

**Colder Carras**

This Stage 2 RR reports results of a longitudinal qualitative study of treatment-seeking gamers and amateur esports players to develop a phenomenological model of normative, intense playing vs. problematic gaming. The report builds on previous research characterizing types of players and seeks to describe players' own meanings of the role of gaming in their lives. The findings suggest several themes that align closely with multiple concepts of normal development as well as mental disorders and should be carefully considered in understanding and treating problematic gaming. The Stage 2 report satisfies the review criteria. I have a few suggestions that the authors might consider in light of their findings. Some are very minor; I will address the more substantive ones first.

Thank you, we appreciate the detailed feedback during this hectic spring semester.

Life thinning seems to me to be very close to the concept of salience of addictions. Apologies that I can't find a great reference for this, but [Griffiths 2005](#) provides a good overview. DSM-IV-TR mentions this on page 195, "In some instances of Substance Dependence, virtually all of the person's daily activities revolve around the substance. Important social, occupational, or recreational activities may be given up or reduced because of substance use (Criterion 6)." To me, this behavioral aspect has always been separate from the cognitive preoccupation, and seems to fit well with the idea of life thinning. I was surprised that this wasn't discussed, but I did appreciate the developmental aspects and the attention also to ways in which depression might develop. I hope the authors will consider whether there might be something worth including regarding theories of addiction and salience, especially as there seems to be a different type of salience involved in the esports players meaning-making.

This was a very valuable observation and indeed we missed the link previously. We went through the literature and likewise found the conceptual range slightly mixed: preoccupation, intruding thoughts, salience, and activities taking over one's life all seem to be associated to some degree and are defined in various ways, yet there appear to be no consensus. We fully agree that a reference to this literature is necessary. We have added a new section about this.

One other theoretical concept that I didn't see discussed was perhaps my own inkling of an idea based on the initial results presented through the stories of Caius and Frederika: the idea that there is a push-pull relationship to gaming.

We also agree that the push-pull pattern is relevant here. In fact, when we reviewed the literature at Stage 1, we found this pattern explicitly discussed by Shi et al (2019) but those findings had been simply forgotten because of our focus on longitudinal patterns here. We have brought back Shi et al (2019) and link the push-pull theme to our findings.

Less important suggestions:

It might be worthwhile to have a native English speaker provide a review of the text. It is certainly readable, but every once in a while there is a phrase that is not perfectly clear and requires a bit of rereading.

The manuscript has now been proofread by a professional. Some changes to Stage 1 sections too were introduced by the proofreading. They have been marked in a separate file at OSF: <https://osf.io/pa64r>

It might also be worth checking the numbers for time gaming. It's not clear to me how Frederika's 10-16 hours of gaming relates to the table reported time, both on page 6.

We have now clarified this. In this section she referred to her gaming before treatment and she quit gaming entirely when the treatment started (with periodic exceptions, as discussed elsewhere).

Could you please discuss where the scale results can be found? It would also be worth including the table note about scales in the first table where they are reported. I also found myself guessing at gaming time/other gaming time, which I don't think is explained until the note in Table 2.

We have moved Table 2 to earlier, which now allows readers to access the information at an earlier point. We apologise for the missing file and now have also created a data file of the survey responses and stored it at OSF like in the previous study. The sensitive qualitative data will go through a detailed de-identification process at the FSD archive, as in the previous study.

For accessibility purposes, it might be worth considering whether to use something other than shading to add meaning to the cells in tables. Also, in some cases people may still be printing out papers (I do so when reviewing), and the shading is barely visible when printed.

This is a valid point; we have revised the formatting. We also revised the colour codes in the other table.

Where "perceived availability" is used, this might be explained a bit more.

We have clarified this and made the text more consistent by systematically using "experienced availability".

The idea of life domains is important. These are mentioned once or twice but I believe are vital to the idea of conative dynamics, so worth bringing out a bit more.

We have further clarified the life domains as a plural concept and expanded on them.

In the General Discussion, I note that clinical interviews weren't carried out, which is fair. I also note that it seems that at baseline, Frederika only satisfied 4/5 IGD criteria. The concurrence between treatment seeking and scale scores might be worth considering, as subthreshold symptoms are a big topic of debate.

We agree with the relevance of addressing the IGD and GD criteria. Regarding the baseline symptoms, we addressed this issue to some degree in the first Stage 2 report that was based on the first-round interviews (section "Instruments, Limitations, and New Hypotheses"). Perhaps the most interesting finding, pointed out by our participants, was the problematic conception of time in the instruments, e.g. "official diagnostic criteria such as *increasing* interest in gaming simply did not apply" (because their gaming hadn't increased anymore in a long time, they had been gaming heavily for decade/s). We compared all 4 instruments across our 16 baseline participants and concluded that single-item THL1, which asks to self-rate problems, worked best to identify treatment-seekers. We hope this contributes to the ongoing clinical validation efforts in the field.

Because the second-round interviews didn't add much new knowledge to this, we decided not to revisit the issue. However, in one of the interviews, we had a long explicit discussion about the measures and discussed their wordings at length. These data could be utilized in a future study (by any interested party).

There are a few instances in which a lack of studies is asserted, eg page 16 “a lack of clinical validation studies with current quantitatively used gaming disorder scales”. It would be good to verify that.

We agree that more accuracy is needed for such claims. At the same time, as the saying goes, it is difficult to evidence the nonexistence of something. This partially applies to clinical validations: two years ago we systematically checked the references to all major scales in order to find clinical validations (when mapping the literature for our content validity analysis), and the result was close to zero. A couple of exceptions emerged, like the C-VAT 2.0, but they are pre-ICD-11 and thus not GD-scales. The most popular GD scales, like GDT, do not list clinical validations. As a compromise, we've added a reference to a WHO team paper from 2022 where lacking validations are noted to be fixed with a new fully validated scale (i.e., the paper doesn't systematically review validations and it's not a perfect reference, but if the WHO team isn't aware of properly validated scales, that's relatively good evidence for their absence). We have also removed a sentence noting the absence of qualitative longitudinal studies with treatment-seekers; we do not believe many or any of them exist, but we also agree that the sentence makes a claim that is not needed and remains hard to back up without a systematic review.

Thank you for the opportunity to review this excellent paper.

One more thank you for all valuable comments both at Stage 1 and Stage 2!

### **Branney**

Thanks for the opportunity to review this Registered Report at Stage 2. As a longitudinal study, this is the second Stage 2 report I have reviewed. The first Stage 2 report was the first set of data (DOI: 10.1525/collabra.38819) and this current report includes data from one-year follow up. I have two potential conflicts of interest that I want to declare. First, I have recently published with Karhulahti (10.12688/openreseurope.15532.1). Indeed, the topic of the publication - registered reports of qualitative research - was sparked from my role reviewing the Stage 1 and first Stage 2 reports for this longitudinal study (DOI: 10.1525/collabra.38819). Second, I was a co-editor of a special section in the British Journal of Social Psychology (DOI: 10.1111/bjso.12628) in which Karhulahti published a paper (DOI: 10.1111/bjso.12573). If I recall correctly, Karhulahti and I discussed the risks of collaboration to the Stage 2 review and only worked on the 'primer' paper once the first Stage 2 paper had been completed and reviewed. I am tempted to share that the presence of conflicts of interest at this point of this longitudinal registered report is, for me, a happy coincidence. Nevertheless, it is perhaps less coincidence than a confluence of forces that sees qualitative researchers and open science taking an interest in each other and the new technologies of registered reports and open review, where reviewer and reviewed get to learn something about each others interests and perspectives at the early stage of a project. I share this information so that you, dear reader can judge my review. Below, I have structured my review according to the criteria for assessing a Registered Report at Stage 2 from PCI RR (accessed 2023-05-27; [PCI Registered Reports (peercommunityin.org)]([https://rr.peercommunityin.org/about/full\\_policies](https://rr.peercommunityin.org/about/full_policies)))

We appreciate the transparency and add that we also disclosed the same sequence of events in our cover letter to this Stage 2 report.

*# 2A. Whether the data are able to test the authors' proposed hypotheses (or answer the proposed research question) by passing the approved outcome-neutral criteria, such as absence of floor and ceiling effects or success of positive controls or other quality checks. \*\* This criterion addresses whether the data quality is sufficient to be able to test the stated hypotheses, according to the pre-specified conditions in 1E. Since not all protocols are able to pre-specify outcome-neutral tests, this assessment is not relevant*

*to all forms of RRs. Where it is relevant and pre-specified, it is possible that the failure of a crucial outcome-neutral test could, in severe circumstances, lead to the rejection of a Stage 2 manuscript. \*\**

The paper reports that the data has been submitted to an archive but no citation is provided, so it is presumably not available at the time of this Stage 2 report was submitted. The Stage 2 baseline report has a data availability statement clarifying that the data is published in a Finish Social Science Data Archive and commenting on access requests. As far as I can tell, the Stage 2 baseline report does not provide bibliographic information (i.e. the citation) on this data archive. I appreciate that the review of the submission of the data to the Finish archive is a separate process to this Registered Report and it may well be a lengthy process. Nevertheless, the bibliographic information is important in making sure the data is accessible. Indeed, providing this bibliographic information in this Stage 2 registered report of the longitudinal data helps make the data accessible to its readers. Consequently, can I request the authors find a way to delay the final publication of this paper until the full bibliographic information of the data is available?

We agree that sharing sensitive qualitative data remains a technical challenge for publications and will hopefully find better solutions in the future. After the first round (first programmatic outcome), it took several months for FSD to systematically verify that no personal identifiers were left in the shared data files. That said, we are happy to have waited that long, as despite our own double verification, the FSD found a few sentences that could have enabled identification of our participants, and they were corrected. Those final datasets are now available at <https://urn.fi/urn:nbn:fi:fsd:T-FSD3678>

We agree this lengthy de-identification is problematic: although it was possible to update the OSF preprint with the final data URL, the official copyedited publication could not be updated and remains without the final data location. As suggested above, one solution could be to delay publication until all data are published, but this will unlikely be optimal as a long-term solution because many Stage 2 outcomes of PhD and other projects entail timely reporting. Perhaps the “added value” of journals in the future could be to offer pre-agreed post-publication editing, such as verified qualitative datasets?

*# 2B. Whether the \*\* introduction, rationale and stated hypotheses (where applicable) are the same as the approved Stage 1 submission.\*\* \*\*This criterion assesses whether the authors have remained consistent in their framing of the study at Stage 2. Aside from changes in tense (e.g. future tense to past tense), correction of typographic and grammatical errors, and correction of clear factual errors, the introduction, rationale and hypotheses of the Stage 2 submission must remain identical to those in the approved Stage 1 manuscript. To make any changes clear, authors are required to submit a tracked changes version of the manuscript at Stage 2. \*\**

This paper mostly has the same rationale as the Stage 1 report but I would recommend elaborating on the differences. First, I would recommend you elaborate that this is a ‘new’ Stage 2 report rather than an update to the original Stage 1 report. What I’ve learnt from my limited experience with registered reports is that the Stage 1 and Stage 2 reports are the same ‘report’ but that new unknown information is added at Stage 2. That is, the Stage 1 report comprises 1) the methodology and 2) the rationale for this before before we learned about the findings. In the Stage 2 report, the researchers put the methodology into practice and therefore add 1) what we found when we tried to do what we planned to do and 2) what we have learnt from these findings. The recommendation of PCI RR for the inclusion of a tracked changes version highlights that the researchers should be showing, and the reviewers should be mindful of, any changes between Stage 1 and 2. Indeed, I can look back at my comment on the revised Stage 1 submission and I am reminded that I mistakenly thought the ‘final article’ would be something new; “\*You do mention about the challenge of word limit. I imagine that the final article will be much shorter

but readers will benefit incredibly from access to your RR\*” ([Phenomenological Strands for Gaming Disorder and Esports Play: A Qualitative Registered Report (peercommunityin.org)](peercommunityin.org/user/recommendations?articleId=44&\_signature=921e36c21ad6fda1d8ec7919ede11c3aedf0af97)). Nevertheless, this Stage 2 report is a new document and I can understand that this is so because the baseline Stage 2 report has been completed and published (DOI: [10.1525/5ollabra.38819](https://doi.org/10.1525/5ollabra.38819)). As such, we are in new territory, figuring out how to do Stage 2 reports when we are used to persuading journals to publish our work after study completion and we have little ability to change what we did. I do think this issue crops up throughout the paper as it’s difficult a time to see if new detail is a posthoc addition. Consequently, can you succinctly explain why this Stage 2 report is as it is, perhaps ensuring you provide enough clarity that someone unfamiliar with, and indeed those familiar with, the Registered Report format?

These are very important observations and, indeed, the most optimal technical solutions for reporting “programmatically” studies will likely be found later in the future. In our case, at Stage 1 (<https://osf.io/a2rww>), we used red colour to indicate sections that were going to be moved to this second Stage 2, and those sections represent only a very small section of this longish article. We had to write a new introduction, and even though we largely followed the original analysis plan, the methods section was also written anew in order to avoid plagiarism in relation to the first Stage 2 outcome. In the future, when the academic world—which is slow to adapt new publication formats—will be ready for it, it could be feasible to publish “forking” articles that share the same introduction and methods. This might solve issues of repeating the same information in multiple introductions and methods sections when a project produces several analyses over a single data generation process.

To clarify these complexities in this manuscript, we have added a clarifying section that highlights the relationship between this and the previous study. We hope that this addition will be helpful for readers and should be somehow added to programmatic papers in the future, clarifying and communicating which parts were written before and after the planning phase.

Second, this one-year follow up Stage 2 report briefly paraphrases the rationale for this study (in the first Stage 1 report and the published Stage 2 report on the baseline data) that there is 1) contested notion that it is possible to diagnose people who have a 'gaming disorder' despite 2) a dearth of research about people who seek treatment for such a diagnosis. What is new to the rationale - and is highlighted in yellow - is the addition of the notions i) that gaming behaviour may be cyclical rather than continuous and ii) that gaming habits develop slowly. In addition, the term 'because' ("==because== ... we expect" and "==because== ... we do not expect"), indicates a chronological, linear logic that belies the retrospective way in which you learned this new information and updated your hypotheses. While highlighting these in yellow does to some extent help to indicate that they are new or distinct, I would challenge the wisdom of a posthoc update to your Stage 1 qualitative hypotheses in this particular way. I appreciate that in raising this point, I risk a response that you remove this update to your hypothesis, which will in turn be another posthoc change to your now Stage 2 report (i.e., if you delete the addition and make no mention of it, something about how this research developed will be hidden). I hope you can find a compromise where you keep the update while elaborating on why you made it in such a way that stays true to your Stage 1 report and is understandable to both those familiar and unfamiliar with Stage 2 reports.

We hope the clarifications noted earlier have made the initial part of the article better at communicating differences between pre and post registration text. Most critically, while the yellow additions in the QH were indeed new sentences, they were not intended to add new information but only reiterated what was registered (“high-powered studies seem to imply that many such health problems are cyclical and do not continue for 12 months”). We thought that the QH would be clearer with the two explanatory sentences spelled out; however, as we see it now, there could be different interpretations and it might be better to

simply erase those sentences in order to stay as close as possible to the original text (technically, HARKing could apply to qualitative work too and we want to avoid any possible interpretations of that). We thus decided to remove the added parts in the QH but kept the new sentence about epistemology, which is less prone to add confusion.

We appreciate the reflexivity in the call for a compromise, but we also honestly believe that minimising changes to Stage 1 should be a default solution and we (as one of the first qualitative RRs) should serve as an example by not adding new information that is not necessary.

*# 2C. Whether the authors adhered precisely to the registered study procedures.*

*\*\* This criterion assesses compliance with protocol. In cases where the preregistered protocol is altered after IPA due to unforeseen circumstances (e.g. change of equipment or unanticipated technical error), the authors must consult the PCI RR recommender immediately for advice, prior to the completion of data collection. Minor changes to the protocol may be permitted per recommender's discretion. In such cases, IPA would be preserved and the deviation reported in the Stage 2 submission. If the authors wish to alter the study procedures more substantially following IPA but still wish to publish their article as a Registered Report then the manuscript must be withdrawn and resubmitted as a new Stage 1 submission. The outcome of all preregistered analyses must be reported in the manuscript, except in rare instances where a preregistered and approved analysis is subsequently shown to be logically flawed or unfounded. In such cases, the authors, reviewers, and recommender must agree that a collective error of judgment was made and that the analysis is inappropriate. In such cases the analysis would still be mentioned in the Stage 2 method but omitted with justification from the results. Additional unregistered analyses can also be included in a final manuscript (see 2D).*

The description of the methodology mirrors the original Stage 1 report. From what I can tell - though apologies if I have missed it in the original Stage 1 report - the 1) detail of the types of changes (narrative change, etc.) and 2) the meeting of the coders are both new additions. Following my comment above about elaborating on the nature of this Stage 2 report, this is one area where the Stage 2 format risks suggesting that this detail was known in advance of the data collection. A deviation to the plan is noted in the methodology and I wonder if this is one way of explaining the way in which the research and researchers have developed since the original Stage 1 and Stage 2 baseline reports. It would be useful to ensure the elaboration resolves issues - the addition of new information that may or may not have been known in advance - such as these.

We fully sympathize with the complexity of the programmatic design, as we also had to carefully revisit our original registration several times and reflect on to what degree the procedures were in line with the plan and which elements were not. We can only imagine the difficulties involved in tracking details externally, without having been directly involved in the qualitative work (it's so different from running quantitative code!).

Regarding the detail of mapping out types of changes, it is true that this element was not present in the first Stage 2 report but we did separately register that detail for the present study (page 12 in <https://osf.io/a2rwg>). As for the second detail (meeting of all coders), it is true that we did not register the entire team meeting over all cases and this has detail now been corrected. The MS has thus been revised to stress this methodological change. [Extra note: we also do not mention in this Stage 2 report the term "manual" which was used in the original Stage 1 plan. As we look back to that, reporting theme development via evolving versions wasn't very well communicated by the term "manual" and our solution was to simply omit the term, as it was considered to improve readability without adding new unplanned terminology and be less confusing versus reframing the process—all the themes and their development are reported in supplements as consensus meeting outcomes, which should be informative

and transparent to an extent that the qualitative process becomes as visible as possible.]

In the findings, where you write "(it was calculated)" can you provide a little more detail? Who, and how, was quitting Heikki's job calculated?

We have now clarified that he had planned to quit beforehand due to disagreements with the superiors.

It is worth highlighting that cyclicity features in one of the themes and cyclicity was one of the additions to the qualitative hypothesis. Consequently, it is important to elaborate on how this new information featured in the development of the research after the Stage 1 report was given accepted-in-principle.

As we have clarified and revised the changes to our QH (see above), we believe this issue is also resolved.

*# 2D. Where applicable, whether any unregistered exploratory analyses are justified, methodologically sound, and informative. \*\* This criterion addresses the quality and value of any additional data analyses that are reported at Stage 2 but were not included in the registered Stage 1 submission. Such analyses are often highly valuable. For instance, a new analytic approach might become available between IPA and Stage 2 review, or a particularly interesting and unexpected finding may emerge. Alternatively, some unexpected characteristic of the data might suggest that the preregistered analyses, while bias-free, are not as sensitive as planned, and therefore a more sensitive post hoc analysis could be informative. Such analyses are admissible but must be clearly identified (e.g. through a separate heading in the Results for "Exploratory Analyses" or "Unregistered analyses"), justified, and appropriately caveated. Authors should also be careful not to base their conclusions entirely on the outcome of unregistered analyses. \*\* A deviation to the plan is noted in the methodology as arising from the Stage 2 baseline review as seems justified. # 2E. Whether the authors' conclusions are justified given the evidence. \*\* This criterion addresses whether the claims drawn by the authors in their conclusions (including in the Discussion, Abstract, and anywhere else in the paper) are warranted by the data or evidence in hand. Note that PCI RR recommendation decisions will never be based on the perceived importance, novelty, or conclusiveness of the results. \*\**

Yes.

One more thank you for all valuable comments both at Stage 1 and Stage 2!