

Structure and Stability of Internet Gaming Disorder from Childhood to Late

Adolescence: A 5-Wave Birth Cohort Study

Lars Wichstrøm^{1,2} | Beate W. Hygen³ | Daria J. Kuss⁴ | Vasileios Stavropoulos⁵ | Věra Skalická¹ | Rubén Rodríguez-Cano¹ | Ane-Kristine Øien¹ | Frode Stenseng⁵ | Silje Steinsbekk¹

¹ Department of Psychology, Norwegian University of Science and Technology, Trondheim, Norway

² Department of Child and Adolescent Psychiatry, St Olavs Hospital, Trondheim, Norway

³ Department of Diversity and Inclusion, NTNU Social Research, Trondheim, Norway

⁴ International Gaming Research Unit, Cyberpsychology Research Group, Psychology Department, Nottingham Trent University, UK

⁵ Department of Psychology, RMIT University, Australia

⁵ Department of Education and Lifelong Learning, Norwegian University of Science and Technology, Trondheim, Norway

Short title: Stability of Internet Gaming Disorder

Correspondence: Lars Wichstrøm, Department of Psychology, Norwegian University of Science and Technology, 7471 Trondheim, Norway. Email: lars.wichstrom@ntnu.no

Funding information: The Research Council of Norway, Grant number 334119; The Liaison Committee between Central Norway RHA and NTNU, Grant number 2024-36863.

Word count: 4,012

Declaration of competing interests: The authors have no competing interests to declare.

Abstract

Background and Aims: The natural course and longer-term stability of Internet gaming disorder (IGD) and its symptoms are largely unknown, yet such knowledge is essential to aid prevention, treatment, and policy development. This study therefore examined four aspects of IGD symptom stability: of form, at the group level, rank-order, and of within-person changes.

Design: Birth-cohort study.

Setting: Trondheim, Norway

Participants: A stratified sample ($n=812$, 50.9% girls) of the 2003/2004 birth cohorts was examined biennially over five waves from age 10 to 18 ($n=3,236$ observations).

Measurements: IGD was assessed through a semi-structured clinical interview. Latent growth curves were employed to assess the stability of symptoms at the group level, and random intercept cross-lagged models were used to determine how changes in symptoms predicted subsequent changes in symptoms at the within-person level.

Findings: IGD symptoms loaded on two factors, termed 'Heavy involvement' and 'Negative consequences,' consistently across ages and sex. Strong measurement invariance was partly achieved across ages. The point prevalence of IGD diagnosis varied between 1.1%-2.2%, and 5.6% (95% confidence interval [CI]: 4.0, 7.3) had an IGD diagnosis at least once across ages 10-18 (boys: 10.2%, CI: 6.9, 13.5; girls: 1.6%, CI: 0.5, 2.7). Symptoms capturing Heavy involvement increased slightly from age 10-16 and then declined sharply at age 18, whereas symptoms representing Negative consequences remained stable. Regarding stability relative to the group, moderate two-year stability was observed (Heavy involvement $r=.33$ to $.56$, $p<.001$; Negative consequences $r=.19$, $p<.01$ to $.37$, $p<.001$) with increasing stability with age. The stability from age 10 to 16 and from 10 to 18 was weak for Heavy involvement ($r=.14$ and $.17$, $p<.05$) and absent for Negative consequences ($r=.09$ to $.11$, $p>.05$). Increases

in Heavy involvement predicted further increases in Heavy involvement two years later ($\beta=.17, p<.01$ to $.39, p<.001$) and increased Negative consequences at ages 14 and 18 ($\beta=.29$ and $.28, p<.001$).

Conclusions: Although the point prevalence of Internet gaming disorder (IGD) diagnosis in the 2003/2004 birth cohorts of Trondheim, Norway was low, 10% of boys fulfilled the criteria for a diagnosis at least once during preadolescence and adolescence. IGD symptoms at age 10 were only weakly related to IGD symptoms in mid- and late adolescence. However, as Heavy involvement symptoms predicted later increases in Negative consequences symptoms from age 12 onwards, early adolescence may offer a window of opportunity for when interventions might intersect a development toward a full-blown IGD diagnosis before symptoms become more intertwined with additional problems and resistant to change.

Keywords: Adolescents, Cohort, Internet gaming disorder, Longitudinal, Prospective, Stability, Structure, Within-person

Internet gaming disorder (IGD) was proposed as a condition requiring further study in the Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition (DSM-5) (1), and gaming disorder is now fully included in the International Classification of Diseases (ICD-11) (2). For prevention, treatment, policy development, and to aid a decision on including IGD in the DSM-6, the natural course of IGD needs to be identified. A recent review and meta-analysis of prospective studies, typically spanning six months to two years, reported standardized stability coefficients varying from .13 to .92, but with most falling in the .40 to .60 range (3) and that 42% (95% CI: 33, 50) exhibiting problematic gaming, evinced this one to two years later. Of note, although clinical interviews are considered the gold standard for assessing psychiatric disorders (e.g., 4), including IGD (5), research cited above (except one(6)) have applied self-report questionnaires, which tend to produce higher stability coefficients than clinical interviews, for a variety of reasons, including common rater bias inflating stabilities (5). Therefore, the stability of IGD and its symptoms, when assessed through clinical interviews, remains largely unknown.

Some short-term longitudinal studies have assessed whether IGD measures capture that same construct across age (i.e., measurement invariance), yielding mixed results (e.g., 7, 8). Moreover, despite a burgeoning of research on IGD during the past decade, little is known about the trajectory and stability of IGD within the age span when gaming emerges as one of the most important pastimes—especially among boys—(i.e., middle childhood) to when it tapers off for the majority (i.e., late adolescence) (9). IGD was portrayed in the DSM-5 as one coherent disorder without any types/specifiers. Most studies on its structure using short checklists/questionnaires show a one-factor solution, e.g., (10, 11), but several studies consider a two-factor structure. Examining gamer heterogeneity using The Internet Gaming Disorder Scale, the authors found evidence of two types of gamers: risky gamers and disordered gamers (12). Moreover, psychometric evaluation of the Gaming Disorder Scale for

Adolescents showed two dimensions, cognitive-behavioral symptoms and negative consequences (13, 14). Finally, when applying a diagnostic interview, one community study also found a two-factor solution, with the factors being ‘Heavy involvement’ (i.e., preoccupation with gaming, withdrawal symptoms, tolerance, unsuccessful attempts to control gaming, and escape from negative moods) and ‘Negative consequences’ (i.e., loss of interest in other activities, continued use despite psychosocial problems, deception, and jeopardized or lost relationships or opportunities) (15). This factor structure aligns with most dedicated gamers not developing substantial problems (16). In other words, gaming preoccupation/engagement does not necessitate functional impairment. Overall, evidence suggests that a simple unifactorial model may not fully capture IGD complexity. Therefore, IGD symptom stability and progression may vary between individuals who are highly engaged in gaming and those who experience negative consequences from gaming—which has yet to be thoroughly explored.

Scholars have identified several conceptualizations of mental health problem ‘stability’ (17). Here, we address four facets of stability: i) of form (i.e., IGD symptoms capture the same phenomenon across ages—namely, longitudinal measurement invariance), ii) at the population level (i.e., mean levels of symptoms and IGD prevalence over time), iii) relative to others in the population, and iv) of within-person changes (i.e., whether changes predict changes at the individual level). These differing conceptualizations carry distinct implications for diagnosis, prevention, and treatment. For instance, if the IGD construct changes throughout developmental stages (e.g., some symptoms become more important), factors that trigger and maintain it may also shift, necessitating age-appropriate intervention strategies. Moreover, low individual-level stability at certain ages may suggest IGD tends to subside naturally during those periods, indicating prevention efforts should target other developmental windows. Additionally, if changes in IGD are stable over time (e.g.,

reductions predict future reductions) or if increased high involvement in gaming predicts greater negative consequences, prevention and treatment may need to focus on reducing excessive gaming involvement. Finally, boys outnumber girls in the prevalence of IGD (18), but the question of whether the stability of IGD symptoms also differs is underexplored.

In sum, research on the stability of changes, the stability of differing dimensions of IGD, and all types of longer-term stability is lacking. To fill these gaps, we will investigate the structure of IGD and four aspects of its long-term stability across childhood and adolescence while exploring sex differences.

METHODS

Participants and Procedure

Parents and children of the 2003/2004 birth cohorts living in Trondheim, Norway, were asked to participate in the Trondheim Early Secure Study (TESS) (19). A letter of invitation, including the Strength and Difficulties Questionnaire (SDQ) 4-16 version (20), a screening assessment for emotional and behavioral problems, was sent to the participants' homes. Parents brought the completed SDQ to the scheduled routine health checkup for 4-year-olds. Of the 3,456 invited parents, 97% ($n = 3,358$) attended the checkup. Parents who lacked sufficient proficiency in Norwegian to complete the SDQ were excluded ($n = 176$), and 166 parents were missed being asked to participate. Health nurses obtained parental consent for participation ($n = 2,477$; 82.2%). To enhance variability and thus power, children with emotional and behavioral problems were oversampled. Children were divided into four strata according to their SDQ scores (cutoff points 0–4, 5–8, 9–11, and 12–40). The probability of being drawn into the study increased with increasing SDQ scores (.37, .48, .70, and .89, in the respective strata). This oversampling of children with emotional and behavioral problems was adjusted for in the analyses to arrive at correct population estimates.

In all, 1,250 families were drawn to participate. Of these, we obtained information from 1,007 in the first wave (T1; $M_{\text{age}}=4.7$, $SD=0.3$; 50.9% girls). The participants were reassessed biennially until age 18 (See Online Figure 1 for more details). At age 16, the children provided written informed consent. The Regional Committee for Medical and Health Research Ethics, Mid-Norway, approved the study.

Measures

Internet gaming disorder. The Internet Gaming Disorder Interview (IGDI) was used to assess DSM-5-defined symptoms of IGD (15). The IGDI is an interviewer-based psychiatric interview of youth, where the interviewer poses mandatory and optional follow-up questions and probes until enough information is obtained to decide whether a symptom has been present during the previous 12 months. Interviews were conducted face-to-face by interviewers with a higher degree in health or education sciences and prior extensive experience in conducting psychiatric interviews of children and adolescents. At ages 10 and 18, 88 and 102 audio-taped interviews, respectively, were randomly drawn and re-coded by raters who were blind to all information about the participants. The reliability of IGD diagnosis was $\kappa=.65$ at age 10 and $\kappa=.95$ at age 18, whereas the reliability for the number of symptoms was $ICC=.90$ at age 10 and $ICC=.93$ at age 18.

Analysis plan

Analyses were conducted in Mplus 8.11 using a robust maximum likelihood estimator. A detailed description of planned analyses is available in the Online supplement, along with Mplus codes for additional analyses. A full information maximum likelihood (FIML) procedure was applied to handle missing data. Hence, missing data were not imputed, but all available data were used. FIML is asymptotically equivalent to multiple imputation, and with the present sample size, FIML and multiple imputation are expected to yield very similar results (21). Population weights were applied to correct for initial oversampling, corresponding

to the number of children in a stratum divided by the number of participants in that stratum, applying a sandwich estimator. The factor structure of the IGD I will first be examined with an exploratory factor analysis (EFAs) for up to four factors across all time points. The resulting preferred factor solution will be compared to the solution with one less factor, at each time point and across all time points through a confirmatory factor analysis (CFA). Gender differences in factor loading will be examined in a multigroup analysis of pooled data.

Three aspects of stability of form (measurement invariance) were examined: basic (analogous to configural invariance with continuous data—i.e., symptoms informing the same factors), equal slopes (analogous to metric—i.e., symptoms load equally onto the factors, indicating that they contribute with the same strength to the identified dimensions) invariance, and equal slopes and thresholds (analogous to scalar invariance—i.e., same scores in the latent construct indicating the same level of severity). To assess whether invariance is achieved, the difference in χ^2 between nested models will be compared. However, given that $\Delta\chi^2$ between models depends on sample size, we relied on recent recommendations for differences in fit indices between models involving categorical data and two or more factors (22): a good-fitting model for basic invariance; Differences in Root Mean Square Error of Approximation (ΔRMSEA) $\leq .05$ in conjunction with significant $\Delta\chi^2$ for equal slopes invariance; and $\Delta\text{RMSEA} \leq .01$ and differences in the Comparative Fit Index (ΔCFI) $\geq -.004$ for equal slopes and thresholds invariance. To describe the development of symptoms at the group level, latent growth curves were fitted and compared using the Satorra-Bentler test. Stability relative to the group was estimated with Pearson's correlations. To investigate whether changes in the two IGD dimensions predict further changes in such symptoms, we fitted a random intercept cross-lagged panel model (RI-CLPM) (23). In this model, between-person associations between Heavy involvement and Negative consequences are filtered out, so the cross-lagged predictions from Heavy involvement to Negative consequences, and vice versa, are at the within-person

level, thus estimated net of all unmeasured time-invariant confounding effects. Of note, the analyses were not pre-registered, and the results should be considered exploratory.

RESULTS

As IGD was assessed from age 10, children were included over five waves: age 10 ($n=702$; $M_{\text{age}}=10.51$, $SD=0.17$; 52.3% girls), 12 ($n=668$; $M_{\text{age}}=12.49$, $SD=0.15$; 51.9% girls), 14 ($n=628$; $M_{\text{age}}=14.33$, $SD=0.59$; 53.0% girls), 16 ($n=666$; $M_{\text{ag}}=16.98$, $SD=0.31$; 55.0% girls), and 18 ($n=606$; $M_{\text{age}}=18.6$, $SD=0.24$; 56.0% girls). A total of 812 children provided valid data on at least one occasion ($n=3,236$ observations), forming the analytical sample. Participants were predominantly of Norwegian descent (91.0%), with 5.8% from Europe, the USA, Canada, Australia, or New Zealand, and 3.2% from Africa, Latin America, or Asia.

Structure of IGD symptoms

An EFA with up to four factors did not converge. Details on the three-factor solution are shown in Supplementary Text 2 (Online). This solution had a better fit than a two-factor solution, $\Delta\chi^2=58.70$, $p<.001$. However, according to both Kaiser's criterion and a scree test, the eigenvalues supported a two-factor solution (4.63, 1.11, 0.95, 0.67, and 0.61 for the first five factors). Moreover, the three-factor solution contained several cross-loadings and a factor that was not clinically meaningful. The two-factor solution was therefore preferred. In a CFA, the previously reported two-factor solution consisting of Heavy involvement and Negative consequences (15) fitted the data better than a one-factor solution, for all ages pooled, at each individual age, and for boys and girls (Table 1). Factor loadings were moderately high and of similar magnitude, except for jeopardized or lost relationships or educational career opportunities, which loaded particularly high on the Negative consequences factor (Table 2). A model where the factor loadings among boys and girls were fixed to be equal was compared to a model where they were freely estimated; the freely estimated model fitted the

data best ($\Delta\chi^2=86.67$, $df=1$, $p<.001$), indicating that the factor loadings differed between the two sexes. Inspecting the factor loadings revealed that tolerance, unsuccessful attempts to control, and escapism seemingly loaded higher among girls than boys. Setting these free and fixing the remaining symptoms to load equally among boys and girls resulted in a model that did not have a worse fit than a freely estimated model ($\Delta\chi=0.87$, $df=6$, $p=.990$), which indicates that these loadings were indeed higher among girls, whereas the others were similar between the sexes.

Stability of form

Upon inspecting the prevalence of individual symptoms, withdrawal symptoms were rare in adolescence (Online Supplementary Table S1), and CFAs including this symptom did not converge, possibly due to its low frequency. Hence, withdrawal was omitted when performing the remaining CFAs. A CFA model was constructed to estimate the latent Heavy involvement and the Negative consequences at each age. The Heavy involvement factors at each age were allowed to covary, as were the Negative consequences factors, and concurrent covariation between the two was also allowed. In a basic model estimation, loadings and thresholds were freely estimated. This model fitted the data well (Table 3). Although the difference between this basic model and a model where loadings were freely estimated was significant (Table 3), the recommended differences in fit indices between models were satisfactory, suggesting loading (metric) invariance. Fixing thresholds to be equal across time deteriorated fit. However, as seen in Supplementary Table S2, the thresholds for symptoms #3 (tolerance) and #4 (unsuccessful attempts to control) seemingly increased with age. Setting these thresholds free and fixing the remaining thresholds to be identical across age resulted in a (partial) strong measurement invariance (Table 3), suggesting more Heavy involvement was needed for youths' categorization as having developed tolerance/unsuccessfully trying to control their gaming as they became older. Due to reduced sample size when analyzing each

sex at each age, measurement invariance across age and sex had to be assessed on pooled data employing the WLSMV estimator. Given the categorical variables and the required theta parametrization to estimate variances, this analysis is currently impossible. However, the CFAs among boys and girls (Table 2) suggested weak invariance was not achieved because some symptoms were more critical to the Heavy involvement construct among girls than boys.

The level of Negative consequences symptoms seemingly remained constant during the study period, whereas Heavy involvement tapered off at age 18 (Table 4). Indeed, a linear growth curve model with these two symptom categories did not fit the data well (Table 5), whereas a model including curvilinear development did. This model portrayed a linear increase and a convex quadratic growth in Heavy involvement (Table 6), but neither a linear nor a quadratic change in Negative consequences. In boys, Heavy involvement evinced a convex curvilinear development, whereas Negative consequences increased. A different pattern was found in girls, as linear declines were seen in Heavy involvement and Negative consequences, with no curvilinear development in Heavy involvement. A model with quadratic growth in Negative consequences did not converge.

Although the prevalence of the IGD diagnosis increased toward early adolescence and then decreased (Table 7), it was not statistically significant, as neither a linear ($M_{\text{slope}}=1.35$, $p=.212$) nor quadratic ($M_{\text{slope}}=-0.19$, $p=.177$) growth was seen in IGD diagnosis across development. In all, 5.62% (95% CI:3.96, 7.28) qualified for an IGD diagnosis at at least one time point during the ages 10–18 years, with boys more frequently affected (10.20%, CI:6.94, 13.45) than girls (1.64%, CI:0.54, 2.73).

Stability relative to the group

The two-year stability was moderate (Heavy involvement $r=.33$ to $.56$; Negative consequences $r=.19$ to $.37$, Table 8), with slightly increasing stability with age. The stability from age 10 years to adolescence was weak or absent. These patterns were similar for boys and girls (Supplementary Table S3).

Stability of within-person changes

The RI-CLPM of changes in Heavy involvement and Negative consequences fitted the data well ($\chi^2=26.25$, $df=21$, $p=.197$, RMSEA=.018, 90% CI:.000, .036, CFI=.992, Tucker-Lewis Index (TLI)=.982). Increases in Heavy involvement forecasted further increases two years later (Figure 1), a prediction that was not seen in Negative consequences, except for age 14-16. Increased Heavy involvement predicted increased Negative consequences from age 12-14 and 16-18. Moreover, Negative consequences at age 16 were also predicted by prior Heavy involvement at age 12 via an indirect effect involving the stability of Negative consequences from age 14-16 ($\beta=.05$, $p=.036$). A model where boys' and girls' cross-lagged paths were set to be equal did not fit the data well ($\chi^2=162.93$, $df=58$, $p<.001$, RMSEA=.067, 90% CI:.055-.079, CFI=.841, TLI=.754) and was worse than a freely estimated model ($\Delta\chi^2=57.94$, $df=16$, $p<.001$), indicating sex differences. However, inspecting model fit indices revealed a possible autoregressive path from Negative consequences from age 14-18 among girls, and including this path resulted in an adequate model fit that did not differ from the freely estimated model ($\chi^2=95.86$, $df=57$, $p=.001$, RMSEA=.041, 90% CI:.026-.055, CFI=.941, TLI=.907; $\Delta\chi^2=11.89$, $df=15$, $p=.688$) and with this path significant in girls ($\beta=.48$, $p=.012$) but not boys ($\beta=.13$, $p=.345$), with no other different paths in boys than girls.

DISCUSSION

Despite a burgeoning amount of research into IGD during the past decade, its natural course and longer-term stability are largely unknown. We therefore examined four aspects of

stability in a birth cohort sample that was followed over five waves from age 10-18. Using a clinical interview to assess IGD, symptoms loaded on two factors: Heavy involvement and Negative consequences. Regarding stability of form, strong measurement invariance was achieved across ages with one exception: With increasing age, more Heavy involvement was needed with age to be assigned tolerance and unsuccessful attempts to control gaming symptoms. The prevalence of IGD diagnosis varied between 1.1% and 2.2%, and 5.6% met the diagnostic criteria at least once during the study period. At the group level, Heavy involvement symptoms increased slightly from age 10-16, then declined sharply, whereas the frequency of Negative consequences symptoms remained at the same level. Considering stability relative to the group, moderate two-year stability was observed. However, there were only weak or no relations between the symptom number at age 10 and in late adolescence. Regarding the stability of changes, increasing Heavy involvement not only predicted future increases in Heavy involvement two years later, but also increased Negative consequences at ages 14, 16, and 18.

Stability of form

Our measurement invariance findings indicate that IGD predominantly captures the same phenomenon across childhood and adolescence, with exceptions being that the intercepts of tolerance and unsuccessful attempts to control gaming increased, meaning more Heavy involvement was needed to achieve these two symptoms. Tolerance inherently involves a progression toward higher levels of involvement, suggesting it naturally requires more gaming time to manifest as a symptom. Similarly, new attempts to control gaming (often unsuccessful) may be instigated as involvement escalates beyond prior levels, indicating a pattern tied to natural IGD progression rather than measurement flaws. The prevalence of withdrawal symptoms was very low in the present study, a finding in line with some questionnaire-based studies (24). Together with the uncertain clinical utility of

withdrawal symptoms for distinguishing healthy engagement from pathology in behavioral addictions in general (25) and in IGD in particular (26), this may point to omitting it in further revisions of the IGD in the DSM. Overall, although both main dimensions appear to inform IGD presentations, Negative consequences may be more indicative of diagnosable IGD levels, which supports the literature that proposes a refinement of IGD diagnostic criteria in favor of the Gaming Disorder's framework for greater accuracy (27). Finding a two-factor solution, with one representing being preoccupied with gaming and the other tapping into the aftermath of intensive gaming involvement, also corroborates the World Health Organization's demarcation of Hazardous Gaming from Gaming Disorder in ICD-11 (2).

Stability at the group level

A 1–2% point prevalence found in the present study is considerably lower than the 5–10% population estimates from meta-analyses derived from studies predominantly using brief screens or checklists, including measures of ‘video game dependency’ and ‘problematic online gaming’ (28, 29). When using such checklists, the adolescent must determine the required intensity and frequency (e.g., what qualifies as ‘losing interest in previous hobbies’ and how often is ‘often’ or ‘very often’?)—a respondent-based approach. Here, we used an interviewer-based approach, where trained interviewers prompted and queried until concluding whether a symptom was present at the required severity, intensity, frequency, and duration. Hence, it is possible that the different methodologies applied can explain the prevalence discrepancy.

Thresholds of tolerance and unsuccessful attempts to control gaming increased with age (above). The convex group-level development in Heavy involvement may hence reflect a normative development in which gaming initially allows connection with peers, especially

among boys, through shared activity. In late adolescence, however, gaming as an avenue to maintain social ties is possibly replaced by adult pathways, e.g., intimate relationships and increased school/work time. Nevertheless, tolerance and unsuccessful attempts to control gaming were among the most frequent symptoms in the Heavy involvement dimension, contributing the most to the overall score. However, they were also among the symptoms evidencing the sharpest decline from age 14, thereby contributing to the inverted U-shaped curve of Heavy involvement across age. Although increasing thresholds may reflect natural development, it should be acknowledged that it may have led to an overestimation of the decline in Heavy involvement from mid-adolescence.

Stability relative to the group

The finding of moderate two-year rank-order stability in IGD symptoms concurs with previous reports (30). However, weak or no relationships were found between symptom counts at age 10 and late adolescence, regardless of sex. Hence, IGD symptoms in middle childhood should cause less concern than later in development, as they are likely to subside or vanish by mid- to late adolescence.

Within-person stability

Increased Heavy involvement at ages 12-16 predicted future increases in both Heavy involvement and Negative consequences. This finding aligns with foundational addiction models, which propose that addictive behaviors, including IGD, often begin with preoccupation and high involvement (31). Initially, this involvement may provide a sense of satisfaction. Over time, however, it can evolve into a means of mitigating the negative feelings of prior overuse consequences. In this cyclic process, overuse typically precedes the emergence of adverse outcomes, and those consequences can, in turn, reinforce further overuse as individuals seek to avoid confronting them emotionally. In the present sample, this

development seemed to start in early adolescence (age 12), suggesting that successful interventions at this age, targeting heavy involvement in gaming, might intersect a development toward continued overuse and incurring negative consequences from it.

Sex differences

Although IGD was 3–5 times more prevalent in boys than girls, and symptoms tended to decline in girls while increasing in boys across development (despite a decline in Heavy involvement in late adolescence), most aspects of stability were similar across sexes. Furthermore, increased Heavy involvement predicted a higher risk of Negative consequences equally for both boys and girls. Therefore, when IGD symptoms are present in girls, they warrant the same level of attention and intervention as in boys.

Strengths and Limitations

Despite several strengths, such as involving a birth cohort followed over eight years from age 10-18 with five data collection waves, clinical interviews to assess IGD, and advanced statistical approaches to study within-person relations, we acknowledge some limitations. First, as the prevalence—and thus possibly the stability—of IGD differs between regions (23), generalizations to other countries should be made with caution.

Notwithstanding, it should be noted that the percentage of adolescents having access to a game console (47% in Norway vs. 52% in the UK and 44-53% in Germany—depending on age) and having one's own PC/Mac (56% in Norway, 60% in the UK and 36% in Germany) (32-34) were similar in Norway and several OECD countries. Second, we were unable to formally test full measurement invariance across both age and sex simultaneously, primarily due to the dichotomous nature of symptoms, which represents a methodological limitation. We therefore had to rely on exploring sex differences by examining factor loadings in the pooled sample and by modelling sex-specific developmental trajectories for each IGD

dimension, revealing meaningful sex differences. Future symptom factor structure research should therefore involve even larger samples. Third, although the extensive follow-up time frame, spanning 2013-2021, is an asset, period differences, e.g., changes in games played, e-sport evolution, may have influenced gaming prevalence and continuity, so that period changes affecting all child and adolescent gamers masquerade as developmental changes.

Conclusions

Although IGD diagnosis point prevalence in the population was low (1–2%), 10% of boys fulfilled diagnostic criteria at least once across ages 10-18. This finding, along with limited rank order and within-person stability, supports moderate short-term IGD symptom stability. In the long term, symptoms appearing in preadolescence (i.e., age 10) were only weakly related to IGD symptoms in mid- and late adolescence. However, from age 12 onwards, increased Heavy involvement symptoms predicted later increases in Negative consequences. Early adolescence may thus be a period when interventions may intersect a development toward a full-blown IGD diagnosis—before symptoms become intertwined with additional problems and more resistance to change.

References

1. American Psychiatric Association. Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition. Arlington, VA: American Psychiatric Association; 2013.
2. World Health Organization (WHO) International Classification of Diseases, Eleventh Revision (ICD-11): World Health Organization; 2019/2021. Available from: <https://icd.who.int/browse11>.
3. Sun A, Ho C, Kuss D, Cross C. The temporal stability of problematic gaming and gaming disorder: A systematic review and meta-analysis. *Addictive Behaviors Reports*. 2025;21.
4. Silverman W, Ollendick T. Evidence-based assessment of anxiety and its disorders in children and adolescents. *J Clin Child Adolesc Psychol*. 2005;34(3):380–411.
5. Pontes HM, Griffiths MD. Measuring DSM-5 internet gaming disorder: Development and validation of a short psychometric scale. *Comput Human Behav*. 2015;45:137–43.
6. Hong Y, Hwang H, Starcevic V, Choi T, Kim T, Han D. Which is more stable and specific: DSM-5 internet gaming disorder or ICD-11 gaming disorder? A longitudinal study. *Psychiatry Clin Neurosci*. 2023;77(4):213–22.
7. Stavropoulos V, Bamford L, Beard C, Gomez R, Griffiths M. Test-Retest Measurement Invariance of the Nine-Item Internet Gaming Disorder Scale in Two Countries: A Preliminary Longitudinal Study. *Int J Mental Health Addict*. 2021;19(6):2003–20.
8. Cosa I, Dobrean A, Balazsi R. Measurement Invariance of the Lemmens Internet Gaming Disorder Scale-9 Across Age, Gender, and Respondents. *Psychiatr Q*. 2024;95(1):137–55.
9. Borodovsky JT, Squeglia LM, Mewton L, Marsch LA. Longitudinal Use Patterns of Technology Subtypes During the Transition Into Early Adolescence: Results From the Adolescent Brain Cognitive Development Study. *J Adolesc Health*. 2024;75(5):809–18.
10. Stavropoulos V, Beard C, Griffiths MD, Buleigh T, Gomez R, Pontes HM. Measurement Invariance of the Internet Gaming Disorder Scale-Short-Form (IGDS9-SF) Between Australia, the USA, and the UK. *Int J Mental Health Addict*. 2018;16(2):377–92.
11. Király O, Slecza P, Pontes HM, Urbán R, Griffiths MD, Demetrovics Z. Validation of the Ten-Item Internet Gaming Disorder Test (IGDT-10) and evaluation of the nine DSM-5 Internet Gaming Disorder criteria. *Addict Behav*. 2017;64:253–60.
12. Lemmens J, Valkenburg P, Gentile D. The Internet Gaming Disorder Scale. *Psychol Assess*. 2015;27(2):567–82.
13. Paschke K, Austermann M, Thomasius R. Assessing ICD-11 Gaming Disorder in Adolescent Gamers: Development and Validation of the Gaming Disorder Scale for Adolescents (GADIS-A). *J Clin Med*. 2020;9(4).
14. Wu T, Huang S, Chen J, Kamolthip R, Kukreti S, Strong C, et al. Translation and Validation of the Gaming Disorder Test and Gaming Disorder Scale for Adolescents into Chinese for Taiwanese Young Adults. *Compr Psychiatry*. 2023;124.
15. Wichstrøm L, Stenseng F, Belsky J, von Soest T, Hygen BW. Symptoms of Internet Gaming Disorder in Youth: Predictors and Comorbidity. *J Abnorm Child Psychol*. 2019;47(1):71–83.
16. Kuss DJ, Pontes HM. Internet addiction. *Advances in Psychotherapy - Evidence-based practice* Göttingen, Germany: Hogrefe; 2019.
17. Morken IS, Viddal KR, Ranum B, Wichstrøm L. Depression from preschool to adolescence - five faces of stability. *J Child Psychol Psychiatry*. 2021;62(8):1000–9.

18. Ropovik I, Martoncik M, Babincák P, Baník G, Vargová L, Adamkovic M. Risk and protective factors for (internet) gaming disorder: A meta-analysis of pre-COVID studies. *Addict Behav.* 2023;139.
19. Steinsbekk S, Wichstrøm L. Cohort Profile: The Trondheim Early Secure Study (TESS)-a study of mental health, psychosocial development and health behaviour from preschool to adolescence. *Int J Epidemiol.* 2018;47(5):1401–+.
20. Goodman R, Ford T, Simmons H, Gatward R, Meltzer H. Using the Strengths and Difficulties Questionnaire (SDQ) to screen for child psychiatric disorders in a community sample. *Br J Psychiatry.* 2000;177:534–9.
21. Schafer JL, Graham JW. Missing data: Our view of the state of the art. *Psychol Methods.* 2002;7(2):147–77.
22. Svetina D, Rutkowski L, Rutkowski D. Multiple-Group Invariance with Categorical Outcomes Using Updated Guidelines: An Illustration Using Mplus and the lavaan/semTools Packages. *Struct Equ Modeling.* 2020;27(1):111–30.
23. Hamaker EL, Kuiper RM, Grasman R. A Critique of the Cross-Lagged Panel Model. *Psychol Methods.* 2015;20(1):102–16.
24. Kiraly O, Bóthe B, Ramos-Diaz J, Rahimi-Movaghar A, Lukavska K, Hrabec O, et al. Ten-Item Internet Gaming Disorder Test (IGDT-10): Measurement Invariance and Cross-Cultural Validation Across Seven Language-Based Samples. *Psychol Addict Behav.* 2019;33(1):91–103.
25. Kardefelt-Winther D, Heeren A, Schimmenti A, van Rooij A, Maurage P, Carras M, et al. How can we conceptualize behavioural addiction without pathologizing common behaviours? *Addiction.* 2017;112(10):1709–15.
26. Castro-Calvo J, King D, Stein D, Brand M, Carmi L, Chamberlain S, et al. Expert appraisal of criteria for assessing gaming disorder: an international Delphi study. *Addiction.* 2021;116(9):2463–75.
27. Király O, Potenza MN, Demetrovics Z. Gaming disorder: current research directions. *Current Opinion in Behavioral Sciences.* 2022;47.
28. Gao YX, Wang JY, Dong GH. The prevalence and possible risk factors of internet gaming disorder among adolescents and young adults: Systematic reviews and meta-analyses. *J Psychiatr Res.* 2022;154:35–43.
29. Fam JY. Prevalence of internet gaming disorder in adolescents: A meta-analysis across three decades. *Scand J Psychol.* 2018;59(5):524–31.
30. Peeters M, Koning I, van den Eijnden R. Predicting Internet Gaming Disorder symptoms in young adolescents: A one-year follow-up study. *Comput Human Behav.* 2018;80:255–61.
31. Griffiths M. A ‘components’ model of addiction within a biopsychosocial framework. *J Subst Abuse.* 2005;10(4):191–7.
32. Medietilsynet. Barn og medier 2022. In: Medietilsynet, editor. *Barn og medier.* Oslo: Medietilsynet; 2022.
33. LÖSUNGEN MiI-SU. Ab 10 Jahren haben die meisten Kinder ein eigenes Smartphone. 2024.
34. Ibbetson C. How many children have their own tech? In: YouGov, editor. 2020.

TABLE 1. Model testing of two correlated versus one-factor solutions for IGD symptoms.

	$\Delta\text{chi-sq}$	$\Delta\text{parameters}$	p-value
Across ages 10–18	54.18	1	<.001
Age 10	10.17	1	.001
Age 12	9.20	1	.002
Age 14	15.53	1	<.001
Age 16	9.68	1	.002
Age 18	20.83	1	<.001
Boys	22.12	1	<.001
Girls	6.12	1	.013

TABLE 2. Standardized factor loadings for a two-factor solution of Internet gaming disorder symptoms.

Internet gaming disorder symptoms	Total sample		Boys		Girls	
	Heavy involvement	Negative consequence	Heavy involvement	Negative consequence	Heavy involvement	Negative consequence
1. Preoccupation with Internet games	.67		.57		.60	
2. Withdrawal	.61		.52		.68	
3. Tolerance	.69		.57		.76	
4. Unsuccessful attempts to control	.67		.54		.73	
5. Loss of interest		.61		.78		.75
6. Continued excessive use despite psychosocial problems		.89		.87		.86
7. Deception		.77		.74		.79
8. Escapism or to relieve negative mood	.75		.65		.84	
9. Jeopardized or lost relationship or educational career opportunity		.92		.90		.89

TABLE 3. Longitudinal measurement invariance of a two-factor solution of Internet gaming disorder symptoms across ages 10–18 years.

	χ^2	<i>df</i>	CFI	Δ CFI	RMSEA	Δ RMSEA	$\Delta\chi^2$		
							χ^2	<i>df</i>	<i>p</i>
Basic invariance	740.53	695	.986		.009				
Loading invariance	790.38	719	.984	.002	.011	.002	51.77	24	<.001
Loading and threshold invariance	1012.67	743	.915	.069	.021	.010	573.90	24	<.001
Loading invariance and partly threshold invariance ¹	873.42	735	.956	.023	.015	.004	236.09	16	<.001

Note: CFI=Comparative Fit Index; RMSEA = Root Mean Square Error of Approximation.

¹ Thresholds for symptoms #3 and #4 vary across time

TABLE 4. Frequency of the sum of the number of Heavy involvement and Negative consequences symptoms of IGD.

	10 years		12 years		14 years		16 years		18 years	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Heavy involvement										
All	0.64	0.99	0.67	1.00	0.64	0.99	0.71	0.98	0.39	0.71
Boys	0.83	1.10	1.03	1.13	1.08	1.12	1.20	1.05	0.69	0.85
Girls	0.47	0.84	0.34	0.74	0.27	0.67	0.33	0.72	0.16	0.48
Negative consequences										
All	0.11	0.43	0.10	0.40	0.16	0.56	0.14	0.51	0.11	0.45
Boys	0.16	0.53	0.14	0.45	0.29	0.73	0.29	0.72	0.22	0.59
Girls	0.06	0.30	0.07	0.34	0.05	0.30	0.02	0.16	0.04	0.28

TABLE 5. Fit indices for growth models of Heavy involvement and Negative consequences symptoms of Internet gaming disorder from 10–18 years.

	χ^2	<i>df</i>	CFI	RMSEA	90% CI RMSEA	$\Delta\chi^2$	χ^2	<i>df</i>	<i>p</i>
Total sample									
M1: Linear growth	150.36	41	.816	.057	.048-.067				
M2: Quadratic growth	61.24	28	.944	.038	.025-.051	84.53	13	.001	

Note: CFI=Comparative Fit Index; RMSEA = Root Mean Square Error of Approximation.

TABLE 6. Linear and quadratic growth in Heavy involvement and Negative consequences symptoms of Internet gaming disorder from 10–18 years in total sample, boys, and girls.

Sample	Heavy involvement				Negative consequences			
	Linear		Quadratic		Linear		Quadratic	
	Mean	<i>p</i> -value	Mean	<i>p</i> -value	Mean	<i>p</i> -value	Mean	<i>p</i> -value
All	0.07	<.001	-0.12	<.001	0.02	.111	0.00	.130
Boys	-0.19	.196	-0.03	<.001	0.24	.044	-0.08	.360
Girls	-0.04	<.001	0.00	.396	-0.01	.007	NA	

Note: NA because the quadratic model did not converge

TABLE 7. Prevalence of IGD diagnosis.

	10 years		12 years		14 years		16 years		18 years	
	%	95% CI	%	95% CI	%	95% CI	%	95% CI	%	95% CI
All	2.00	1.00, 3.12	2.42	1.34, 3.64	2.36	1.33, 3.73	2.01	1.08, 3.21	1.00	0.33, 1.87
Boys	3.02	1.45, 5.03	2.50	1.07, 4.47	3.93	1.77, 6.33	4.05	1.47, 6.74	2.05	0.55, 4.45
Girls	0.52	0.07, 1.37	1.06	0.45, 1.87	0.60	0.38, 1.27	0.25	0.07, 0.62	0.43	0.38, 0.88

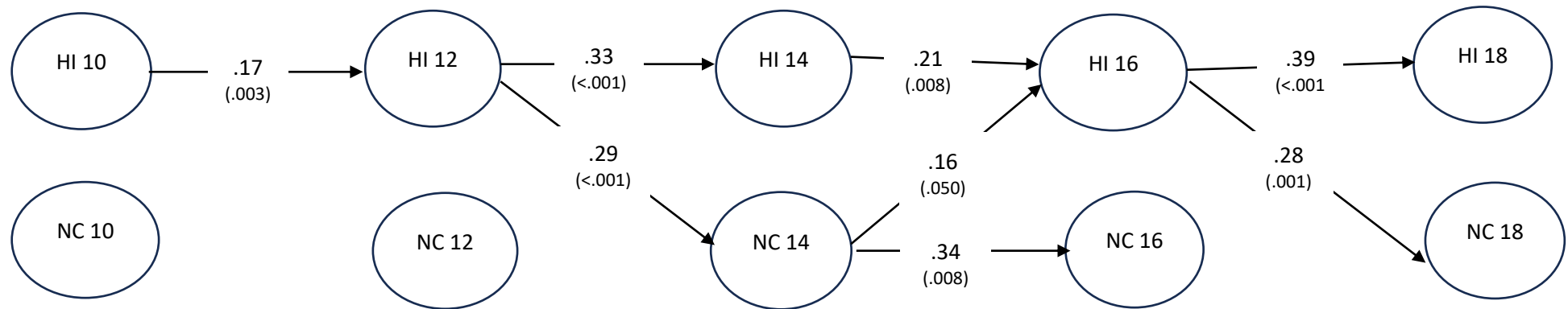
TABLE 8. Pearson correlations between dimensions of IGD at ages 10–18 years.

	Age 10		Age 12		Age 14		Age 16		Age 18	
	HI	NC	HI	NC	HI	NC	HI	NC	HI	NC
HI – age 10	-	.31 (<.001)	.29 (<.001)	.11 (.048)	.24 (<.001)	.12 (.040)	.15 (.004)	.08 (.215)	.11 (.076)	.08 (.178)
NC – age 10			.21 (.001)	.22 (.034)	.15(.004)	.08 (.075)	.10 (.047)	.09 (.117)	.06 (.302)	.11 (.141)
HI – age 12				.37 (<.001)	.46 (<.001)	.34(<.001)	.37 (<.001)	.21 (.004)	.28 (<.001)	.17 (.012)
NC – age 12					.22 (.005)	.35 (<.001)	.24 (.006)	.17 (.055)	.15 (.028)	.16 (.278)
HI – age 14						.53 (<.001)	.42 (<.001)	.18 (.004)	.39 (<.001)	.29(<.001)
NC – age 14							.31 (<.001)	.21 (.001)	.18 (.007)	.34 (.017)
HI – age 16								.37 (<.001)	.52 (<.001)	.34 (<.001)
NC – age 16									.17 (.053)	.27 (.028)
HI – age 18										.44 (<.001)
NC – age 18										-

Note: NI = Heavy involvement, NC = Negative consequences

Figure 1.

Stability of within-person changes in Heavy involvement and Negative consequences from age 10 to 18 years



Note: Random intercept cross-lagged model. Standardized estimates. Non-significant paths, random intercepts, and correlations between concurrent residuals are estimated but not included in the figure for presentation purposes. Numbers in parentheses are p-values.