

Dear Prof Chambers and the Managing Board,

Thank you for the opportunity to submit a revised version of this stage 1 manuscript to *PCI Registered Reports*.

I include the complete text of Prof Chambers' recommender comments and the two reviews by Prof Etchells and Dr Sauer in *black italics*; my point-by-point responses in *purple*; and amended or newly added text in *blue* below.

In addition, I note that the German age rating system (USK) has announced during the review process of revision 1 of stage 1 that it will also start to attach an informational/warning label for loot boxes, which is similar to the ESRB/PEGI label, starting with new submissions from 1 January 2023. Not enough time would have passed for compliance with this new label to be empirically assessable. The USK label also will not be applied retroactively, so its forthcoming portfolio of games marked with its label will not be comparable to the ESRB's and PEGI's portfolios (which have developed in the past 2.5 years). In my opinion, the proposed study is not affected and can proceed as planned. I intend to detail this in the limitation section following data collection. However, I had to change a few factual assertions regarding the USK in the introduction section of my manuscript as those have since changed (all changes are tracked as usual). I do not expect these to be controversial.

Leon Y. Xiao

--

Recommender Comments by Prof Chris Chambers, 16 Dec 2022 16:00

Minor Revision

Thanks for the careful attention to the previous comments. Both reviewers returned to evaluate your revised manuscript, and I'm happy to say that both are broadly satisfied. You will see some remaining clarifications to address in the review from Sauer, after which I expect we can issue Stage 1 IPA.

I look forward to receiving your revised manuscript. Please note that due to the December shutdown at PCI RR, you won't be able to submit your revised manuscript until we reopen on January 3rd. I will action your manuscript as quickly as possible at that time.

Response 1: I am grateful to Prof Chambers and the reviewers for the comments. Thanks also for sending these through before the holidays. I hope everyone has since had a good new year break. I have addressed Dr Sauer's comments in detail below.

--

Review 1 by Dr Jim Sauer, 14 Dec 2022 22:33

The responses to the previous reviews are thoughtful, open, and conscientious and have largely addressed my concerns. I have only a couple of relatively minor comments on the revised submission.

1. RQ2 and H2

H2 currently reads "Video games previously known to be high-grossing and contain loot boxes and presently containing loot boxes on the Google Play Store will accurately display the IARC 'In-Game Purchases (Includes Random Items)' label".

I think there are a couple of issues here. First, the clause "previously known to be high grossing" feels vague (e.g., feels like you'd need to operationalise "high grossing" and the manner through which this was "previously known"?). Second, the hypothesis is a bit wordy, and a little difficult to follow. Perhaps the hypothesis could be simplified? Maybe something like: "All titles in this sample of video games previously known to contain loot boxes, and available on the Google Play Store, will accurately display the IARC 'In-Game Purchases (Includes Random Items)' label"?

Response 2: Thanks again to Dr Sauer for these helpful comments and suggestions. I agree with Dr Sauer that including the historically high-grossing nature of these games is unnecessary and too wordy. Therefore, I have rephrased H2 as follows to capture these three essential points: (i) previously known to contain loot boxes; (ii) presently available on the Google Play Store; and (iii) presently contain loot boxes. I think the remaining wordiness is unfortunately unavoidable.

Hypothesis 2: All titles in the present sample of video games previously known to contain loot boxes, and which are presently available on the Google Play Store and continue to contain loot boxes, will accurately display the IARC 'In-Game Purchases (Includes Random Items)' label.

2. Representativeness

Given the revisions to the RQ2 and H2, the sample can now provide data capable of testing the hypothesis, so that issue has been resolved. And the revised ms also explicitly considers the issues relating to the representativeness of the sample used, and I appreciate the openness here. The question, however, is will compliance data from this sample of very popular and highly scrutinised games tell us something useful about compliance behaviour more generally? On the one hand, compliance may well be higher for these games than for other, less popular and less scrutinised games, and therefore the current investigation may produce an overestimate of compliance rates. On the other hand, these are the popular games; the games interacted with most by gamers. I've wavered back and forth on this issue while writing this review. Ultimately, I think, the proposed methodology/sample is suitable. Although it may

overestimate compliance for the market as a whole, it addresses the titles most gamers play (and therefore provides data of broad enough relevance to be practically useful) and I'm not unsympathetic to the arguments relating to expediency and efficiency (and the impracticality of assessing "all games").

Response 3: Dr Sauer is right in saying that there are pros and cons to the present sample whose representativeness is somewhat constrained by practical reasons. I have noted these in my previous studies that relied on highest-grossing games (for example, my recent Belgian study: <https://doi.org/10.31219/osf.io/hnd7w> (Version 6, Page 44, Lines 1438–1447). However, I think the most important point which Dr Sauer has also mentioned is that stakeholders (players, parents and policymakers) would likely be more interested in how these popular games are complying rather than how unknown, poor performing games are complying. An additional (unlikely) counterpoint worth considering is that perhaps the high-grossing games are doing well because they are less compliant than average... I will highlight these limitations again in the discussion section after the study has been conducted.

3. 95% cut-off

I understand that any cut-off of this sort is going to be largely arbitrary and, as long as this is acknowledged, all is well.

4. Minecraft and Roblox

Interesting issue, eh? I think the proposed treatment of these titles is a sensible approach.

Response 4: I am glad that Dr Sauer has found my revisions agreeable in relation to these two points.

--

Review 2 by Prof Pete Etchells, 22 Nov 2022 10:33

Thanks to the researcher for addressing my comments in the previous round of reviewer. These have been done to a satisfactory standard, and I have no further suggestions to add. I note that there was some disagreement between myself and the second reviewer regarding how Minecraft and Roblox should be treated in Study 2; I think that the proposed solution has merit, and does a good job of balancing the need to treat these games as unusual entities with regards to loot box systems, while also considering them in the context of future avenues for research.

Response 5: Thanks to Prof Etchells for taking the time to review this manuscript again and helping to improve the methodology.