section they write: “The number of participants were determined to yield statistical powers and data saturation among this population” (line 162). They do not really explain their power calculation. But in the discussion they write: “it is worth remarking that the reason might be due to the fact that the stratification reduced the sample size and, consequently, there was not enough power to detect a significant correlation in each group.” (line 447). Thus, it seems, that they have not conducted a formal power analysis at all.

**REPLY: Since we conducted a convenient sample, we removed the phrases that *“The number of participants were determined to yield statistical powers and data saturation among this population”* And that *“Notwithstanding, it is worth remarking that the reason might be due to the fact that the stratification reduced the sample size and, consequently, there was not enough power to detect a significant correlation in each group”.***

They explain that they believe that Borg RPE and Omni RPE are proxies of physical effort. They measure physical effort by measuring heart rate (HR) (and by applying an individual calibration to HR). (“The Borg RPE was developed in the context of cardiovascular treadmill exercise and intended to be highly correlated with heart rate; that is, the RPE score multiplied by 10 generally represents a person’s actual heart rate in beats per minutes”, line 225.) They never explain the details of the Borg scale. But for the Omni scale they provide an example (figure 1). I am not sure if that scale only consists of this single figure, or if this is just one example. But this figure does not ask about “physical effort”, but about “tiredness”. “Tiredness” is a complex phenomenon including physical and mental parts. I can be tired after hard physical work, or after difficult mental work, or, after a long and boring day, without any hard work at all. When I get tired, my body usually relaxes, and HR will go down. So, how should asking about tiredness be a proxy about physical effort? The authors should really reconsider their wording.

**REPLY: We initially assumed that the readers would already know the Borg scales and we had not needed to explicitly explain the scale. Also, the full explanations and examples were already illustrated in the previous paper as cited. Per your comments, we added the description of the Borg scales, both Borg RPE and Borg CR10 in the methodology section that *“Borg RPE and Omni RPE scales were used as subjective measures of overall effort exerted by the workers, and Borg CR10 scale was used as a subjective measure of local discomfort, particularly the tiredness (cansado in Spanish) they felt at the moment. Borg RPE scale, ranged from 6 to 20, was accompanied by verbal anchors from “no exertion” at 6 to “maximal exertion” at 20. Borg CR10 was also accompanied by verbal anchors from “not tired” at 0 to “severely tired” at 10.”***

There are statements that are wrong or make no sense, and there are sentences in rather poor English language. Just a few examples: Content: “As shown in Figure 4, relative to the Borg RPE ratings collected from the beginning of the work shift (T1), there were no difference between the ratings collected at the beginning of the work shift (T1) and after 30 minutes of rest (T2).” (line 355): T1 was not done at “from the beginning of the work shift”, but at the beginning of the break, at least according to the description in the methods!

**REPLY: Per your comments on wordings, we revised the phrase from *“As shown in Figure 4, relative to the Borg RPE ratings collected at the beginning of the work break (T1), there were no difference between the ratings collected at the beginning of the work shift (T1) and after 30 minutes of rest (T2).”* To *“******As shown in Figure 4, relative to the Borg RPE ratings collected at the beginning of the work break (T0), there were no difference between the ratings collected at the beginning of the first break (T1) and after 30 minutes of rest (T2).”***

“The Omni RPE from the beginning of the work shift were also greater at the end of the work shift (T3) as compared to the other time (p-value < 0.0001)” (line 364). What does this mean? The RPE at the beginning is greater at the end? This does not make any sense at all!

**REPLY: Also, the phrase *“The Omni RPE from the beginning of the work shift were also greater at the end of the work shift (T3) as compared to the other time”* were revised to *“The difference in Omni RPE from the beginning …”***

“Without adjusting for neither work period (T1, T2 and T3) nor harvesting method (Ground, Ladder and Platform), the correlation coefficient between the % HRR and the Borg RPE was insignificant (p-value = 0.23). Similarly, the correlation coefficient between the % HRR and the Omni RPE was positive (p-value = 0.006).” What is “similar” between the correlation with OMNI RPE and that with Borg RPE? There was no correlation with the latter and a significant one with the former!

**REPLY: Moreover, we changed the conjunction for the two sentences *“Without adjusting for neither work period (T1, T2 and T3) nor harvesting method (Ground, Ladder and Platform), the correlation coefficient between the % HRR and the Borg RPE was insignificant (p-value = 0.23). Similarly, the correlation coefficient between the % HRR and the Omni RPE was positive (p-value = 0.006).”* from “similarly” to “in contrast”.**

“There were negative significant correlations between the direct and subjective measures of overall effort when the analysis adjusted for the harvesting method and the time of measurement. That is, both Borg RPE and Omni RPE may be used to predict the outcomes of % HRR.” (line 438). But this statement makes no sense. Borg RPE has been designed to be directly interpreted as HR. Therefore, it should be positively associated with HR, not negatively.

**REPLY: In the discussion, we changed the interpretation which you mentioned as making no sense *“There were negative significant correlations between the direct and subjective measures of overall effort when the analysis adjusted for the harvesting method and the time of measurement. That is, both Borg RPE and Omni RPE may be used to predict the outcomes of % HRR.*” The new paragraph is written as *“This study found significant correlations between the direct and subjective measures of overall effort when the analysis adjusted for the harvesting method and the time of measurement; however, the directions of corrections were contradictory for Borg RPE and Omni RPE. The negative correlations between Borg RPE and % HRR suggested that the Borg RPE may not be useful as subjective measures for this population whereas the positive correlations between Omni RPE and % HRR suggested that the Omni RPE could predict the outcomes of % HRR*.”**

Language: “Field ergonomic assessment can be directly-measured with sensors and subjectively participant’s self-reported ratings.” (line 63). Wrong wording! Assessments cannot be measured!

**REPLY: For another sentence that were incorrect “*Field ergonomic assessment can be directly-measured with sensors and subjectively participant’s self-reported ratings*”, we revised that “*Field ergonomic assessment can be conducted by direct measurement with sensors and by questionnaires of subjective self-reported ratings*”.**

“While the Borg scales has a verbal description for each numeric value,” (line 237). Scales (plural!) have!

**REPLY: We also corrected the grammar from “while the Borg scales has a verbal description …” to “… scales have …”**

I do not understand the real aim of the study. If the authors wanted to test if the subjective RPE scales are able to measure the physical effort of the work, why then stratify by physical effort?

**REPLY: It is already mentioned in the last paragraph of the introduction that this paper’s goals are to “determine 1.) the association between metabolic, i.e. cardiovascular, load and the overall Omni and RPE scales, and 2.) the association between localized muscle fatigue measured through electromyography and the local Borg CR-10 scales.”**

On the whole, the authors make multiple comparisons (stratified by task and by time), and are surprised when some of these comparisons turn out to be “significant”, even if the direction of the association is contrary to their stated expectations. They do not examine the strength of the association but still they claim that “Borg RPE and Omni RPE may be used to predict the outcomes of % HRR.”

**REPLY: As mentioned in the reply to your previous comments. The interpretation that “Borg RPE and Omni RPE may be used to predict the outcomes of % HRR” was removed.**

The findings per se, regarding differences between tasks and changes over time, are interesting for each scale or instrument. But the comparison of the instruments and the wild and crude interpretation make no sense at all!

**REPLY: This paper did not compare the instruments.**