

Author's response: *General response to both reviewers and the recommender: Please accept our sincerest thanks for all your invested time and useful feedback.*

Decision on your Stage 1 Registered Report: Revisions Required

Dear Marcel Martončík and colleagues,

I have now received the review of your revised PCI-RR submission from two reviewers who also reviewed your initial submission. As you will see, both reviewers (and I) think that most previous comments have been addressed very satisfactorily in the revised manuscript. The results of the new pilot studies are also very informative, and certainly addressed some of the potential concerns with the measurements. Both reviewers also provided some new or remaining comments. All comments are relatively minor, and aim at further improving the readability of the manuscript. Addressing these comments will very likely result in an IPA.

Although you are now not planning a direct statistical comparison between different titles, the issue that the ranks are not comparable still remains when you interpret the results across the different games. This is an important point, and Reviewer 1 (Dr. Bonny) suggested a potential metric of long-term skilled performance that may be more comparable across the games. Please carefully consider whether such a metric is available with the three games that you have selected. If yes, I would recommend including such a measurement, even only for exploratory purposes. If such a metric does not exist, or cannot be easily obtained, this point will definitely need to be mentioned when interpreting the results after data collection.

Author's response: *Response included in Dr. Bonny comment n. 2.*

Reviewer 1 suggested a potential distinction between competitive versus non-competitive play. While this distinction may indeed be informative, I feel this is a difficult decision to make at this point. The results of your pilot study show that the reliability of the practice measure is okay, supporting the use of this measurement without further changes. The competitive vs. non-competitive distinction may be addressed in follow-up research. In any case, I agree that psychometric analyses on the questionnaire once the data are in are recommended, as it may further increase the impact of this work.

Author's response: *Response included in Dr. Bonny comment n. 1.*

Reviewer 2 (Dr. Behnke) gave some suggestions on using sub-sections in the Introduction, and I agree that this can further improve the structure and readability of the introduction.

The effect size for the first entry in Table 1 is missing. Most entries in the 'Notes' column succinctly summarised (the direction of) the main finding of each study, which I found very informative. However, some entries did not contain this information. For instance, for Mora-Cantalops & Sicilia (2018) - competence and presence (immersion), the 'Notes' column only mentioned the instrument used. Could you also briefly mention how these two factors relate to a player's rank (e.g., in which direction)? Other entries that may also benefit from adding such info includes the rows from Li et al. (2020) till Trotter et al. (2021).

Author's response: *Thank you for pointing this out. Table 1 was not complete indeed. All the missing information has been added.*

The information in Table 2 (descriptive data from Pilot 2) does not seem to be crucial for the introduction. Perhaps it can be moved into the appendix, to make the introduction more compact?

Author's response: *We have moved Table 2 into the Appendix 2 as suggested also by Dr. Behnke.*

I agree with reviewer 2 that table 3 is difficult to read, possibly because it combined multiple sources of information (i.e., statistically significant results in bold, predictions of the current study highlighted in purple or green, and the smallest effect size of interest and its interpretation for both titles). Personally, I think table 3 will be easier to understand if it would only show the statistical results from Pilot 2. You may consider making a similar table separately for the hypotheses, with the cells in different colours to distinguish the different predictions (null vs. alternative). This table may serve to replace much text in which you now spell out the hypotheses. (A similarly-structured table may also be used to show the results once the data are in, so that the pilot results, the predictions and the confirmatory results may be easily compared.) For the interpretation of the SESOI, I think presenting them in Appendix 5 seems sufficient (see my next comment).

Author's response: *We have also moved Table 3 into the Appendix 2 as suggested also by Dr. Behnke while keeping a more simplified Table in the Results section. The table has some columns populated with exploratory data, while blank columns are left for confirmatory data.*

The rationale behind the smallest effect of interest in Appendix 5 was very nuanced and thoughtful. While I like this information a lot (it certainly made it much more concrete for me what a certain effect size means!), I feel including all this information in the introduction would interrupt the overall flow. Thus, referring the readers to Appendix 5 (as you currently do) seems like a good solution to me.

Very minor point: It will be useful to have a final check of the whole document once you are ready with the edits - sometimes there are two spaces instead of one between two words.

Author's response: *Thank you for your kind words! We are happy that Appendix 5 makes it easier to understand our logic behind setting the SESOI. We have also checked the whole document one more time.*

Kind regards,

Zhang Chen

Reviews

Reviewed by Justin Bonny, 03 Mar 2023 21:45

Overall Reviewer Response

I commend the authors for the improvements throughout the manuscript. Most of my prior comments have been addressed save for a few clarifications on the practice measures and planned analyses.

Practice Questionnaire

This in and of itself could be a valuable contribution to esports research. Having a measure that can assess different facets of esports practice would be useful for subsequent studies. That being said, the 'practice' question, "Routinely playing the game (ranked mode, non-ranked mode, with or without friends, etc.)", seems to span across multiple types of play. A prior study distinguished competitive (ranked matches) and non-competitive (non-ranked

matches) video game play with some evidence of differences in connections between the two with psychological traits (e.g., Bonny et al. 2020 Intelligence). It may be worth considering taking a similar approach here, splitting the question into two different ones that distinguish between competitive and non-competitive play. Although your reliability statistics in Pilot 4 suggests this may not be necessary, doing so could make a bigger impact on the future use of the instrument in esports research.

I do recommend that psychometric analyses be provided in the study for the practice questionnaire items. Specifically, including a factor analysis, in addition to reliability statistics, would provide further evidence of the performance of the measure.

Author's response: *We agree that it could be worth further distinguishing naive practice (NP4), thus we have split this item into two: the play of ranked (NP4) and non-ranked games (NP5):*

“The final two activities do not directly aim at improving esports rank/skills. Please do not include gaming hours that you have already reported in previous questions.

- *Routinely playing the game in ranked mode (alone or with others).*
- *Routinely playing the game in non-ranked mode (alone or with others).”*

We are aware of the fact that proper exploration of the psychometric properties of our instrument requires a sufficiently sized sample, thus we plan to provide (as a Supplementary file) additional information about reliability and validity (EFA/CFA) conducted on confirmatory data from all three games.

Planned Analyses

The number of levels in the dependent variable of rank is still different across each esports game, with 27 in LoL, 18 in CSGO, and 10 in Fortnite. It is still not clear whether the differences between each rank are commiserate across the esports games. It does appear that direct comparisons across esports game are not included in the analysis plan. But it is still worth considering how these differences in skill rank could impact conclusions about cross-title differences. For example, if intelligence is a significant predictor for LoL but not Fortnite, is that due to intelligence being more important for LoL or differences in the ranks between LoL and Fortnite? If there are any other metrics of long-term skilled performance that are common across the games, like MMR for Dota 2 which is modeled on ELO in chess, collecting these could be helpful. Although there are still limitations (e.g., games may use ELO-like ratings, but are calculated differently), this could provide corroborating support for the use of rank as a performance metric and additional support for cross-game comparisons.

I encourage the authors to provide the R-scripts via OSF to provide further information for replicating analyses in subsequent research.

Author's response: *The number of scale points indeed affects the magnitude of the correlation, however, previous studies have shown that the potential change is minor (Norman, 2010 or Rhemtulla et al., 2012 on a similar matter) and even smaller when both of the variables have more than 10 scale points (Martin, 1978), which is the case with all three games included in our study. Citing the work of Martin (1978), „a review of the correlations calculated from sets of scale points of 10 or more shows fairly good agreement with the "true" r. For example, with a "true" r of .9, the 20 points-10 points r is .884, the 16 points-10 points r is .880, the 14 points-10 points r is .880, and the 12 points-10 points r is .875.“ In our*

view, a number of ranks between the three titles should not significantly affect the magnitude of correlation and even less likely the associated *p*-value. However, there is a different thing that could affect the comparability of results between different titles, namely the distribution of ranks. The shape of the distribution of ranks affects the relationships between our DV and predictors. This is not a psychometric artifact but a property of the researched phenomenon. From the publicly available data (e.g., <https://www.ggrecon.com/guides/league-of-legends-rank-distribution/>; <https://www.ggrecon.com/guides/csgo-rank-distribution/>; note that rank distribution for Fortnite is not available), it is evident that distributions differ a lot: for LoL it is highly positively skewed, while for CSGO it resembles a normal distribution. Thus, we consider the distribution of ranks as a more important factor affecting results than the actual number of ranks for a specific title. This, of course, will be properly discussed in the Discussion section.

Before submitting this research proposal as RR we have piloted measurement of ranking using MMR/ ELO/ Hype points. However, a large proportion of players reported that they do not know the precise number of MMR/ELO/Hype points but only their rank (e.g. Gold, Platinum, etc.). It is also true that some of these metrics are either not easily accessible (like ELO in CSGO) or not available for all players (like Hype points in Fortnite are available only for the top 100 players).

R script for confirmatory analyses is on the OSF in the Data folder (<https://osf.io/vbkns/>) and R scripts for exploratory analyses are in the Pilot 2 folder (<https://osf.io/qbd7x/>).

References:

Martin, W. S. (1978). Effects of Scaling on the Correlation Coefficient: Additional Considerations. *Journal of Marketing Research*, 15(2), 304–308.

<https://doi.org/10.1177/002224377801500219>

Norman, G. (2010). Likert scales, levels of measurement and the “laws” of statistics. *Advances in Health Sciences Education*, 15(5), 625–632. <https://doi.org/10.1007/s10459-010-9222-y>

Rhemtulla, M., Brosseau-Liard, P. É., & Savalei, V. (2012). When can categorical variables be treated as continuous? A comparison of robust continuous and categorical SEM estimation methods under suboptimal conditions. *Psychological Methods*, 17(3), 354–373. <https://doi.org/10.1037/a0029315>

Reviewed by Maciej Behnke

The authors did a very good job addressing my concerns, and I enjoyed reading the revised manuscript. Please find my comments below.

1. Intro structure. I’m satisfied with the current shape of the introduction, but I want to suggest some things. These are not requests but only something to consider. We don’t want to create another example for a popular meme that compares the original paper with the revised one.

1.1 I would add the Pilot Studies section after the Literature on Esports Expertise section. Next, I would add Present Study section after the Pilot Studies.

Author's response: *Thank you for the suggestion, we agree and have restructured the Introduction section.*

1.2 Pilot Studies.

- I like the description of Pilot 1.

- I would prefer to have a longer description of Pilot 2 here rather than Tables 2 and 3. I personally prefer the content of appendix 5 than Tables 2 and 3. If I understand it correctly, the pilot studies provide the rationale for some decisions included in the present study, so I would prefer abstract-type paragraphs rather than detailed tables in the introduction.

Ps. The authors wrote: “For the data and detailed results, see Appendix 2 (<https://osf.io/qbd7x/>)”. Appendix 2 doesn’t include the detailed results – they are in the OSF folder –, thus, the sentence can be misleading.

- I would recommend adding the description of Plot 3 here to show the completeness of pilot studies.

Author's response: *Thank you, the sentence related to Appendix 2 has been corrected. We have also transferred Tables 2 and 3 to Appendix 2 and instead we have extended the description of Pilot 2 and have also included descriptions of Pilot 3 and 4.*

1.3 Games description – I would move this section to methods

Author's response: *Moved.*

However, If the authors insist on keeping Tables 2 and 3, I would recommend adding the phrase “daily” to practice (hours), deliberate practice (hours), and physical training (minutes). Furthermore, Table 3 is extremely hard to read, but maybe only for me.

Author's response: *We have moved both tables to Appendix 2 and added “daily” to practice and training. We have kept Table 3 in Appendix 2 in its present form, while still believing it could be informative - containing only necessary information (explanation of SESOI, B, beta, CI) with sign. predictors highlighted with colors. However, we have included a simplified version of Table 3 in the Results section that will be supplemented with confirmatory data.*

2. Independent variables.

2.1 Naive or General practice. Please, be consistent with the used labels. In the hypotheses, you used “naive”, whereas in the last sentence of the first paragraph of the independent variables section, you used “general practice”.

Author's response: *Corrected - changed from general to naive practice.*

2.2 Deliberate Esports Practice (DEP) Instrument. I needed help understanding the measure as presented in the manuscript, so I let myself create the Table presenting it. If keeping it in the form of the table is only my preference, ignore it.

Author's response: *This is definitely a better way of presenting DEP items. We have added Table 2 in the Independent variables section.*

Instrument	Table XXX. Deliberate Esports Practice (DEP)
------------	--

Introduction to items	Item Content	Category (code)
<p>The first two activities require focused attention and directly aim at improving esports rank/skills.</p>	<p>Learning alone (from guides, videos, streams, replays, etc.)? This does not include playing.</p>	<p>Deliberate Practice (DP1)</p>
	<p>Learning with others (getting feedback from teammates or coaches, team discussions, etc.)? This does not include playing.</p>	<p>Deliberate Practice (DP2)</p>
<p>The next three activities do not directly aim at improving esports rank/skills.</p>	<p>Physical practice (gym, running, etc.)?</p>	<p>Naive Practice (NP1)</p>
	<p>Mental practice that is not playing (meditation, breathing exercise, etc.)?</p>	<p>Naive Practice (NP2)</p>
	<p>Relaxing esports activities that are not playing (watching streams, discussing the game, etc.)</p>	<p>Naive Practice (NP3)</p>
<p>The last activities specifically concern playing esports game(s). The first two require focused</p>	<p>Playing with coaches, team, or other experts (with tactical communication, reflection, etc.)</p>	<p>Deliberate Practice (DP3)</p>

attention and directly aim at improving esports rank/skills.		
	Playing the game alone (practicing aim or last-hit, game scenarios/matchups, etc.)?	Deliberate Practice (DP4)
The final activity does not directly aim at improving esports rank/skills.	Routinely playing the game (ranked mode, non-ranked mode, with or without friends, etc.). Please, exclude the hours you reported earlier.	Naive Practice (NP4)
<p>Note. The questionnaire starts with the question: During the past 12 months of playing [GAME NAME], how many hours per week did you spend on the following activities?</p>		

3. The full survey. I'm not an expert in legal issues, but can we share the content of the items from the original questionnaires (e.g., Short Grit Scale, Duckworth & Quinn, 2009)? I just want to draw attention to this issue.

Author's response: *We agree, we did not realize the original article was not OA. We have removed the items.*

4. Design and analysis plan. E1 & E2: For clarity, I would move "persistence and intelligence" to the same place - both at the end or beginning.

Author's response: *Thank you. Corrected.*