Supplementary Research Question 3 Analyses Exploring the Validity of Self-Report to Assess Internet Gaming Disorder

A key feature of the DSM-5 guidance on Internet Gaming Disorder is that diagnosis can be made, in part, on the endorsement indicators of problem gaming. The implicit idea behind this approach is that these criteria equally contribute toward the diagnosis of Internet Gaming Disorder. Statistically, this is the assumption of a Rasch model in item response theory,[29] and can be tested by examining the model fit of a factor model with equal factor loadings. As an exploratory analysis, we applied a factor model with equal factor loadings and dichotomous outcome variables to the data. Results used a maximum likelihood estimation with robust standard errors. The Rasch model showed very good fit to the data, with significant factor loadings across all the studies (0.75, 0.77, 0.70, and 0.70, respectively). Findings supported the validity of these nine items to assess Internet Gaming Disorder (see Supplementary Table 1). Further, multi-group analyses examined the equivalence of factor loadings across gender and countries. Across all studies, the Rasch model with equal factor loadings and intercepts (item difficulty) across gender showed good fit to the data, Comparative Fit Index = 0.97-0.99, Tucker Lewis index = 0.97-0.99, Root Mean Square Error of Approximation = 0.017-0.029.

The same model showed equivalent factor structures across the four countries (US, UK, Canada, and Germany) in Study 3, χ^2 (140) = 408.3, Comparative Fit Index = 0.97, Tucker Lewis index = 0.97, Root Mean Square Error of Approximation = 0.028, suggesting that items assessed Internet Gaming Disorder with the same sensitivity and difficulty across these countries. We also examined whether distributions of item endorsements were equivalent comparing those with and without distress among those who endorsed five or more than five indicators. Results did not show differences in the profile of endorsement at this level with and without distress, χ^2 (8) = 2.23, 0.67, 14.7, and 4.94 (Studies 1-4), suggesting that items assessed Internet Gaming Disorder with the same sensitivity and difficulty across the distress criterion.

Supplementary Research Question 4 Behavioral and Clinical Impact of Internet Gaming Disorder

Behavioral impact. Given that Internet Gaming Disorder is thought to have a practically significant influence on functioning, akin to psychiatric disorders, we tested a preregistered hypothesis that those meeting the diagnostic threshold would show more frequent gaming, and less frequent physical exercise (physical activity) and quality social time with others (social activity), as compared to those who did not. A series of one-way Bayesian t-tests using a default Cauchy prior of 0.707 for the effect size of the alternative hypothesis tested confirmatory relations between Internet Gaming Disorder and behavioral engagement with games and physical and social activity.[30,31] Bayesian t-test was selected for our registered analysis plan in place of null hypothesis testing because it quantifies the relative evidence for the alternate hypothesis with moderately sized effects compared to the null.[32] In line with best practices, if observed Bayes factors were 3 or above we considered our hypotheses to be supported, if Bayes factors were 1/3 or below we considered the null hypothesis to be supported, and if Bayes factors observed were between 1/3 and 3 we considered the results inconclusive.[33] Results strongly supported the hypothesis that Internet Gaming Disorder was significantly linked to higher levels of regular gaming (Bayes Factor = 11.29), showing that engagement levels were higher for those meeting the Internet Gaming Disorder threshold (M = 4.00, SD = 1.04) than not (M = 2.80, SD = 1.67). Analyses concerned with physical activity were not conclusive (Bayes Factor = 1.709). Analysis concerned with social activity provided the strong evidence for the null hypothesis (Bayes Factor = 0.161). Examination of means indicated that those meeting the Internet Gaming

Disorder threshold reported overall lower levels of physical (*M* = 2.92, *SD* = 1.49, vs. *M* = 3.26, *SD* = 1.39) yet higher social activity (*M* = 3.92, *SD* = 1.00, vs. *M* = 3.61, *SD* = 1.14).

Bayesian regressions using model averaging[34] evaluated the relative significance of Internet Gaming Disorder as a predictor of everyday behaviors. Models considering Internet Gaming Disorder were compared to those without Internet Gaming Disorder. We also included age, sex, education level, and household income as controlling variables. Results provided equivocal to strong evidence favoring the null hypothesis (Bayes Factor = 0.495) for both physical (Bayes Factor = 0.095), and social (Bayes Factor = 0.091) activity, results which are consistent with the previous findings.

Clinical impact. In line with the registered analysis plan, one-way Bayesian t-tests quantified the evidence for a hypothesis that Internet Gaming Disorder would link to lower levels of mental, physical, and social health. Results from these analyses indicated that the data favored the null hypothesis for mental, physical, and social outcomes (all Bayes Factors < .283). Examination of means indicated that those meeting the Internet Gaming Disorder threshold reported marginally lower levels of mental health (*M* = 2.77, *SD* = 1.01 vs. *M* = 2.78, *SD* = 1.01), and marginally higher levels of physical (*M* = 2.33, *SD* = 1.23 vs. *M* = 2.31, *SD* = 0.94), and social health (*M* = 2.64, *SD* = 1.03 vs. *M* = 2.23, *SD* = 0.96). Further, Bayesian regressions controlling for age, sex, education level, and household income also provided evidence favoring the null hypothesis for Internet Gaming Disorder (all Bayes Factors < .171). Considering the results together, findings indicated that Internet Gaming Disorder was not linked to behavioral or clinical outcomes in the direction hypothesized.

Sensitivity analysis. We conducted exploratory analyses to test the robustness of the relations between Internet Gaming Disorder and observed outcomes. To do so, we tested the significance of a range of effect sizes smaller and larger than what was predicted relating Internet Gaming Disorder to outcomes (see Supplemental Analyses and Supplemental Figure

1). For gaming behavior evidence strongly supported our hypothesis. For all other outcomes the null hypothesis was supported. These sensitivity analyses provided further evidence there were no reliable relations between Internet Gaming Disorder and behavioral and clinical outcomes.

Sensitivity analysis. We conducted exploratory analyses to test the robustness of relations between Internet Gaming Disorder and observed outcomes. To do so, we tested the significance of a range of effect sizes smaller and larger than what was predicted relating Internet Gaming Disorder to outcomes (Supplemental Figure 1).For gaming behavior evidence strongly supported our hypothesis, for all other outcomes the null hypothesis was supported. In the most favorable conditions for the alternative hypothesis, where the prior effect size specified approached zero, Bayes factors approached a value of 1. These sensitivity analyses provided further evidence for relations between Internet Gaming Disorder and behavioral and clinical outcomes at levels different than those hypothesized were not significant.

References

1. Lord FM. *Applications of Item Response Theory to Practical Testing Problems*. Routledge; 1980.

2. Morey RD, Chambers CD, Etchells PJ, et al. The Peer Reviewers' Openness Initiative: incentivizing open research practices through peer review. *R Soc Open Sci*. 2016;3(1):150547. doi:10.1098/rsos.150547.

3. Rouder JN, Morey RD, Speckman PL, Province JM. Default Bayes factors for ANOVA designs. *J Math Psychol*. 2012;56(5):356-374. doi:10.1016/j.jmp.2012.08.001.

4. Morey RD, Romeijn J-W, Rouder JN. The philosophy of Bayes factors and the quantification of statistical evidence. *J Math Psychol*. January 2016. doi:10.1016/j.jmp.2015.11.001.

5. Wetzels R, Matzke D, Lee MD, Rouder JN, Iverson GJ, Wagenmakers E-J. Statistical evidence in experimental psychology an empirical comparison using 855 t tests. *Perspect Psychol Sci*. 2011;6(3):291-298.

6. Love J, Selker R, Marsman M, et al. *JASP*.; 2016.

Supplementary Table 1. Rasch Equal Factor Loading Models for Internet Gaming Disorder Indicators by Study

Study	χ^2	*df*	Comparative Fit Index	Tucker-Lewis Index	Root Mean Square Error for Approximation
1	64.6	36	0.98	0.98	0.026
2	70.9	36	0.99	0.99	0.020
3	286.4	36	0.98	0.98	0.027
4	72.7	36	0.99	0.99	0.017

Supplemental Figure 1.