Cover Letter

Dear Prof. Chambers and Reviewers,

Thank you for considering our work for further assessment. We have now revised based on the highly useful feedback. A separate table below includes a point-by-point response to each comment. We also provide two revision files: one with tracked changes to provide an overview of changes, and another in clean format without change tracking for the ease of reading. A few general things are listed in this letter.

1. We have pursued to meet the programmatic criteria by adding significantly more content and details. However, we also understand that the study design might not be a good fit with the programmatic approach due to the second outcome being related to the first outcome (which we do not have yet). We offer a partial solution by suggesting two alternative reporting paths with otherwise similar methodology, being aware that this may still not be satisfactory. We let the reviewers/recommender decide whether the programmatic approach applies here, and if not, all the red font text (including the separate reference list) can simply be deleted, which leaves a clean autonomous Stage 1 RR for the first round.
2. Related to the above, we have removed the second-round interviews of medical experts from the second outcome. The expert' role in Study 2 was not clear, and after further thinking, we felt they would not clearly contribute to the longitudinal design. We have thus removed Group 3 second-round interviews from this registration.
3. To update: we have now carried out two Group 2 interviews and one Group 3 interview, maintaining our data at Level 4. The data have not been transcribed or analyzed, and the collection procedure has followed Stage 1 registration.
4. Just to make a note of this: none of us authors are native English speakers, and our university provides English proofreading for accepted articles. The local policy does not discuss RRs. I do not recall proofreading being discussed in the PCI RR guidelines; however, corrections to language being allowed, perhaps proofreading corrections can be integrated at Stage 2, with the recommender confirming that the changes do not modify Stage 1 content.
5. The manuscript is now 10,000 words long. Even though this includes appendices and the second outcome, there is a risk that our methodological accuracy has become a tradeoff against some otherwise potential journals and their word limits. Perhaps moving some of the Stage 1 content into supplement files will be allowed to access journals with word limits.
6. As a lead author, I have a final detail to inquire about. One of our authors, Jukka Vahlo (JV), is part of a company Kinrate Analytics, which does player profiling for game companies. Although our work has no direct financial impact to any direction, and I have no personal reason to suspect that JV's industry position leads to any bias in our study, I must ask what to do with this fact, considering the PCI RR requirement "Authors have no financial conflict of interest relating to the article." Would this be considered a financial

conflict of interest, considering the topic of the study? JV has not contributed yet, as his role is to carry out external assessment of our analysis after coding. If JV is considered having a conflict of interest that is against the PCI RR, we are likely able to negotiate him another role in the project outside this study. On the other hand, I believe JV's expert input would be highly useful, especially considering his research background in enactivism. Because we document our analysis step by step, the changes made to the original coding rounds by JV's contribution will be made visible.

Sincerely, on behalf of the team,
Veli-Matti Karhulahti

Reviews	
	We thank all the reviewers for the constructive and detailed feedback, which highly improves the accuracy of this RR.
Review 1: Malte Elson	
Question	**Response**
1. The incoherence and lack of precision on the conceptual level of IGD is inherited by attempts to operationalize and measure it. The arbitrariness of measures, cutoffs, thresholds, computations etc results in a literature that is difficult to synthesize -- not only because such flexibility offers opportunity for p-hacking, but even more so because it is not sufficiently clear whether different studies on IGD are actually measuring the same thing. These larger points are discussed, e.g., in van Rooij et al (2018); and recently, Satchell et al. (2021) published a study supporting the problem of arbitrariness with the mock-development of a friendship addiction scale.	We have added a new section dicussing these ambiguities in the introduction.
2. It is not sufficiently established whether "problematic game-playing" (however it is operationalized) should be its own diagnosis (which implies that the medium itself may be an immediate cause), or whether it is merely a symptom of another underlying condition (e.g., depression). I think this point deserves particular attention not because it is simply part of the larger discourse around IGD, but also because it has implications for the present study design:	We have added a new section dicussing the lack of causal and theoretical evidence in the introduction.
The authors are aiming to recruit participants for group 1 by inviting gamers who report mental health problems related to their gaming habits. I don't think this by itself is problematic at all compared to, e.g., using a screening instrument, but I think the point that gaming itself may not necessarily be the proximate or distal cause of the disorder should be further discussed.	We now highlight (in several sections) that treatment-seeking serves merely as a proxy for assumed problems, as we cannot assess the clinical validity of those assumed problems. Also note the added Appendix 2 (which will not be used for verification but improves the accuracy of describing the health status of the participants).

While I find the comparison "gamers with disorders" with "esport players without disorders" an interesting approach, it naturally has some limitations for the goal of the proposed research. For example, it might be conceivable that a disorder is a major hurdle in becoming a competitive, (semi-) professional esport player. As an experimental psychologist, my toolbox woud suggest sampling gamers from a group that is similar except for the variable of interest (whether or not they suffer from mental health problems); of course, I understand that the nature of this study does not demand such strict control of confounding factors, but maybe this point deserves some further attention beyond what is currently proposed in the paper.	We have added discussion about this, especially in the participants section. It is now clearly stated that none of the Group 2 participants will be professional or semi-professional players, streamers or other people for whom gaming is a major part of their profession in order to minimize the differences except for the variable of interest. It must be noted, however, that we do not know the Group 1 at this point, i.e. some of them might be professional or semi-professional players, streamers etc. with related health problems (in which case excluding such individuals from Group 2 might be unwise). That said, we consider this scenario improbable, and thus exclude the professionals from Group 2 (we cannot afford to exclude them from Group 1 in case such individuals happen to be treatment-seekers).
My feeling was that the route for study 1 was quite clear, but that study 2 was quite vague in contrast; part of the reason is of course that the outcome of the first round of interviews could dramatically change what is currenty planned for the second round. That, however, sort of defeats the purpose of Registered Reports.	We have now exapanded Study 2 significantly, through the manuscript.
I also see some concerns regarding potential dropouts, which are currently not accounted for. Assuming, for example, the authors are able to recruit 7 interview candidates for round 1, of which 3 are no longer available for round 2, the outcome of study 2 could be quite limited.	This is a highly important comment; we simply forgot these key details. The dropouts are now discussed explicitly and taken into consideration in the follow-up.
(My own feeling about this is that the authors should rewrite this as an RR for study 1 and use their findings to write a much clearer, separate RR for study 2 at a later point, but maybe this too strongly interferes with the authors' project roadmap.)	We fully sympathize with this alternative, and leave it for the recommender and reviewers to decide whether the two studies are accepted as programmatic or the latter be turned into a separate RR later. In case of a separate RR, the red parts can simply be deleted and a coherent cross-sectional RR will be left (the red parts also have a separate reference list).
I'm not 100% clear which parts of the data the authors will be able to share. The recordings or literal transcripts of the interviews seem out of question, so maybe this point could be clarified.	In fact, we do intend to share transcripts. This has now been clarified, and there is now more discussion about it. We understand this is a relatively unique element in qualitative research, but we have put a lot of effort in making that possible (including previous experiences of it).
Review 2: Peter Branney	
The use of 'qualitatively' (penultimate paragraph before the methods section) is extremely broad and means the reader doesn't really have a sense of the approach you are taking. 'Qualitatively' could encompass everything from positivist to social constructionist approaches.	We have specified our use of the term and removed it whenever it has not served the correct purpose.
You mention that you will conduct phenomenological interviews in the abstract and in the Sample Justification you mention IPA. Can you instead describe your specific phenomenological approach and how and why this is appropriate for your research questions?	We now systematically refer to IPA throughout the manuscript. As a small caveat, we will develop an iterative "manual" to document our analysis and carry out the comparisons. We consider this an important addition, as IPA analyses are typically rather closed and not very transparent. We hope sharing the manual and its versions help improving these issues.
Plus, can you consistently refer to this phenomenological approach throughout the paper?	Done.

Personally, it would be good to see a brief elaboration on how you can use a phenomenological approach longitudinally perhaps with reference to the longitudinal qualitative, such as the Timescapes projects: https://timescapes-archive.leeds.ac.uk/timescapes/ or others that are specifically phenomenological.	We now provide further context and discussion for longitudinal IPA methodology. Instead of Timescape projects (which we were not deeply familiar with before), we cite a longitudinal IPA review and its methodological discussion in previous literature.
can you specify your aim or aims for this study?	The aims are now specified so that our contribution to the overall field is more clear.
The use of hypotheses is interesting. My first thought that hypotheses unnecessary. Next, I was persuaded by your argument as significance as meaning, so I was open to see how you used them. The hypotheses outline what you 'expect', so I can see how they might be comparable to a confirmatory hypothesis. Last, I'm sitting on fence. I can see the potential benefit of outlining your expectations; but I also think they could also be presented as expectations rather than hypotheses. As such, I'm not saying you should or shouldn't change it - just giving my perspective on reading it.	We have kept the QH concept; however, we are still open for re-naming these in case the recommenders and reviewers so prefer. The term is not important here; we could also refer to "non-testable hypotheses," "predictions," "expectations," etc. But currently, to build a bridge between quant RR and qual RR (while also distinguishing between the two), it feels logical to have the word "hypothesis" here.
Archiving qualitative data is a delicate topic; is it worth exploring this briefly in the ethics section and explaining how you plan to get consent and the approach to anonymity and confidentiality?	This is certainly a significant piece of feedback. We now have elaborated on this issue. Sharing complete transcripts is indeed a complex and unorthodox feat in the field, and we could easily write a full article about the challenges and solutions related to this process (the draft is on the desk!). If more details are needed, we might add one more supplement regarding these issues at Stage 2. Because most journals of PCI RR do not have unlimited word limits, we must navigate the tradeoffs in the length of the manuscript carefully.
I thought the red font was unnecessary	We agree the current readability does not benefit from the red font, however, we did not find a better way to separate the two studies -- considering that Study 1 should be written into a finalized form, without notable modifications after IPA. So we still kept the red font in a way that *all red text can be entirely removed* (including the separate reference list) and the remaining text remains an autonomous Stage 1 RR without further amendments.
Can you follow the Journal Article Reporting Standards for Qualitative Methods (https://apastyle.apa.org/jars/qual-table-1.pdf0 and include the researcher description and the researcher-participant relationship.	We have included these details from the APA guidelines, in addition to which there is now a full 32-point COREQ list as an appendix.
In the abstract, can you clarify your how you will analyse the data. I note that you specify you will approach the questions with 'phenomenological interviews'; a phenomenological data analysis is therefore implicit. There are a wide range of phenomenological approaches to data analysis, so could be make it explicit by specifying?	These details have been added.
You mention IPA in the method section - if this is an IPA study, can you mention this consistently throughout? From my reading, this seems to be a longitudinal IPA study.	Now referring to IPA systematically.
In the abstract, the group 'those who play esports more than 4 hours per day without problems' is 'without problems' too simplistic? While playing, they may still have problems, such as neglecting childcare or other responsibilities.	We now refer to the two groups systematically as "treatment-seekers" and "players without self-reported health problems related to gaming."
Later in the sample size justification, you use the term 'self-reported health problems'. Is there are better way of describing this group (and can you use the same term consistently throughout the paper)?	In addition to the above, we now discuss the challenges related to this terminology in the introduction. It is difficult to balance between accuracy and readability; we hope the solution works out.

As you have two time points, can you clarify how you will synthesise the data between the two time points? Will you, for example, keep the idiographic focus usually seen in IPA and, if so, how? In the analysis section, you mention that 'the process will be repeated in 12 months'. Wouldn't this give you two separate analyses (which would be incredibly useful) but doesn't explain how you will consider time and/or synthesis findings from the two time points.	We now significantly expanded the Study 2 descriptions. Additionally, we refer to the attached POP-UP interview frame, which includes explicit questions related to the changes over the past 12 months. We believe this will allow us to analyze the responses idiographically *and* including the change element.
For group 3, the interview type is 'non-phenomenological'. Can you specify what it is rather than what it is not?	This has now been fixed.
Interesting use of the Phenomenology of Play interview frame; can you highlight this in the abstract?	This has now been added.
As the interviews will be in Finnish, can you clarify which language the analysis will be conducted in and when translation will occur. E.g., will you conduct the IPA in Finnish and translate theme names and illustrative quotes into English for publication?	These details have been added.
For the analysis plan, can you clarify that it will be idiographic (e.g. One person at a time); from your description, this could also be thematic analysis.	We have significantly clarified the idiographic element in the analysis throughout the manuscript.
Can you provide the reference for the 'phenomenological manual'?	This element is our own; we have further clarified its role.
Reviewer 3: Michelle Colder Carras	
One area that I think should be re-evaluated is the hypothesis (QH2) that suggests that participants with clinical-level gaming-related health problems will be likely not to express the pursuit of self-development and social value. This is contrary to previous research that suggests that gaming offers benefits as well as problems, even to those who may have clinical-level problems. As with other forms of "addiction", even at the level where use is compulsive and out of control, people may still be experiencing benefits. This may be even greater for gamers who belong to guilds or stream regularly to an established community.	This was a truly valuable observation; we agree with it and have reformulated the QH accordingly.
Consider adding design elements or reporting some information that can be used later to meet standards for qualitative reporting. Some information (per JARS-qual) that could be added is reflexivity on the researchers' experiences and backgrounds. One design element that could be added is member checking. I realize this is somewhat contentious in phenomenological qualitative research, but to me as a gamer/researcher/person who has experienced some "addiction" symptoms as well as mental health challenges, the idea of using clinician interviews after the fact seems less supportable. Just a bit more explanation of this choice would help, or some consideration of additional methods of triangulation outside the research group.	We added details regarding our member checking and attach a full COREQ list as an appendix.
What is the study procedure? It was not clear to me until the final page that the Group 3 interviews would take place after analysis of Groups 1 and 2.	We have clarified the procedure.

More information on the recruitment and screening process is needed. How will people be screened into the study (i.e., how will inclusion/exclusion criteria be assessed) if they respond to advertisements? How will it be determined that esports players do not have gaming "related health problems"? What exactly is the definition of esports? Please give some example games. Will it be assessed whether people are playing esports competitively/professionally? What if there are options in the game to play solo vs competitively; will that be assessed? Will professional esports athletes or sponsored streamers be included? How will the study purpose be explained to participants? Will the interviewers reveal their own perspectives/backgrounds during the screening or interviews?	We have added this information, some of which are in the COREQ list to save manuscript space. Also note the added Appendix 2 (which will not be used for recruitment but improves the accuracy of describing the health status of the participants).
What is the expectation of retaining members for the 2nd interview and what will be done to enhance this probability?	This is now discussed explicitly.
"which can be seriously harmful when over and under medicating…" implies that medication is the treatment that will be given. This should be expanded to (potentially unnecessary) psychosocial treatments. It may also good to discuss this in light of children/adolescents being brought to treatment and that as a coercive practice.	These elements are now noted.
Please support "which has recently become one of the most prevalent forms of videogame play" with a citation. I'm not sure I agree with this.	The industry statistics are generally not very reliable; instead of a citation, we simply toned down the sentence into "one of the major forms of videogame play."
The explanation of qualitative hypotheses as disclosing biases is a bit confusing to someone with little knowledge of IPA. How does this differ from the researchers' inherent biases and how these are addressed in other forms of qualitative analysis (e.g., reflexivity)?	The QH appraoch is novel and, indeed, overalps with reflexivity (while being primarily based on literature). The COREQ appendix adds to our interviewer reflexivity.
"We are not aware of reviews that summarize these results" – my unpublished thesis discusses this in the section "Incidence and transition between states of problematic gaming, at-risk and non-problematic gaming" on page 28. https://jscholarship.library.jhu.edu/bitstream/handle/1774.2/39616/COLDERCARRAS-DISSERTATION-2015.pdf?sequence=1	We now discuss this previous literature in more detail.
Are there figures saying that gaming-related clinical services are rarely used in Finland? Even a personal communication would help.	We now cite personal communication. Indeed, there are no citable national statistics about this, but our mapping of the local field systematically yields the same answer: people very rarely seek these services, and service providers thus struggle to maintain their practice.
I didn't realize that the ICD-11 diagnosis is not yet in clinical use; is it still in field trials? Please cite	This is now cited.
Please provide a link to Pelit and give 1-2 more examples of forums that will be used to recruit gamers	We have elaborated on this, with one central change: we recently found a local amateur esports league for adults at work, so we use this channel as our starting point instead of Pelit, which today is less focused on competitive play.
Please clarify whether full time students will be included in the esports players sample or only those with a full time job.	They are not included, and this is now stated.
Will only Finnish-speaking people be recruited?	Also English people can participate if they meet other criteria; these details have been now added.

Have you considered using DeepL or some other supported AI translation to translate interviews into English as an intermediate step, with the caveat that they are not quality-checked? This could be a very valuable dataset. Perhaps such transcripts could be posted as well, and it could be made clear that the interviews were machine translated. Eventually the authors may find collaborators willing to provide funding support for further translation/quality checking.	We definitely would love to translate and share the data in English, however, there are practical challenges in this. First, in order to maximize participant privacy and safety, the transcripts are shared in collaboration with FSD who has experience and infrastructure in sharing qualitative health data. The FSD, however, only shares quality-checked transcripts, so in order to share any English versions, we would first need the funding for secured translation/back-translation. So currenty, we unfortuantely do not see a way to share translated versions. That said, if any external researcher or team happens to have funds for translation, we are more than happy to collaborate in the translation process to make the data reusable internationally.
My apologies if this was covered, but what will happen to the audio recordings? How long will they be retained?	We did not meantion this earlier -- it has now been included (according to the local ethics standards, the audio is a personal identifier that should not be stored without special reasons).
Please provide more description of the documentation of the coding process.	Further details have been added.
Please clarify the role of the Stage 1 interview phenomenological manuals in the Stage 2 interview analysis.	This has been now clarified.
Typos "those who play videogames daily with better" -> have better	Fixed.
neglection of needs/responsibilities -> neglect	Fixed.
a priory I think should be a priori	Fixed.