April 28, 2022

Jamilah Foucher and Benoît Bardy
EuroMov Digital Health in Motion
University of Montpellier
700 Av. Pic Saint Loup
34090 Montpellier, France

To

Aasia Khanum
Associate Editor, IEEE Access

Dear Dr. Khanum,

Thank you for your time in reviewing our paper submission. We are now preparing our revision incorporating the reviewers' recommendations. We would normally respond to the reviewer's comments without contacting the Editor, however it is unclear how to incorporate the major concern of Reviewer 2. We believe that there are inaccuracies in Reviewer 2's main concern (and this may be due to some unclear sections of our text), and we are worried that if we responded to these inaccuracies normally in our paper revision, we may have less opportunity for publication success than if we discussed these concerns with you (before submitting the revision) to help us to reconfirm and interpret how to best respond.

Reviewer 2's major concern implies that she/he believes that we used the physical disorientation and motion detection performance measures as features for an ML model, because she/he recommends using Deep Learning to improve the relationship between the measures. However, in the paper we state that we only use transformations of joystick sensor data as features for ML and DL models (see section III.ANALYSIS D. CLASSIFICATION MODELS, FEATURE & LABEL CREATION), and the motion detection performance measures are model labels; physical disorientation measures are not used for modeling. This paper only states that we investigated physical disorientation as a potential feature/marker for future ML/DL modeling using statistical analysis.

We used a well-known conventional questionnaire to measure a physical disorientation score called SSQ disorientation; we used a questionnaire to measure physical disorientation because questionnaires are commonly used in the aviation industry for monitoring SD. The physical disorientation measure consisted of two scalar values per participant, obtained by a questionnaire before and after the whole-body experiments. Motion detection performance categories from 1 to 10 were assigned per trial per participant, as shown in Figure 4. Across participants we calculated the mean of trial counts for two negative and positive physical disorientation questionnaire categories and four motion detection categories. Using statistical analysis, we compared the eight coupled physical disorientation and motion detection categories, and no statistical significance was found.

Additionally, Reviewer 2 says that our 2-layer Neural Network (NN) model that is a Deep Learning model in the scikit-learn python toolkit is not a sufficient Deep Learning model. However, he/she does not state why the model is insufficient for their liking or how we should improve it so it will be acceptable to them in the future revision. From our professional perspective, considering our data size, a 2 to 4 layer NN is typically used; we do not have Big Data for this dataset so a many layer 'deep' NN model was not needed.

What we could propose is to (1) mention this discussion and our feeling about Reviewer 2's main comment in the authors' response accompanying the revision. In the revised version, we could (2) be clearer about the relation between motion detection and physical disorientation as it is in our analysis to avoid the current misunderstanding, (3) dig deeper in the discussion about how to further investigate this relation through deep learning models in future research.

Thank you for your time in hearing our concerns. We thought that the revision process would be more fair if we explained our concerns and asked for more clarification on how to best respond; we took the concerns of Reviewer 2 very seriously and did not want any inaccuracies to jeopardize our chances of success. Any advice and your professional opinion on how to respond would be greatly appreciated.

Sincerely,

Jamilah Foucher & Benoît Bardy

Below are copied excerpt from both reviewer 2 and the paper about the physical disorientation and motion detection result.

1) Reviewer 2's main concern: Given only traditional ML model was applied, the conclusion "no significant relationship between physical disorientation and motion detection was found." might be true. However, what about some deep learning model? I don't think the 2-layer NN mentioned in the paper should be identified as a deep model. I think the discussion of deep learning models will be helpful. If not, deep-diving into the reason why physical disorientation and motion detection is not strongly related is necessary.

2) Citations from the paper supporting the description of the physical disorientation analysis: On page 9, we describe the physical disorientation questionnaire measure and how it was compared with motion detection performance. "Physical disorientation for accurate and non-accurate motion detection performers were compared, to quantify whether physical disorientation report could also be a marker for SD, like the perceptual joystick. Again, Wilcoxon signed-rank or rank-sum non-parametric distribution tests were used to evaluate comparisons, as the KS test only found non-parametric distributions. The mentioned statistical p-value reporting convention was used."

On page 14 under subsection 'D. Motion Detection Performance and Physical Disorientation' of the submitted manuscript, we dedicate a short sub-section to explain the statistical comparison between physical disorientation and motion detection performance.

Finally on page 16 in the last paragraph, we discuss why there may not have been a significant relationship between physical disorientation and motion detection. We state that "More sample points regarding physical disorientation are needed during the experiment, instead of a sample before and after the task, in order to determine if physical disorientation is correlated with motion detection....We do not claim that questionnaire methods can not quantify SD, however before and after questionnaire samples may not produce enough data to find statistically significant correlations with other SD measures especially when population sample size is small. A physiological sampling measure that implies physical discomfort, with a sampling rate comparable to that of the joystick, such as EEG, NIRS, heart rate, or electrodermal activity, could provide more insight into correlations with physical and perceptual disorientation."