

Do different screening instruments for ‘gaming disorder’ measure the same or different construct(s)?

Charlotte Pennington based on reviews by Daniel Dunleavy, David Ellis, Linda Kaye and 1 anonymous reviewer

A recommendation of:

STAGE 1

Open Access

Identifying Gaming Disorders by Ontology: A Nationally Representative Registered Report

Veli-Matti Karhulahti, Jukka Vahlo, Marcel Martončík, Matti Munukka, Raine Koskimaa, Mikaela von Bonsdorff (2022), *OSF mpz9q*, Stage 1 preregistration, in principle acceptance of version 4 by Peer Community in Registered Reports <https://osf.io/mpz9q/>

Published: 18 January 2022

Copyright: This work is licensed under the Creative Commons Attribution-NoDerivatives 4.0 International License. To view a copy of this license, visit <http://creativecommons.org/licenses/by-nd/4.0/>

Submitted: 25 August 2021, Recommended: 17 January 2022

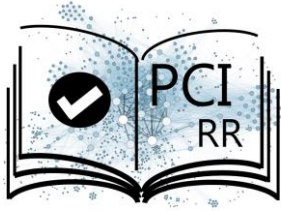
Cite this recommendation as:

Charlotte Pennington (2022) Do different screening instruments for ‘gaming disorder’ measure the same or different construct(s)? *Peer Community in Registered Reports*, 100004. <https://doi.org/10.24072/pci.rr.100004>

Recommendation

There is considerable debate regarding the relationship between excessive gaming and mental health problems. Whilst the diagnostic classification of “gaming disorder” has now been included in the WHO’s International Classification of Diseases (ICD-11), the APA decided not to include this diagnosis in their Diagnostic and Statistical Manual of Mental Disorders (DSM-5) because the literature “suffers from a lack of a standard definition from which to derive prevalence data” (APA 2013, p. 796). Furthermore, screening instruments that aim to provide diagnostic classifications derive from different ontologies and it is not known whether they identify equivalent prevalence rates of ‘gaming disorder’ or even the same individuals.

In this Stage 1 Registered Report, Karhulahti et al. (2022) aim to assess how screening instruments that derive from different ontologies differ in identifying associated problem groups. A nationally representative sample of 8000 Finnish individuals will complete four screening measures to assess the degree of overlap between identified prevalence (how many?), who they identify (what characteristics?) and the health of their identified groups (how healthy?). If these four ontologically diverse instruments operate similarly, this will support the notion of a single “gaming disorder” construct. If, however, the instruments operate differently, this will suggest that efforts should be directed toward assessing the clinical (ir)relevance of multiple constructs. This rigorous study will therefore have important implications for the conceptualisation and measurement of



“gaming disorder”, contributing to the debate around the mixed findings of gaming-related health problems.

Four expert reviewers with field expertise assessed the Stage 1 manuscript over three rounds of in-depth review. Based on detailed and informed responses to the reviewers' comments, the recommender decided that the manuscript met the Stage 1 criteria and therefore awarded in-principle acceptance (IPA).

URL to the preregistered Stage 1 protocol: <https://osf.io/usj5b>

Level of bias control achieved: **Level 6.** *No part of the data or evidence that will be used to answer the research question yet exists and no part will be generated until after IPA.*

List of eligible PCI RR-friendly journals:

- [Addiction Research & Theory](#)
- [F1000Research](#)
- [Peer Community Journal](#)
- [PeerJ](#)
- [Royal Society Open Science](#)
- [Swiss Psychology Open](#)

References

1. APA (American Psychiatric Association). (2013). Diagnostic and Statistical Manual of Mental Disorders (5th Edition). APA.
2. Karhulahti V-M, Vahlo J, Martončík M, Munukka M, Koskimaa R and Bonsdorff M (2022). Identifying Gaming Disorders by Ontology: A Nationally Representative Registered Report. OSF mpz9q, Stage 1 preregistration, in principle acceptance of version 4 by Peer Community in Registered Reports. <https://osf.io/mpz9q/>

Reviews

Evaluation round #3

DOI or URL of the report: [https://mfr.de-](https://mfr.de-1.osf.io/render?url=https://osf.io/k84tb/?direct%26mode=render%26action=download%26mode=render)

[1.osf.io/render?url=https://osf.io/k84tb/?direct%26mode=render%26action=download%26mode=render](https://mfr.de-1.osf.io/render?url=https://osf.io/k84tb/?direct%26mode=render%26action=download%26mode=render)

Version of the report: v3

Author's Reply, 17 Jan 2022

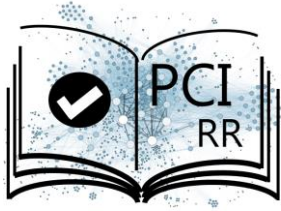
[Download tracked changes file](#)

Dear Charlotte Pennington,

Thank you for noticing the error -- I have now revised the section so that all scales are written to make it coherent. I have also added page numbers to the document.

Best wishes,

Veli-Matti Karhulahti



Decision by *Charlotte Pennington*, 17 Jan 2022

Dear Veli-Matti Karhulahti,

Thank you for satisfying all of the reviewer and editor comments.

Can I please just request one last minor revision - that the GDT (Pontes et al., 2019) and the THL1 (Salonen & Raisamo 2015) are not abbreviated in the first instance on Page 3 (under the research questions). I found that I had to search for these questionnaires to find out their names.

Best wishes,

Charlotte

Evaluation round #2

DOI or URL of the report: [https://mfr.de-](https://mfr.de-1.osf.io/render?url=https://osf.io/8mgns/?direct%26mode=render%26action=download%26mode=render)

[1.osf.io/render?url=https://osf.io/8mgns/?direct%26mode=render%26action=download%26mode=render](https://osf.io/8mgns/?direct%26mode=render%26action=download%26mode=render)

Version of the report: v2

Author's Reply, 13 Jan 2022

[Download tracked changes file](#)

Dear Charlotte Pennington,

Thank you for the (literally) fastest-ever feedback round and the related detailed comments. I will use this format due to the relatively few number of issues.

1. Of crucial importance, there is nowhere in the manuscript that confirms that the data and analysis scripts will be made available, nor is there any indication of this (e.g., folders) on the OSF page. Can you please revise this in line with the PCI RR policy / TOP guidelines. These can be found here: https://rr.peercommunityin.org/help/top_guidelines

--> This has been added on the first page.

2. Page 3: "which in turn was built on studies on "gaming addiction" – the second 'on' should be 'of'.

--> Fixed.

3. Throughout: a comma should not be included within a quotation mark but rather after it, e.g., "gaming disorder," should be "gaming disorder", This appears in many places in the text.

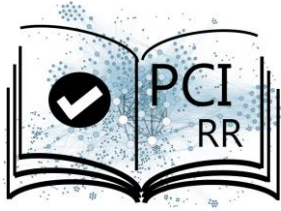
--> Fixed.

4. Page 4, "Because "gaming disorder" is a mental disorder" – I recommend changing this to "is conceptualised as a mental disorder".

--> Fixed.

5. Page 6, Please can you not abbreviate global physical health items (GPH-2) and two global mental health (GMH-2) in the first instance? I had to search for this to understand what they referred to. The H3 statement may therefore need a wording change to clarify this.

--> Fixed, and H3 elaborated.



6. Can you please clarify what is the reason for different alphas under H3? (0.0125 and 0.025)

--> Fixed to consistent 0.025. This was our mistake. Due to the structural similarity of H3a and H3b (as they are addressed together in Appendix 2), we had used Bonferroni correction for their combined number of tests, but indeed the hypotheses are separate and they have two tests each.

7. Page 7, could the following sentence be (a) checked for accuracy and (b) clarified; the repetitive use of double negatives makes it very hard to follow: "If either sample will be smaller than the one needed to detect the observed effect, we will not make inferences in H3a–b unless the upper bound of the effect's confidence interval does not exceed $d=0.22$ (which would support an uninteresting effect)". Can I please check that the authors don't mean "exceeds"? Shouldn't an inference be made if the CI exceeds $d = 0.22$ as this is your smallest effect of interest? (the "unless" and "does not" throw me – if this is indeed correct, is there a better way of putting this?).

--> Indeed, this was a confusing sentence. We have revised it so that only the main point remains: "We will only make inferences based on the effects that are reliably detectable by the final identified groups." I.e., we are aware of the possibility that the groups identified in the final sample may be smaller than the ones in the pilot, thus providing less power, which will be taken into consideration.

Sincerely, on behalf of the team,

Veli-Matti Karhulahti

Decision by *Charlotte Pennington*, 13 Jan 2022

Dear Veli-Matti Karhulahti,

Thank you for submitting the revisions of your Stage 1 Registered Report: "Identifying Gaming Disorders by Ontology: A Nationally Representative Registered Report to PCI Registered Reports".

As the assigned "Recommender" of this manuscript, I have now reviewed your response to reviewers letter and tracked changes under revision 2. I thank you for your attentiveness to the reviewer comments, most of which have been satisfied. At this stage there are some minor revisions to implement before I can make a final decision. I append these below.

Best wishes,

Charlotte Pennington

1. Of crucial importance, there is nowhere in the manuscript that confirms that the data and analysis scripts will be made available, nor is there any indication of this (e.g., folders) on the OSF page. Can you please revise this in line with the PCI RR policy / TOP guidelines. These can be found here: https://rr.peercommunityin.org/help/top_guidelines

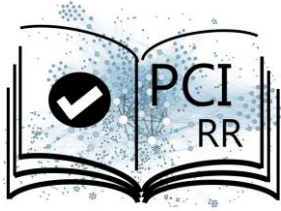
2. Page 3: "which in turn was built on studies on "gaming addiction" – the second 'on' should be 'of'.

3. Throughout: a comma should not be included within a quotation mark but rather after it, e.g., "gaming disorder," should be "gaming disorder", This appears in many places in the text.

4. Page 4, "Because "gaming disorder" is a mental disorder" – I recommend changing this to "is conceptualised as a mental disorder".

5. Page 6, Please can you not abbreviate global physical health items (GPH-2) and two global mental health (GMH-2) in the first instance? I had to search for this to understand what they referred to. The H3 statement may therefore need a wording change to clarify this.

6. Can you please clarify what is the reason for different alphas under H3? (0.0125 and 0.025)



7. Page 7, could the following sentence be (a) checked for accuracy and (b) clarified; the repetitive use of double negatives makes it very hard to follow: “If either sample will be smaller than the one needed to detect the observed effect, we will not make inferences in H3a–b unless the upper bound of the effect’s confidence interval does not exceed $d=0.22$ (which would support an uninteresting effect)”. Can I please check that the authors don’t mean “exceeds”? Shouldn’t an inference be made if the CI exceeds $d = 0.22$ as this is your smallest effect of interest? (the “unless” and “does not” throw me – if this is indeed correct, is there a better way of putting this?).

Evaluation round #1

DOI or URL of the report: [https://mfr.de-](https://mfr.de-1.osf.io/render?url=https://osf.io/72kxy/?direct%26mode=render%26action=download%26mode=render)

[1.osf.io/render?url=https://osf.io/72kxy/?direct%26mode=render%26action=download%26mode=render](https://mfr.de-1.osf.io/render?url=https://osf.io/72kxy/?direct%26mode=render%26action=download%26mode=render)

Author's Reply, None

[Download author's reply](#)[Download tracked changes file](#)

Decision by [Charlotte Pennington](#), 01 Dec 2021

Dear Veli-Matti Karhulahti and co-authors,

Thank you for submitting your Stage 1 Registered Report “Identifying gaming disorders by ontology: A nationally representative Registered Report” for consideration by PCI Registered Reports.

I have now received comments from four expert reviewers in this field. As you will see, these reviews are overall positive and based on these reviews and my own assessment, I would like you to revise your manuscript. Please note that in-principle acceptance (IPA) and progression to Stage 2 is not guaranteed at this point and your revisions may be sent back to the reviewers for further evaluation. You should, therefore, include a point-by-point response to the reviewers' comments, outlining each change made in your manuscript or providing a suitable rebuttal.

You will see below that the majority of comments are minor, but I would like you to pay particular attention to the following:

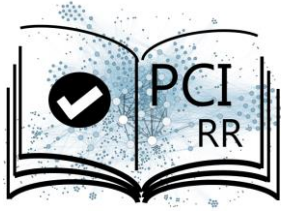
Reviewer 1 makes some important points regarding the methodology, with one comment referring to the different scoring approaches used in the field to categorise problematic gaming, and the potential for these to impact the findings/interpretations. This is where Registered Reports are particularly beneficial, as you justify the scoring approach in advance. I ask that you are particularly attentive in responding to this concern. Please note that Reviewer 1’s comments have been provided in a Word document rather than in-text.

Both Reviewer 1 and 2 note some confusion with the terminology of “remarkably lower” in the hypotheses, and Reviewer 2 asks whether this could be changed to “significantly lower”. Can you please clarify the wording of these hypotheses on Page 4?

Reviewer 1, 2 and 3 note a lack of clarity with regards to the recruitment timeline and strategy. The recruitment timeline could be included within a table and, with your concerns around word count, perhaps this could be included as supplementary material on your OSF page.

Reviewer 4 makes a very good suggestion about conclusions – could the findings suggest the abandonment of such constructs altogether rather than any attempts at improved conceptualisation?

In my own assessment of your manuscript, I noted the following which should be addressed in a revision:



The Abstract is a little unclear in parts, e.g., the term “related ontologically diverse screening instruments” is difficult to read – could the word ‘related’ be removed here? And “each of which representing a different ontological basis” – is this meant to be “each of which represent”?

Page 2, Introduction: What do you mean by “essence” when outlining the research questions? This term isn’t used elsewhere until a lot later, so this can be confusing for the reader. I also recommend changing the word ‘good hypotheses’ to ‘informed hypotheses’ or something similar.

What will be the conclusions, as like in your pilot research, some of the measures overlap (e.g., in prevalence rates) but one or more do not?

I am a little confused by the difference in H1 and H2: if you expect the ICD-11 and DSM-5 based “gaming disorder” prevalence rates to be lower than DSM-IV (H1), then how does it follow that the ICD-11 and DSM-5 will overlap with DSM-IV in essence (H2)? Perhaps this is just my misunderstanding, so I am happy for you to just respond to this in your response to reviewers if it is. For H3, you state that this is compared to the general population – where are you getting the estimates for the general population?

In the Introduction, H1 specifies that ICD-11 and DSM-5 will have lower prevalence rates than DSM-IV, however, in the Method section, a reiteration of this hypothesis also mentions THL1. These should be consistent throughout.

Page 7, “If neither mental nor physical effects are nonsignificant or below $d=.22$, we consider H3a/H3b not supported”, do you mean “significant” here? Can you also provide a reference to equivalence testing so a reader could read up on this to understand this (relatively new) technique (e.g., will you be using guidance by Lakens, 2017?).

Page 9, you state that it would be “impossible for us to collect a representative sample” – is this based on resources (too large sample size)? Please clarify why this is impossible (e.g., “the number of resources required makes it impossible.”).

The hypotheses are outlined in both the Introduction and Methods section, but this is a little confusing because the Methods further outline sub-hypotheses (e.g., H1a-H1d etc.). There needs to be consistency. One option is to move all hypotheses to the Introduction and remove these from the Method, and the other option is to remove from Introduction and just have in the Method. In the former, the Method could then simply refer back to these hypotheses, e.g., “To test the hypotheses of XX to XX, we will..”. I will leave the authors to decide what is best to do here.

In the Method, what do you mean by “nothing will be corroborated”: do you mean that the results (prevalence) would be inconclusive?

I also note that you would like to submit to a particular journal upon Stage 2 acceptance – please note that the word counts for RRs at this journal include the references section too. I was unable to check the word count due to your document being in an online PDF, but I just wanted to notify you of this. I think you could remove the code detail within Page 9 H3 – if you make your analytical code available (which is a mandate to adhere to TOP guidelines), then readers/reviewers can see the calculations/code there instead.

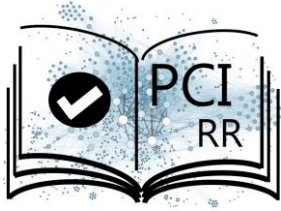
Please can you add the funding information to the title page of the manuscript.

Best wishes,

Charlotte Pennington - Recommender at PCI RR

Reviewed by Linda Kaye, 18 Oct 2021

[Download the review](#)



Reviewed by [Daniel Dunleavy](#), 03 Nov 2021

I thank the authors for their submission and the recommender for the opportunity to review this interesting submission. I hope the following comments, suggestions, and questions help strengthen and clarify components of this submission:

General Comments and Summary:

1. The authors are proposing to study how different screening instruments (for gaming-related health problems) differ and perform in identifying risk groups. The authors astutely note that gaming-related health problems have been cashed out in a variety of different ways, each with unique ontological commitments. Each of the four measures discussed here, corresponds to a different one of these commitments.

Background and Justification:

1. The overarching research question and three sub-questions are clear and reasonable. The authors rightly note that the instruments may yield differing prevalence rates (**RQ-A**), but also those who are identified as such (**RQ-B**). Interestingly, the authors are proposing (**RQ-C**) to explore how different screening instruments differ with respect to the health statuses of those identified as being at risk.

Design/Methodology:

1. I found each of the four sub-hypotheses for H1 to be clear (p. 4). I was initially hesitant, finding them to be poorly defined (i.e., what "remarkably lower" meant). However, I found the description of the interval-based method on p. 5 to be sufficiently clear to satisfy my concerns. I feel that the method used here is described in enough detail to be replicated by another set of researchers and further that it is an adequate method for assessing the hypotheses. The other two hypotheses (H2 and H3) are clear and testable - particularly H3, which clearly specifies the smallest effect size of interest (and justifies its selection).

2. Recruitment - p. 5 - The authors are recruiting using a company called Bilendi. I'd just like to see a little more detail about the recruitment timeline (expected start-finish dates) and how participants will 1) be incentivized, and 2) complete the survey (i.e., what software or platform will they use to complete the survey?).

3. Sampling plan - pp. 7-8 - The authors provide a reasonable sample size justification and the required sample size to estimate prevalence rates.

Other Comments:

I don't have any other concerns at this time. I thank the authors for their clearly written Stage 1 submission and the recommender for their consideration of the above review.

Reviewed by anonymous reviewer, 01 Nov 2021

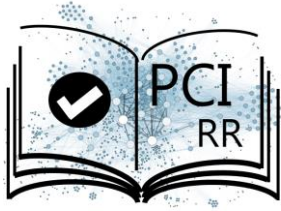
This proposed study presents a timely and important investigation into the ontological differences between screening instruments for gaming disorder, and the effects on prevalence. The following sections outline how the study adheres to the key questions by PCI RR:

1. Does the research question make sense in light of the theory or applications? Is it clearly defined? Where the proposal includes hypotheses, are the hypotheses capable of answering the research question?

The objectives and hypotheses outlined within are clearly defined.

2. Is the protocol sufficiently detailed to enable replication by an expert in the field, and to close off sources of undisclosed procedural or analytic flexibility?

The methods section largely contains enough detail that replication would be possible. Sources of ambiguity may include how and from where participants will be recruited via the 3rd party (online advertisements), and



how the survey data will actually be collected; i.e., whether in-person or online. These considerations may impact overall representativeness.

3. Is there an exact mapping between the theory, hypotheses, sampling plan (e.g. power analysis, where applicable), preregistered statistical tests, and possible interpretations given different outcomes?

Where possible, the hypotheses and proposed testing frameworks are adequately covered. The requisite sample size has been calculated appropriately. pre-registered tests and interpretations have been sufficiently provided.

4. For proposals that test hypotheses, have the authors explained precisely which outcomes will confirm or disconfirm their predictions?

In most cases, explanations are adequate. There may be questions about what constitutes *exploratory* investigation in some tests, but the authors have justified the need for exploratory analyses where necessary.

5. Is the sample size sufficient to provide informative results?

Sample size appears to have been calculated appropriately.

6. Where the proposal involves statistical hypothesis testing, does the sampling plan for each hypothesis propose a realistic and well justified estimate of the effect size?

The effect sizes proposed are realistic and justified.

7. Have the authors avoided the common pitfall of relying on conventional null hypothesis significance testing to conclude evidence of absence from null results? Where the authors intend to interpret a negative result as evidence that an effect is absent, have authors proposed an inferential method that is capable of drawing such a conclusion, such as Bayesian hypothesis testing or frequentist equivalence testing?

All hypotheses have been clearly defined and tests of significance are adequate.

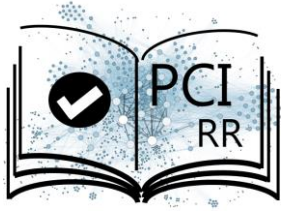
8. Have the authors minimised all discussion of post hoc exploratory analyses, apart from those that must be explained to justify specific design features? Maintaining this clear distinction at Stage 1 can prevent exploratory analyses at Stage 2 being inadvertently presented as pre-planned. Have the authors clearly distinguished work that has already been done (e.g. preliminary studies and data analyses) from work yet to be done?

All post hoc explanatory analyses have been minimised where possible. The results from the pilot study have been sufficiently distinguished.

9. Have the authors prespecified positive controls, manipulation checks or other data quality checks? If not, have they justified why such tests are either infeasible or unnecessary? Is the design sufficiently well controlled in all other respects?

In some cases, for example the case of mischievous reporting has been accounted for.

10. When proposing positive controls or other data quality checks that rely on inferential testing, have the



authors included a statistical sampling plan that is sufficient in terms of statistical power or evidential strength?

All proposed statistical analyses are adequate.

11. Does the proposed research fall within established ethical norms for its field? Regardless of whether the study has received ethical approval, have the authors adequately considered any ethical risks of the research

All ethical considerations have been met.

One future suggestion, that may be worth bearing in mind, is the digital convergence between gaming and gambling, with microtransactions and other predatory features that resemble gambling, including lootboxes. It strikes me that, given the overlap between gambling and gaming disorders in the DSM-IV and DSM-5 criteria, there might be the potential for confounding among participants. Controlling for this possibility might be something to think about.

Reviewed by [David Ellis](#), 01 Dec 2021

Identifying Gaming Disorders by Ontology:
A Nationally Representative Registered Report

This is a really interesting study that has been carefully considered. Beyond quantifying ontological overlap, I especially appreciate how the authors plan to compare mental and physical health differences between classifications from different measures. My comments are therefore relatively minor and perhaps reflect my own lack of understanding rather than any major issue.

1. It would be useful if the Abstract and Introduction made it clear what 'risk groups' might look like. I assumed these were groups whereby their scores suggest some form of diagnosis/classification? However, as I read on I realised this is also about other aspects of physical and mental health. Is a 'risk group' both of these? A table summarising all measures would be helpful within the method section to provide further clarity in this regard.

2. The research questions could also do with some clarification or possible re-ordering. 'Who' really refers to overlap (or lack of ontological overlap) in the sample as I understand it. It feels like it should be the first research question. The second question could then consider prevalence as this is making more population-based inferences (I think). I am not familiar with the interval-based method outlined so can't comment further. Wondered a bit as to how this goes beyond a series of chi-squared tests that could also consider ontological similarities. Research question C is clearly about Health and that is clear.

3. I suspect the results, regardless of direction, will have implications for many notions of technological 'addiction' or related 'disorders'.
If ontologies are different, then this suggests a problem for comparisons across different studies. If they are all the same, then this doesn't only indicate that scholars should direct efforts toward assessing the clinical relevance of multiple constructs as suggested. On the contrary, it begs the question as to why researchers have continued to re-invent the wheel and made little progress regarding a consensus. Creating a new disorder or ontology doesn't create new resources for practitioners or clinics. What are the societal costs of stigmatising the most popular form of play in children and adolescents? I guess what I am getting is that results from this work might suggest the abandonment of such constructs altogether rather than any attempts at improved conceptualisation.