

Late-Night Gaming and Sleep in Adults: A Registered Report Using Multi-Platform Telemetry

Chronotype, Sleep, and Wellbeing in Adult Gamers

Tamás A. Földes^{1*}  & Other Authors²

¹University of Oxford, Oxford Internet Institute

²Second University

Late-night gaming has been linked to disrupted sleep and poorer wellbeing, but most evidence relies on self-reported play time in adolescent samples. Using the Open Play dataset—a three-month longitudinal study of adult gamers in the UK and US that paired session-level digital-trace data from Nintendo, Xbox, and Steam with biweekly panel and daily-diary self-reports—this Registered Report tested whether late-night gaming (23:00–06:00) predicted (H1a) poorer sleep quality, (H1b) shorter sleep duration, (H1c) greater daytime sleepiness, and (H1d) lower mental wellbeing, and whether these associations were amplified among individuals with more evening-leaning chronotypes (H2). Outcomes were assessed with the PSQI, ESS, and SWEMWBS across six biweekly waves; chronotype was indexed by the MCTQ-derived MSF_{sc}. Only H1a was supported: each additional 10 minutes of daily late-night gaming was associated with an approximate 0.9 percentage-point increase in the probability of reporting poor sleep quality, consistent across imputed, complete-case, and exploratory diary specifications. With a smallest effect size of interest set to $\pm 0.1 \times \text{SD}$ of the outcome, TOST equivalence tests accepted the null for H1b–H1d, bounding any true effects below ± 7 minutes of nightly sleep, ± 0.38 Epworth points, and ± 0.53 SWEMWBS points per one-SD (≈ 36 min/day) increase in late-night gaming. None of the H2 chronotype \times late-night gaming interactions reached significance in the pooled imputation models; equivalence tests accepted the null for H2b–H2d, with H2a undecided. The “quality down, duration preserved” pattern is inconsistent with sleep displacement but compatible with pre-sleep arousal, attribution, or compensatory phase delay — mechanisms self-report alone cannot adjudicate.

Keywords: late-night gaming, sleep quality, sleep duration, daytime sleepiness, mental wellbeing, chronotype, digital trace data, registered report

Words: 10239

Introduction

Concerns have been raised about the potential negative impacts of video gaming on sleep and overall wellbeing, particularly for adolescents and young adults and especially when gaming occurs late at night (Hale & Guan, 2015; Muppalla et al., 2023; Peracchia & Curcio, 2018). Pre-sleep technology use is also the modal evening behaviour in adults: a recent nationally representative US survey found that 89.9% of adults use an electronic device within an hour of bedtime, with 71.3% using a TV or gaming console specifically and 49.3% doing so inside the bedroom (Carlson et al., 2026). Heavy or problematic gaming has been shown to disrupt sleep patterns, reduce sleep duration, lower sleep quality, and increase daytime sleepiness (Exelmans & Van Den Bulck, 2015; Han et al., 2024; Kemp et al., 2021; Kristensen et al., 2021). Crucially, however, Kristensen et al. (2021) noted that none of the studies in their review registered the time of day gaming took place, meaning that timing-specific effects — particularly those concentrated in the pre-sleep window — remain poorly characterised. This is especially concerning given the far-reaching effects of sleep disturbances on cognitive and emotional functioning (McCoy & Strecker, 2011; Simon et al., 2020; Vriend et al., 2013). For instance, habitual gaming between 10 p.m. and 6 a.m. has been associated with an increased risk of depressive symptoms, partially mediated by daytime sleepiness (Lemola

*Send correspondence to: Tamás A. Földes, contact@tamasfoldes.mozmail.com.

et al., 2011). Understanding the consequences of late-night gaming is thus vital for both gamers and health professionals.

Mechanisms Linking Late-Night Gaming to Sleep Disturbance

Two key mechanisms have been proposed to explain the impact of late-night digital engagement (including gaming) on sleep (Cain & Gradisar, 2010; LeBourgeois et al., 2017). The first is the displacement hypothesis, which argues that late-night gaming is more harmful than daytime gaming because it cuts into sleep time (Exelmans & Van Den Bulck, 2015; Lemola et al., 2011; Twenge, 2019). Gamers often feel compelled to continue playing and may struggle with self-regulation, which can lead to insufficient sleep (King & Delfabbro, 2009; Pirrone et al., 2024; Spada & Caselli, 2017). For example, adolescents with high trait flow delayed bedtime by ~90 minutes when playing challenging games (Smith et al., 2017).

The second mechanism involves arousal-related disturbances in sleep architecture caused by late-night gaming. Empirical studies have shown that extended gaming, especially when involving violent content, significantly decreases REM sleep and total sleep time (King et al., 2013). Polysomnographic evidence in school-aged children indicates that a single bout of evening computer gaming reduced slow-wave sleep, lengthened sleep-onset latency, and impaired next-day verbal memory (Dworak et al., 2007). Weaver et al. (2010) found that pre-sleep gaming extended sleep latency by approximately 5–10 minutes, while King et al. (2013) demonstrated that such arousal-related disturbances can also alter the natural progression into sleep stages. This delay in sleep onset could be exacerbated by lower melatonin levels following an evening of gaming, compared to neutral activities like board games, which are crucial for regulating the sleep-wake cycle (Hartmann et al., 2019).

The Moderating Role of Chronotype

Negative effects of late-night gaming may be compounded among individuals with an eveningness chronotype—a group naturally predisposed to staying up late and consistently shown to spend more time on screen-based media in the pre-sleep window (Kortesoja et al., 2023; Reardon et al., 2023). Pre-sleep technology use, in turn, appears to affect this group more strongly: in a large adult cohort, daily screen use before bed was associated with later bedtimes and shorter sleep among both chronotypes, but the delay and sleep loss were markedly greater in evening types — a pattern the authors interpret as evidence of compounded social jetlag, i.e. the misalignment between endogenous circadian rhythms and socially imposed schedules (Zhong et al., 2025). Adolescent evidence is consistent: late-night digital media use mediates the link between evening chronotype and poorer sleep quality and daytime tiredness, and these effects are most pronounced for evening types (Kortesoja et al., 2023).

The downstream picture for psychosocial outcomes is more nuanced. Reardon et al. (2023) found that shorter sleep on weekdays was associated with greater psychological distress, but technology medium and chronotype themselves were not direct predictors of distress. Gumport et al. (2021) reported that, in evening-type adolescents, technology use was linked to better emotional, social, cognitive, and physical health but worse behavioral health (sensation-seeking, ADHD diagnosis, and alcohol/substance use); notably, electronic game use specifically was not associated with behavioral health in that study the behavioral-health association was driven by other technology uses. The adult evidence reviewed above (Zhong et al., 2025) suggests that the chronotype-amplified effect of pre-sleep screen exposure on sleep timing and duration is not confined to adolescence, but adult evidence specifically isolating gaming (rather than screen use broadly) within the pre-sleep window remains sparse.

The Present Study

In sum, the literature indicates that video gaming, particularly when it occurs late at night, has significant implications for sleep quality, sleep duration, and overall wellbeing. This disruption can be attributed to both the displacement hypothesis (Exelmans & Van Den Bulck, 2015; Lemola et al., 2011; Twenge, 2019) and arousal-related disturbances in sleep architecture (King et al., 2013). Individual differences, such as chronotype, may moderate these effects, with eveningness

chronotypes particularly vulnerable to the negative consequences of pre-sleep screen exposure (Kortesoja et al., 2023; Zhong et al., 2025). The present study aims to empirically test the following hypotheses regarding the relationship between late-night gaming and sleep outcomes:

H1: Late-night gaming is associated with:

- H1a: Poorer sleep quality
- H1b: Shorter sleep duration
- H1c: Higher daytime sleepiness
- H1d: Lower wellbeing

In addition to testing direct relationships between late-night gaming and various sleep-related outcomes are critical to understand, we further assess the potential moderating role of chronotype, which refers to a person's natural preference for activities during certain times of the day (morningness or eveningness). Individuals with an evening chronotype tend to stay up later and may be more inclined to engage in late-night gaming, potentially exacerbating the negative impacts on sleep and wellbeing. The combination of an evening chronotype and late-night gaming may even have a compounded effect on overall wellbeing, as both factors are independently associated with poorer mental health outcomes. Given this, we propose the following:

H2: Chronotype moderates the relationships between late-night gaming and all outcomes in H1 (sleep quality, sleep duration, daytime sleepiness, and wellbeing), such that these negative associations are stronger for individuals with more of an eveningness chronotype.

By examining chronotype on a continuous scale as a moderating factor, this study seeks to provide a more nuanced understanding of the potential risks associated with late-night gaming and to identify individuals who may be most vulnerable to its negative effects.

Methods

Data Source and Measures

The analyses reported here are part of a Stage 1 Registered Report (Ballou et al., 2024) and utilize data from the Open Play dataset (Ballou et al., 2025a), a longitudinal study that collected multi-platform video game digital trace data alongside psychological measures from adult gamers in the UK and US over a three-month period. The study combined objective behavioral telemetry from gaming platforms with repeated self-report surveys administered biweekly across six waves. Importantly, the present analyses use only a subset of the Open Play dataset, specifically data from Nintendo, Xbox, and Steam platforms, as these provide session-level (Nintendo, Xbox) or near session-level (Steam) temporal granularity necessary for hourly aggregation of playtime to operationalize late-night gaming (23:00–06:00). The following validated measures were administered via panel surveys at multiple timepoints: Wellbeing was assessed using the Short Warwick-Edinburgh Mental Well-being Scale [SWEMWBS; Tennant et al. (2007)], a 7-item measure of mental wellbeing covering psychological functioning and subjective well-being over the past 2 weeks, with responses on a 5-point Likert scale ranging from “None of the time” to “All of the time” (score range: 7–35).

Sleep quality and duration were assessed using the Pittsburgh Sleep Quality Index [PSQI; Buysse et al. (1989)], a 19-item questionnaire evaluating sleep quality over the past month. The measure yields seven component scores (sleep quality, sleep latency, sleep duration, sleep efficiency, sleep disturbances, use of sleep medication, and daytime dysfunction) and a global score (range: 0–21), with scores above 5 indicating poor sleep quality. Excessive daytime sleepiness was measured using the Epworth Sleepiness Scale [ESS; Johns (1991)], an 8-item scale assessing the likelihood of dozing off in various situations (score range: 0–24). Higher scores indicate greater propensity for daytime sleepiness, with scores above 10 typically indicating clinically significant excessive sleepiness. Chronotype was measured at baseline (Wave 1) using the Munich Chronotype Questionnaire [MCTQ; Roenneberg et al. (2003)]. The key metric used in this study is MSF_{sc} (Mid-Sleep on Free Days corrected for sleep debt on work days), which represents an individual's natural sleep-wake preference when not constrained by social obligations. Higher MSF_{sc} values indicate a preference for eveningness (later sleep-wake times).

Handling Missing Data

Missingness in the longitudinal self-report outcomes (PSQI components, PSQI sleep duration, Epworth Sleepiness Scale, and SWEMWBS) was addressed via hierarchical two-level multiple imputation by chained equations [MICE; mice v3.16.0 and miceadds v3.19 in R; van Buuren & Groothuis-Oudshoorn (2011); Robitzsch & Grund (2024)] using predictive mean matching for multilevel data (2l.pmm). This method models participants as clusters with repeated waves nested within, preserving between-person variation that single-level PMM would attenuate. We imputed under a Missing at Random assumption conditional on rich auxiliary information. Because sleep measures (PSQI, ESS) were only collected at waves 2, 4, and 6 by design, the imputation was run in two passes: biweekly variables (SWEMWBS, measured at all six waves) and monthly variables (PSQI components, sleep duration, ESS, measured at waves 2, 4, 6 only). This two-pass approach avoids exposing the mixed-effects model inside 2l.pmm to the structural NAs that arise at non-measurement waves. The PSQI global score was derived via passive imputation (sum of seven imputed component scores). Following van Buuren's 2018 multilevel imputation recipe, predictor matrices coded participant ID as the random-intercept cluster (code -2), level-2 (person-level) predictors as fixed effects only (code 1), and level-1 outcome cross-predictors with disaggregated cluster means (code 3). In the biweekly pass, biweekly gaming exposure averages also received code 3; in the monthly pass, wave, gaming, and lag/lead terms remained at code 1 to avoid inflating the predictor count. This coding ensures that both the raw within-person value and its person mean enter the imputation model where appropriate, preserving the between vs. within-person decomposition critical for contextual effects. Continuous outcomes were grand-mean centred before imputation and back-transformed afterwards to improve stability of the mixed-effects models inside 2l.pmm (van Buuren, 2018).

Quality control diagnostics (convergence, density, strip, missingness, and between-imputation variability plots) were generated and inspected for both the panel and diary imputations and found to be acceptable. Level-2 (person-level) predictors included static demographics (age, BMI, SES, region, gender recoded as male/female/other) and chronotype (MSF_{sc}); level-1 time-varying predictors included wave number, biweekly gaming exposure averages (total and late-night minutes over the preceding 14 days), and ± 1 -wave temporal lag and lead terms for each outcome variable to capture within-person trajectories. To ensure each participant contributed a full six-wave panel, we first expanded the self-report data to the complete participant \times wave grid and inferred survey timestamps for missing waves by aligning observed dates with wave-specific medians; those inferred dates were then used to compute the rolling gaming exposures. For the primary panel dataset we generated 20 imputed datasets with 20 iterations per pass—sufficient to stabilize estimates given wave-specific missingness rates up to ~56%. Diagnostic density and strip plots confirmed plausible imputations. All regressions were fit separately in each imputed dataset and combined using Rubin's rules.

Missingness in the adjustment covariates was limited in the original (pre-imputation) analytical dataset; exact per-covariate counts are reported in the Results section below. The imputed outcomes serve as the primary analytic dataset throughout the main text. Complete-case versions of every regression (using only non-imputed observations for each outcome) were re-estimated in parallel and are reported in the Appendix (Table A1, Table A2, Table A6). Wave-level outcome missingness rates are reported in Table A3.

Results

Data Quality Controls

Prior to hypothesis testing, we conducted three positive data quality controls (DQCs) to check the directions and magnitudes of three expected associations in the data. First, self-reported playtime was positively correlated with digital trace playtime ($r = 0.49$, 95% CI [0.47, 0.51], $p < .001$). Second, social jetlag was positively associated with daytime sleepiness (Spearman's $\rho = 0.09$, $p < .001$, one-sided), in the same direction reported for circadian misalignment and sleepiness in prior work (Fernandes et al., 2023; Wu et al., 2025). Third, sleep quality was negatively associated with wellbeing (Spearman's $\rho = -0.25$, $p < .001$, one-sided), in the same direction reported for sleep and

mental health in prior work (Gadie et al., 2016). All three associations were in the preregistered direction.

Sample Demographics

Of 34,922 participants who signed up for the study, 1,948 completed at least one biweekly panel survey. The analytical sample was derived from these participants by applying a three-step filter. First, participants were required to have at least one valid outcome measure (SWEMWBS, PSQI, or ESS) at wave 1, which excluded 107 participants ($n = 1,841$). Second, participants needed valid timezone data (either self-reported or imputed for UK participants) to accurately classify late-night gaming sessions, which excluded a further 120 participants ($n = 1,721$). Third, participants were required to have contributed at least one valid gaming session during the study period, which excluded 143 additional participants, yielding a final analytical sample of $N = 1,578$. Table 1 presents the demographic composition of both the total survey sample and the analytical sample.

Table 1. Sample Characteristics

| Characteristic | Total | Analytical |
|----------------------------------|---------------|---------------|
| A. Sociodemographics | | |
| N | 1948 | 1578 |
| Age | 26.8 (5.0) | 27.1 (5.1) |
| Gender | | |
| Woman | 518 (26.6%) | 444 (28.1%) |
| Man | 1211 (62.2%) | 1035 (65.6%) |
| Other | 111 (5.7%) | 99 (6.3%) |
| Region | | |
| UK | 719 (36.9%) | 672 (42.6%) |
| US | 1121 (57.5%) | 906 (57.4%) |
| BMI (kg/m ²) | 22.0 (7.0) | 22.1 (7.0) |
| SES index | 2.27 (0.54) | 2.26 (0.54) |
| B. Chronotype | | |
| No alarm on free days | 1240 (74.4%) | 1141 (74.5%) |
| MCTQ-MSFsc (HH:MM) ¹ | 06:00 (03:23) | 05:52 (03:05) |
| C. Gaming | | |
| Gaming (min/day) ¹ | 59.6 (138.3) | 83.7 (137.2) |
| LN gaming (min/day) ¹ | 4.4 (24.1) | 9.3 (30.7) |
| % nights LN gaming | 13.3 (16.3) | 16.4 (16.7) |
| D. Outcomes | | |
| Sleep (h) | 7.2 (1.2) | 7.2 (1.1) |
| PSQI global | 6.7 (2.9) | 6.7 (2.8) |
| Sleep quality | 1.3 (0.6) | 1.3 (0.6) |
| Poor sleep (PSQI>5) | 793 (63.8%) | 754 (63.8%) |
| ESS | 5.6 (3.5) | 5.6 (3.5) |
| Excessive sleepiness (ESS>10) | 126 (10.2%) | 120 (10.2%) |
| SWEMWBS | 23.2 (5.0) | 23.2 (5.0) |

Values are M (SD) unless noted. ¹ Mdn (IQR). LN = late-night.

Self-reported sleep duration in the analytical sample was 7.2 hours (SD = 1.1), mean PSQI sleep-quality component scores were 1.3 (SD = 0.6), mean daytime sleepiness was 5.6 on the Epworth Sleepiness Scale (SD = 3.5), and mean wellbeing was 23.2 on the SWEMWBS (SD = 5.0). Figure 1 displays the distributions of gaming patterns and outcomes across the analytical sample.

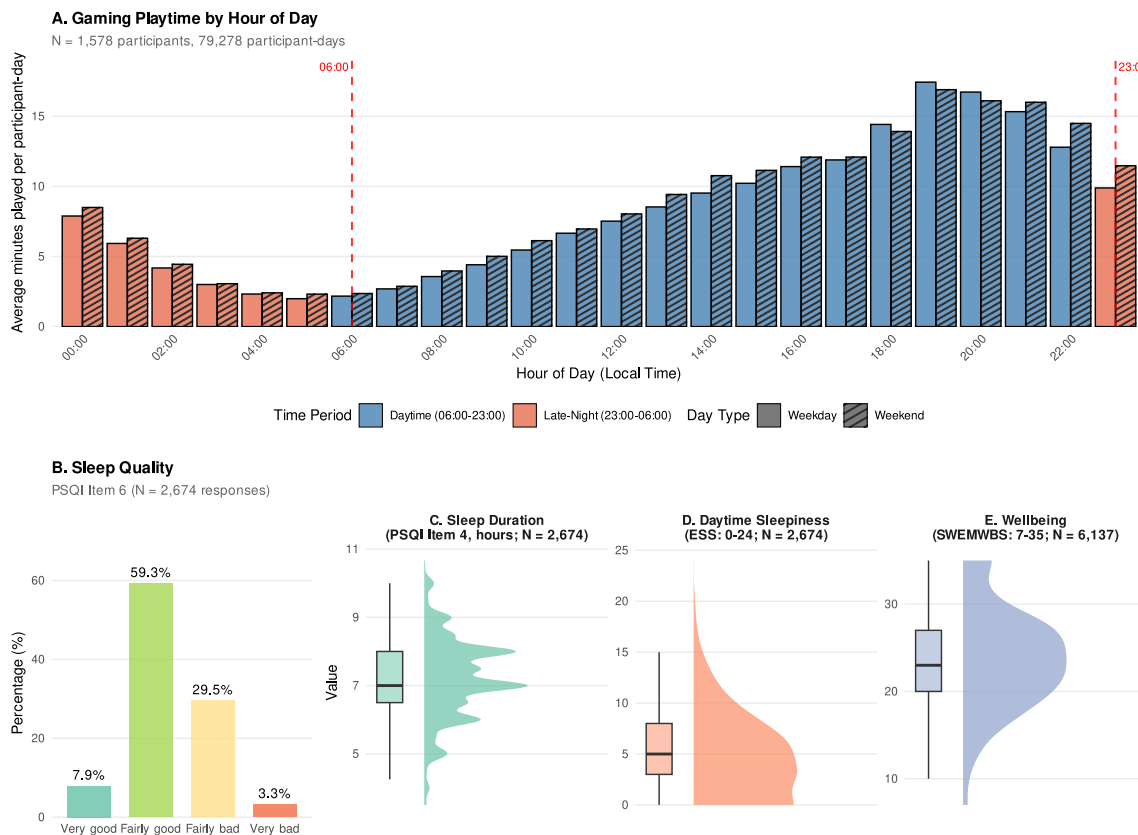


Figure 1. Gaming patterns and outcome distributions in the complete-case analytical sample. (A) Hourly playtime: average minutes played per participant-day (total minutes summed across all sessions divided by the number of contributing participant-days), shown by hour of day, with grouped bars distinguishing weekdays (Sunday-Thursday nights, solid fill) from weekend nights (Friday-Saturday, striped fill); the pre-registered late-night window (23:00-06:00 local time) is delimited by red dashed lines and shaded in red. (B) Self-reported sleep quality (PSQI Item 6) shown as the percentage of person-wave responses in each ordinal category; missing responses are excluded. (C-E) Continuous outcomes displayed as raincloud plots: half-violin shapes show the kernel density estimate (bandwidth multiplier = 2, normalised within panel), boxplots show the median (centre line), interquartile range (box) and whiskers extending to the most extreme value within 1.5 x IQR of the hinges; outliers beyond the whiskers are not plotted. Self-reported sleep duration in panel C is derived from PSQI Item 4. Panel-specific sample sizes are reported in the panel subtitles. Abbreviations: PSQI = Pittsburgh Sleep Quality Index; ESS = Epworth Sleepiness Scale (range 0-24, higher = greater daytime sleepiness); SWEMWBS = Short Warwick-Edinburgh Mental Well-Being Scale (range 7-35, higher = greater wellbeing); IQR = interquartile range.

Missingness in the adjustment covariates was minimal in the original (pre-imputation) analytical dataset of 1182 participants: BMI was missing for 80 (6.8%), while age, SES index, region, and the weekend/weekday indicator were complete.

Panel

H1

For sleep quality (H1a), each additional 10 minutes of average daily late-night gaming was associated with a probit-scale change of $b = 0.051$, 95% CI [0.023, 0.079], $p < .001$, corresponding to an

approximate 0.9 percentage-point increase in the marginal probability of reporting fairly bad or very bad sleep quality (Figure 2 translates this coefficient to the probability scale, with marginal predicted probabilities computed in each of the 20 imputed datasets, pooled via Rubin's rules, and 95% confidence intervals derived from the combined within- and between-imputation variance using the delta method). The corresponding pooled effects were $b = -0.005$, 95% CI $[-0.019, 0.008]$, $p = 0.456$ (hours per 10 min/day) for sleep duration (H1b), $b = 0.013$, 95% CI $[-0.034, 0.061]$, $p = 0.577$ (Epworth points per 10 min/day) for daytime sleepiness (H1c), and $b = -0.004$, 95% CI $[-0.038, 0.030]$, $p = 0.805$ (SWEMWBS points per 10 min/day) for wellbeing (H1d) — each indicating negligible variation in the outcome as a function of late-night play. Full coefficient estimates, confidence intervals, and variance components are reported in the H1 regression summary table.

These estimates depart from the Stage 1 protocol (Ballou et al., 2024) in two respects. First, the preregistered by-participant random slopes on late-night minutes were dropped after they produced convergence problems and boundary estimates (near-zero variance components) in the Open Play data — most acutely in the H1a cumulative link mixed model — so all four H1 models retain only random intercepts for participants, and for gender in H1c and H1d where it remained supported. Second, the multiply imputed outcomes (20 datasets pooled via Rubin's rules (Rubin, 1987)) are used as the primary analytic dataset rather than the incomplete original outcomes. All models additionally adjust for age, BMI, SES index, region, and weekend versus weekday timing, and use rolling 14- and 28-day windows of late-night play anchored to each survey date.

Two further sensitivity checks support the linear parameterization used in the preregistered models. The H1a model was refit using the PSQI global score (sum of all seven components, range 0–21) as a continuous alternative to the ordinal item-6 outcome; results are reported in Table A4. A natural cubic spline sensitivity analysis (Table A7 Figure A1), evaluated by BIC under the rule of Jones et al. (2001) ($|2\Delta\text{BIC}| < 2 = \text{indifferent}$, $2-6 = \text{positive evidence}$), returned outcome-specific verdicts rather than a blanket endorsement of linearity: evidence was indifferent between the linear and spline parameterizations for H1a (sleep quality, $|2\Delta\text{BIC}| = 0.6$) and H1d (wellbeing, $|2\Delta\text{BIC}| = 0.6$); positive but weak evidence favored the linear specification for H1b (sleep duration, $|2\Delta\text{BIC}| = 3.6$); and positive but weak evidence favored a non-linear (spline) specification for H1c (daytime sleepiness, $|2\Delta\text{BIC}| = 2.8$). We retained the linear parameterization in the preregistered models for all four outcomes; for H1c on the basis of parsimony and direct comparability with the other H1 models.

Table 2. Summary of H1 Hypotheses: Effects of Late-Night Gaming on Sleep and Wellbeing

| | H1a: Sleep Quality | H1b: Sleep Duration | H1c: Daytime Sleepiness | H1d: Wellbeing |
|--|------------------------|-------------------------|-------------------------|----------------------|
| Daily LN gaming (per 10 min/day, monthly) | 0.05 [0.02, 0.08]*** | -0.01 [-0.02, 0.01] | 0.01 [-0.03, 0.06] | |
| Daily LN gaming (per 10 min/day, biweekly) | | | | -0.00 [-0.04, 0.03] |
| Age (scaled) | 0.16 [-0.17, 0.50] | -0.45 [-0.62, -0.27]*** | -0.68 [-1.27, -0.09]* | 0.07 [-0.59, 0.72] |
| BMI (scaled) | 0.18 [0.05, 0.31]** | -0.07 [-0.14, -0.00]* | 0.18 [-0.05, 0.41] | -0.21 [-0.48, 0.07] |
| SES (scaled) | -0.24 [-0.39, -0.09]** | -0.09 [-0.16, -0.01]* | 0.09 [-0.17, 0.34] | 0.98 [0.70, 1.26]*** |
| Region: US | -0.14 [-0.37, 0.09] | 0.04 [-0.08, 0.17] | 0.25 [-0.18, 0.69] | 0.19 [-0.29, 0.68] |
| Day: Weekend | 0.03 [-0.16, 0.23] | -0.00 [-0.09, 0.09] | 0.04 [-0.25, 0.33] | -0.01 [-0.27, 0.25] |
| N Obs | 3561 | 3553 | 3551 | 7425 |
| N Participants | 1294 | 1293 | 1290 | 1469 |
| ICC | 0.75 | 0.69 | 0.69 | 0.71 |

+ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

LN = late-night. Confidence intervals shown in brackets.

ICC = Intraclass Correlation Coefficient (adjusted).

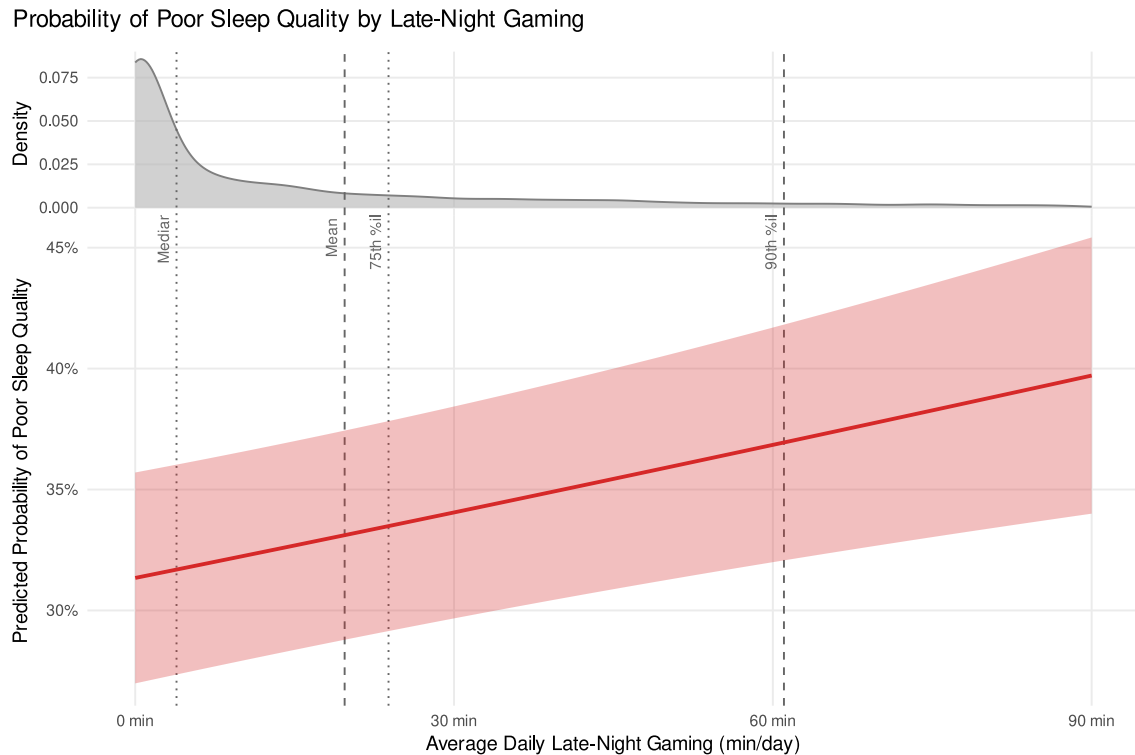


Figure 2. Marginal predicted probability of poor sleep quality (Fairly bad or Very bad) as a function of late-night gaming. Predictions are derived from the H1a probit cumulative link mixed model fitted separately on each of 20 multiply imputed datasets; predicted probabilities and their within-imputation variances are then pooled via Rubin’s rules. The solid line shows the pooled point estimate and the shaded ribbon the 95% confidence interval, both computed on the probability scale using the delta method applied to each imputation’s threshold, gaming coefficient, and variance–covariance matrix, then combined with Rubin’s within- and between-imputation variance components. Probabilities are population-average (marginalised over the participant random intercept) with other covariates held at their reference or mean values. The top panel shows the marginal density of late-night gaming in the sample; vertical reference lines mark the median, mean, 75th, and 90th percentiles.

We also conducted frequentist equivalence tests (TOST; [Lakens \(2017\)](#)) on the focal H1 predictors against a Region of Practical Equivalence (ROPE) derived from $\pm 0.1 \times \text{SD}(\text{outcome})$ rescaled to the raw coefficient’s native units (see Section A1.8 for derivation, the rationale for the ROPE width, and response-scale thresholds. The pooled 90% CIs for H1b, H1c, and H1d fell entirely within their respective ROPEs. The H1a sleep-quality CI lay partly outside the stricter ordinal-probit ROPE, returning an “Undecided” TOST decision. Full per-hypothesis TOST decisions, ROPE widths, and pooled 90% CIs are reported in Table A8.

H2

Having established a small but consistent late-night gaming \rightarrow poor-sleep-quality association in H1, we next asked whether this association is moderated by chronotype, on the prior expectation that evening types (higher MSF_{sc}) may pay a larger sleep cost for the same late-night play.

None of the preregistered chronotype \times late-night gaming interactions reached conventional significance after pooling across imputations, so H2 was not supported for any outcome (H2a sleep quality: $b = -0.003$, 95% CI $[-0.013, 0.006]$, $p = 0.487$ probit per 10 min/day per hour of MSF_{sc} ; H2b sleep duration: $b = 0.003$, 95% CI $[-0.002, 0.008]$, $p = 0.199$ hours per 10 min/day per hour of MSF_{sc} ; H2c daytime sleepiness: $b = -0.009$, 95% CI $[-0.024, 0.006]$, $p = 0.225$ Epworth points per 10 min/day per hour of MSF_{sc} ; H2d wellbeing: $b = -0.009$, 95% CI $[-0.019, 0.002]$, $p = 0.101$ SWEMWBS points per 10 min/day per hour of MSF_{sc}). Equivalence tests (Table A8) go further and provide positive evidence that the three linear interactions are practically negligible: applying the TOST

Table 3. Summary of H2 Hypotheses: Chronotype Moderation of Late-Night Gaming Effects

| | H2a: Sleep Quality | H2b: Sleep Duration | H2c: Daytime Sleepiness | H2d: Wellbeing |
|--|------------------------|-------------------------|-------------------------|----------------------|
| Daily LN gaming (per 10 min/day, monthly) | 0.06 [0.02, 0.11]** | -0.01 [-0.03, 0.01] | 0.02 [-0.04, 0.09] | |
| Daily LN gaming (per 10 min/day, biweekly) | | | | 0.04 [-0.01, 0.08] |
| Chronotype (h, centered) | 0.03 [-0.02, 0.08] | -0.02 [-0.05, 0.01] | 0.04 [-0.04, 0.12] | -0.07 [-0.16, 0.03] |
| LN gaming × Chronotype (h, monthly) | -0.00 [-0.01, 0.01] | 0.00 [-0.00, 0.01] | -0.01 [-0.02, 0.01] | |
| LN gaming × Chronotype (h, biweekly) | | | | -0.01 [-0.02, 0.00] |
| Age (scaled) | -0.06 [-0.53, 0.40] | -0.45 [-0.69, -0.21]*** | -0.93 [-1.71, -0.16]* | 0.14 [-0.74, 1.01] |
| BMI (scaled) | 0.06 [-0.10, 0.23] | -0.04 [-0.13, 0.05] | 0.12 [-0.15, 0.39] | -0.06 [-0.41, 0.28] |
| SES (scaled) | -0.28 [-0.50, -0.07]** | -0.09 [-0.19, 0.01]+ | 0.13 [-0.20, 0.46] | 1.08 [0.71, 1.44]*** |
| Region: US | -0.17 [-0.47, 0.14] | -0.03 [-0.19, 0.13] | 0.38 [-0.17, 0.93] | 0.30 [-0.34, 0.93] |
| Day: Weekend | 0.07 [-0.19, 0.32] | 0.00 [-0.11, 0.12] | 0.04 [-0.31, 0.39] | -0.06 [-0.38, 0.27] |
| N Obs | 2580 | 2580 | 2580 | 5160 |
| N Participants | 860 | 860 | 860 | 860 |
| ICC | 0.76 | 0.70 | 0.70 | 0.70 |

+ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

LN = late-night. Confidence intervals shown in brackets.

ICC = Intraclass Correlation Coefficient (adjusted).

procedure to the chronotype × late-night gaming coefficients against their interaction-scale ROPEs, the pooled 90% CIs for H2b, H2c, and H2d all lie entirely within the ROPE; the H2a interaction is undecided under the stricter ordinal-probit ROPE. The H2a main effect of late-night gaming was $b = 0.064$, 95% CI [0.024, 0.105], $p = 0.002$ (probit per 10 min/day, ≈ 1.0 percentage-point increase in the probability of reporting fairly bad or worse sleep), closely mirroring the H1a estimate and indicating that the small late-night-gaming → sleep-quality association is not concentrated in chronotype-specific subgroups. Full model summaries are reported in Table 3.

The H2 models added a late-night gaming × chronotype (MSF_{sc}) interaction to each H1 specification; chronotype had a median of 5.9 hours (IQR 3.1) in the analytical sample and was mean-centred before entering the interaction. The Open Play data made the moderation structure more demanding than H1, so we applied the same random-effects simplification (random intercepts for participants, and for gender where supported, dropping the by-participant random slopes on late-night minutes) while retaining the gaming × chronotype interaction for all four outcomes. With these adjustments the ordinal H2a CLMM converged cleanly alongside the H2b–H2d linear mixed-effects models, with finite standard errors for all fixed effects (including the interaction term) in both the complete-case and imputed fits.

Diary

H1 and H2

The diary stage tests whether the panel-level H1a/H2a finding holds when sleep quality and late-night gaming are measured day-by-day, with the trimmed late-night exposure (capped at the 99th percentile) decomposed into a within-person component (daily deviation from the participant's own mean) and a between-person component (participant mean centred at the grand mean).

For the direct-effects model (H1), the within-person effect of late-night gaming on sleep quality was $b = -0.003$, 95% CI $[-0.007, 0.002]$, $p = 0.214$ (probit coefficient per 10 additional minutes); on days when participants gamed more than their own average, the association with sleep quality was not statistically significant. The between-person effect was $b = 0.029$, 95% CI $[0.012, 0.047]$, $p = 0.001$ (probit coefficient per 10 min/day) and was statistically significant, indicating that participants who habitually engaged in more late-night gaming tended to report worse sleep quality on average. On the probability scale, each additional 10 minutes of habitual daily late-night gaming is associated with an approximate 0.6 percentage-point increase in the probability of reporting poor or very poor sleep, broadly consistent with the panel-level H1a estimate (0.9 pp per 10 min; Figure 3, Panel A).

For the chronotype moderation model (H2), the interaction between within-person late-night gaming and chronotype was $b = -0.002$, 95% CI $[-0.004, -0.000]$, $p = 0.024$ and the interaction between between-person late-night gaming and chronotype was $b = -0.001$, 95% CI $[-0.009, 0.007]$, $p = 0.730$. The within-person interaction was statistically significant, indicating that the day-to-day link between late-night gaming and sleep quality was somewhat weaker for evening types; the between-person interaction was not statistically significant, suggesting that the trait-level association between habitual late-night gaming and sleep quality did not meaningfully vary across chronotypes. The main effect of chronotype was $b = 0.035$, 95% CI $[0.005, 0.065]$, $p = 0.022$, with a later chronotype associated with worse sleep quality (statistically significant; Figure 3, Panel B).

These diary analyses were not formally pre-registered and are exploratory; they re-use the H1a/H2a probit CLMM specification (random intercept for participant, late-night gaming scaled per 10 min/day, same covariate set) but operate over the preceding 24 hours rather than the 14- or 28-day windows used in the panel, with sleep quality recorded on a 5-level ordinal scale (Very poor, Poor, Fair, Good, Very good). The analytical diary sample comprised 1271 participants contributing 15,842 diary entries with valid sleep quality ratings (Table A5 in the Appendix). Region was dropped because all diary participants are US-based, and a random intercept for gender failed to converge with only three levels, so gender was excluded as well; continuous covariates (age, BMI, SES) were rescaled within the diary subsample, and SES values missing for participants with unrecognised employment categories ($n = 23$) were imputed via standard PMM ($m = 5$) using age, BMI, gender, and region as predictors. Missing diary outcomes and daily predictors were handled with the same hierarchical two-level `2l.pmm` approach and predictor-matrix coding used in the panel imputation, with participants as clusters and diary days nested within (`miceadds`); level-2 variables were age, BMI, SES, chronotype, gender, and person-mean late-night gaming, level-1 daily predictors with genuine missingness (gaming played, basic psychological needs, stress, day type, late-night gaming hours) were imputed via standard PMM within the same MICE run, and ± 5 -day (rather than ± 1 -wave) lag and lead terms provided temporal context. We generated 60 imputed datasets with 20 iterations each – more than for the panel because of higher per-day missingness – inspected QC diagnostics and judged them acceptable, and combined diary regression estimates across imputations using Rubin's rules.

We compared the linear specification against natural cubic spline alternatives ($df = 2-6$). Among the splines, $df = 2$ was preferred, but the linear model had a lower BIC by 7.7 units ($2|\Delta BIC| = 15.4$) – very strong evidence for the linear specification under the [Jones et al. \(2001\)](#) rule of thumb. We therefore retained the linear specification, both as a parsimonious summary of the average within- and between-person associations and for direct comparability with the panel H1a/H2a estimates.

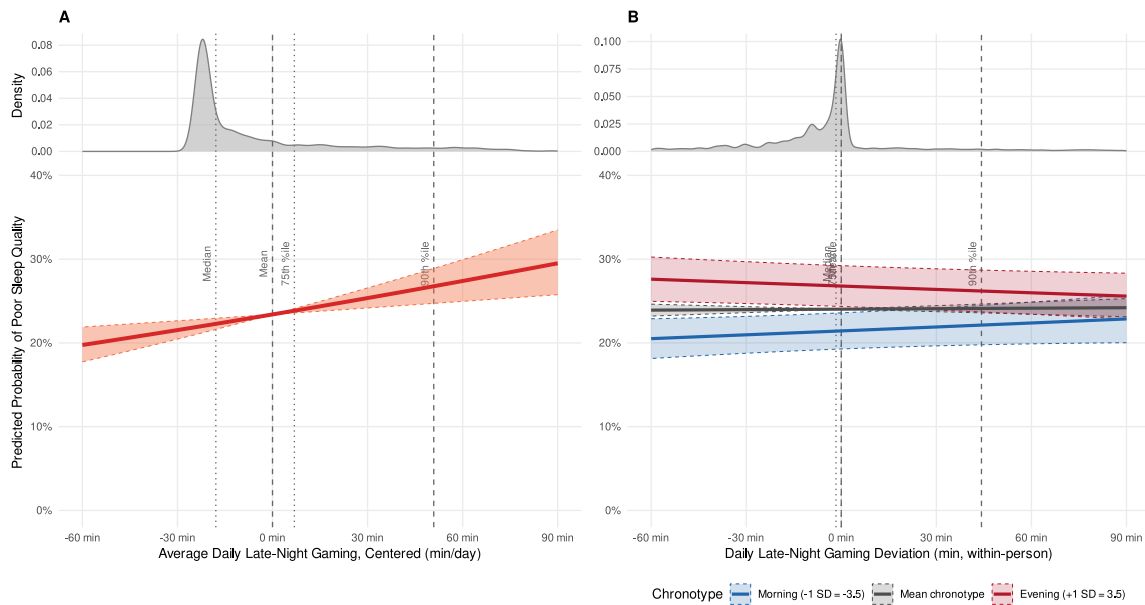


Figure 3. Predicted probability of poor sleep quality from the diary CLMM. (A) Between-person effect (H1): marginal probability of reporting Poor or Very poor sleep quality as a function of participants' average daily late-night gaming, with the within-person component held at zero. The ribbon shows 95% CIs from the between-person coefficient SE. (B) Within-person x chronotype interaction (H2): marginal probability of poor sleep quality as a function of daily within-person gaming deviation at three chronotype levels (morning = -1 SD, mean, evening = $+1$ SD), with the between-person component held at zero. Ribbons show 95% CIs via the delta method pooled across 60 imputations using Rubin's rules. Vertical reference lines mark the median (dotted), mean (dashed), 75th percentile (dotted), and 90th percentile (dashed) of the respective predictor distributions. Density plots (top) show predictor distributions.

Table 4. Diary CLMM Regression Results: H1 (Direct Effects) and H2 (Chronotype Moderation) for Sleep Quality

| | H1: Sleep Quality | H2: Chronotype Moderation |
|---------------------------------------|----------------------------|----------------------------|
| LN gaming within-person (per 10 min) | -0.00 [-0.01, 0.00] | 0.00 [-0.00, 0.01] |
| LN gaming between-person (per 10 min) | 0.03** [0.01, 0.05] | 0.02 [-0.01, 0.05] |
| Chronotype (h, centered) | | 0.04* [0.00, 0.07] |
| LN within × Chronotype | | -0.00* [-0.00, -0.00] |
| LN between × Chronotype | | -0.00 [-0.01, 0.01] |
| Age (scaled) | 0.01 [-0.06, 0.07] | -0.06 [-0.16, 0.04] |
| BMI (scaled) | 0.07* [0.01, 0.14] | 0.08 [-0.02, 0.18] |
| SES (scaled) | -0.13*** [-0.20, -0.07] | -0.17** [-0.28, -0.07] |
| Day: Weekend | -0.29*** [-0.33, -0.24] | -0.28*** [-0.34, -0.22] |
| N Obs | 14690 | 8399 |
| N Participants | 1132 | 509 |
| ICC | 0.48 | 0.53 |

+ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

LN = late-night. Confidence intervals shown in brackets below each estimate.

Estimates pooled across 60 multiply imputed datasets using Rubin's rules.

Cumulative link mixed models (random intercept for participant) on 5-level ordinal sleep quality (positive coefficients = higher probability of worse sleep). Both H1 and H2 use probit link. Late-night gaming expressed per 10 minutes; chronotype in centered hours; age, BMI, SES scaled within the diary subsample. Region excluded: all diary participants are US-only.

Within = daily deviation from person mean; Between = person mean - grand mean.

Discussion

This preregistered study examined whether late-night gaming is associated with poorer sleep quality, shorter sleep duration, greater daytime sleepiness, and lower wellbeing among adult gamers, and whether chronotype moderates these associations. Using objective telemetry linked to repeated self-report surveys, we found a consistent but small association between late-night gaming and sleep quality and little evidence that late-night gaming affects sleep duration, daytime sleepiness, or wellbeing directly. Under the same, pooled multiple-imputation analysis, chronotype did not meaningfully moderate any of the sleep- or wellbeing-related associations.

Of the four preregistered direct-effect hypotheses (H1a–H1d), only the association between late-night gaming and sleep quality was supported (H1a). Across the panel models (Table 2), each additional 10 minutes of average daily late-night gaming was associated with an approximate 0.9 percentage-point increase in the marginal probability of reporting fairly bad or very bad sleep quality. This estimate was consistent across analytic specifications: the complete-case panel model (Table A1) produced a closely matching 1.0 percentage-point increase; the exploratory diary

between-person component (Table 4) (capturing stable individual differences in habitual late-night gaming) gave a comparable 0.6 percentage-point increase in the probability of reporting poor or very poor daily sleep.

A sensitivity analysis substituting the continuous PSQI global score (Table A4) yielded directionally consistent positive fixed effects of late-night gaming on total PSQI points (imputed: $b = 0.089$, 95% CI [0.055, 0.123], $p < .001$; complete-case: $b = 0.036$, 95% CI [0.002, 0.070]), where the coefficient denotes the change in PSQI global score per additional 10 late-night minutes per day. Translated to the probability scale, these correspond to an approximate 1.0 (imputed) / 0.4 (complete-case) percentage-point increase in the probability of exceeding the PSQI poor-sleeper cutoff (> 5) at the sample mean. Linearity checks for the primary H1a specification, comparing natural-spline alternatives to the linear form on the gaming-exposure term, produced similar slopes at representative exposure levels (Table A7 Figure A1), supporting the preregistered linear parameterisation.

The convergence of these findings across different operationalisations of sleep quality, analytic strategies, and study designs strengthens our confidence that the association, though small, is genuine. Our observational design does not permit causal inference: although the association is robust across specifications, we cannot determine whether late-night gaming itself degrades sleep quality, whether poor sleepers are drawn to late-night gaming, or whether a third, uncontrolled factor drives both.

While remaining direct-effect hypotheses were not supported, these are not merely underpowered nulls: frequentist equivalence tests (TOST; [Lakens \(2017\)](#)) allow us to actively accept the null hypothesis of a negligible effect. For all three non-significant direct effects (H1b–H1d), the 90% confidence intervals fell entirely within the region of practical equivalence ($\pm 0.1 \times SD(y)$; see Table A8), providing positive evidence that any true effects of late-night gaming in this sample are too small to be of practical consequence. In substantive terms, we can rule out effects larger than roughly ± 7 minutes of nightly sleep per one-SD increase in late-night gaming (H1b) – far below the amount that would plausibly affect next-day functioning; ± 0.38 points on the 0–24 Epworth Sleepiness Scale (H1c) – much smaller than the ~ 5 -point gap separating normal from clinically excessive daytime sleepiness ([Johns, 1991](#)); and ± 0.53 points on the SWEMWBS wellbeing scale (H1d), well under the ~ 1 to 3-point change regarded as clinically meaningful ([Maheswaran et al., 2012](#)). In other words, even the largest effects compatible with our data would be undetectable against the everyday variability in sleep, mood, and alertness.

These findings speak to the sleep displacement hypothesis, which posits that late-night gaming displaces sleep time, thereby shortening sleep duration and producing daytime sleepiness ([Exelmans & Van Den Bulck, 2015](#); [Lemola et al., 2011](#); [Twenge, 2019](#)). Reduced sleep duration, the core prediction of that account is directly contradicted by our equivalence-confirmed null for H1b, and the absence of any detectable effect on daytime sleepiness (H1c) removes a key downstream consequence that would be expected if meaningful displacement were occurring. Taken together, the pattern of results argues against a displacement explanation: sleep duration and alertness were preserved, while subjective sleep quality alone was modestly worse. The residual quality effect is multiply realisable: at least three non-displacement mechanisms predict the same “quality down, duration preserved” signature. Pre-sleep physiological arousal or cognitive stimulation from gaming ([King et al., 2013](#)) is one; an attribution mechanism whereby players who game late at night perceive their sleep as poorer regardless of whether its duration or architecture is objectively altered is a second; and compensatory phase delay, in which gamers delay rather than shorten their sleep window so that duration and alertness are preserved but the sleep occurs at a circadian phase misaligned with their work-week schedule – leaving subjective quality degraded even when the hours are intact, is a third. The classical arousal account is not unambiguously preferred among these: it typically predicts disrupted sleep architecture – slow-wave and REM suppression and lengthened sleep-onset latency – with downstream daytime costs such as poorer next-day memory and alertness ([Dworak et al., 2007](#); [Higuchi et al., 2005](#); [Weaver et al., 2010](#)), and our equivalence-confirmed null on daytime sleepiness (H1c) does not see that cost. Within the preregistered equivalence bounds, the data therefore rule out displacement effects of practically meaningful magnitude and are compatible with, but do not positively single out, any one of the surviving

candidates; self-report instruments alone cannot adjudicate among them. This quality-without-duration dissociation is mirrored in contemporary adult evidence. A nationally representative US survey found that pre-bedtime device use was associated with poorer self-reported sleep quality and greater sleep deficits but showed no association with sleep duration after adjustment for sociodemographic covariates (Carlson et al., 2026); another adult sample contrasting habitual gamers (≥ 7 h/week) with occasional and non-gamers found no group differences in PSQI global score, sleep duration, or daytime sleepiness, and weekly hours of play actually predicted lower daytime sleepiness and dysfunction in the continuous analyses (De Rosa et al., 2023). Yet another adult cohort sits at the edge of this pattern, reporting elevated odds of excessive daytime sleepiness for adults who often play games before bed (Schrempft et al., 2024); the divergence may reflect design choices (dichotomised Epworth cutoff, categorical Likert exposure rather than a continuous one) that concentrate signal in heavy users in a way our continuous, population-mean equivalence test explicitly bounds.

The null finding for wellbeing (H1d) sits within a growing body of trace-data evidence that objectively logged gameplay is at most weakly related to mental wellbeing. A narrative review of 13 proposed mechanisms argues that gaming operates through offsetting pathways that may involve relaxation, need satisfaction, and social connection on one side; displacement of sleep, exercise, and social activity, plus pre-sleep arousal on the other, making a small or null net association with global wellbeing the expected result (Ballou et al., 2025b). The empirical record agrees: large telemetry studies consistently find trivial playtime–wellbeing links (Ballou et al., 2025c; Johannes et al., 2021; Vuorre et al., 2022), with perceived value of play and not hours logged driving wellbeing outcomes. Our null fits this pattern: the hedonic benefits of late-night gaming appear to offset any modest cost routed through perceived sleep quality.

Chronotype Moderation

The preregistered moderation hypotheses (H2) predicted that evening chronotype would amplify the negative associations between late-night gaming and all four outcomes. These predictions were grounded in theories of circadian misalignment and social jetlag, which hold that evening chronotypes are especially affected by late-night stimulation because it exacerbates the existing misalignment between their endogenous sleep–wake rhythm and socially imposed schedules (Kortesoja et al., 2023; Zhong et al., 2025). None of the preregistered moderation hypotheses were supported under the pooled multiple-imputation analysis: the chronotype \times late-night gaming interaction was non-significant for sleep quality (H2a: $b = -0.003$, 95% CI $[-0.013, 0.006]$, $p = 0.487$), sleep duration (H2b: $b = 0.003$, 95% CI $[-0.002, 0.008]$, $p = 0.199$), daytime sleepiness (H2c: $b = -0.009$, 95% CI $[-0.024, 0.006]$, $p = 0.225$), and wellbeing (H2d: $b = -0.009$, 95% CI $[-0.019, 0.002]$, $p = 0.101$).

Frequentist equivalence tests with a ROPE rescaled to the coefficient's native units ($\pm 0.1 \times SD(y) / SD(x)$; see Table A8) supported practical equivalence for the H2b, H2c, and H2d interactions. The H2a interaction was undecided: the 90% CI was narrow but extended marginally beyond the lower bound of the latent-scale ROPE.

The exploratory diary data told a similar story at the trait level: the between-person chronotype \times late-night gaming interaction was small and non-significant ($b = -0.001$, 95% CI $[-0.009, 0.007]$, $p = 0.730$), indicating that the habitual association between late-night gaming and nightly sleep quality did not meaningfully vary across chronotypes. The within-person interaction did reach conventional significance ($b = -0.002$, 95% CI $[-0.004, -0.000]$, $p = 0.024$), suggesting that on days when evening-type participants gamed more than usual, the day-to-day link to sleep quality was paradoxically somewhat weaker rather than stronger; a finding that runs against the preregistered prediction and that we treat as exploratory. Notably, the diary model did show a significant main effect of chronotype on nightly sleep quality ($b = 0.035$, 95% CI $[0.005, 0.065]$, $p = 0.022$), with later chronotypes reporting worse sleep on average. A pattern of trait-level chronotype penalty for sleep quality, but no amplification of the late-night-gaming effect by chronotype is consistent with the broader circadian-misalignment literature without supporting the more specific claim that evening types are differentially harmed by late-night play.

The largely null moderation results have several possible explanations. Evening-type gamers may have developed coping mechanisms or adapted sleep routines that buffer them against the effects of late-night play, for instance, by habitually sleeping in later or napping to compensate. More likely, the circadian misalignment pathway may simply be weaker than previously assumed in adult populations who have more autonomy over their schedules than the adolescent samples on which much of the chronotype–technology literature is based (Bruni et al., 2015; Reardon et al., 2023). Consensus on how to operationalise problematic gaming in adults with greater scheduling autonomy is also still being established (Costa & Kuss, 2019; Männikkö et al., 2020), which complicates direct read-across from adolescent-focused effect estimates. This reading is corroborated by independent adult-cohort evidence: Schrempft et al. (2024) explicitly tested chronotype as both a mediator and a moderator of the bedtime-media-sleep link in 4,188 adults, and found that evening chronotypes engaged in pre-sleep media more frequently but were not differentially harmed by it, exactly the pattern our pre-registered H2 equivalence tests recover.

Contextualising the effect

Viewed at the per-person, per-night scale, the association is small in absolute terms. At the marginal rate of 0.9 percentage points per 10 minutes, even a gamer averaging an hour of play past 11pm would see only a roughly 5 percentage-point increase in the probability of reporting poor sleep quality. Against the backdrop of 64% of participants already qualifying as poor sleepers by PSQI criteria, late-night gaming appears to be a minor per-person contributor to an issue with broader origins that likely involves work schedules, general screen use, caffeine consumption, and other lifestyle factors not captured here.

It is useful to anchor the implied per-hour magnitude against other modifiable lifestyle influences on adult sleep. At roughly 5 percentage points per hour, the sleep-quality cost of late-night gaming appears to be in the same broad range as the perceived sleep-quality cost reported for a high (400 mg) dose of caffeine consumed within four hours of bedtime (Gardiner et al., 2025). This comparison is most directly interpretable for our exploratory diary substudy, which used a structurally similar single-item nightly sleep-quality rating; there, the between-person estimate (0.6 percentage points per 10 min of habitual late-night gaming) aligned closely with the panel PSQI result. A comparable benchmark on the continuous PSQI scale comes from Yan et al. (2024), who report a ≈ 0.39 -point higher PSQI total score for adults in the latest vs earliest tertile of last-meal time (a ≈ 3 -hour shift). For a contrast of similar practical magnitude in our data – moving from ~ 10 min/day to ~ 3 h/day of late-night gaming – the panel PSQI sensitivity model implies a ≈ 1.52 -point increase, of comparable order but somewhat smaller, suggesting that the per-person sleep-quality penalty of habitual late-night gaming sits below that of habitually late evening meal timing.

That we found this effect specifically for sleep quality – rather than duration, sleepiness, or wellbeing – narrows the field to mechanisms that disturb perceived sleep without producing a detectable downstream cost, namely the three candidates flagged above (pre-sleep arousal, attribution, and compensatory phase delay). The exploratory diary decomposition further constrains which of these is most plausible. The H1a-analog signal is carried almost entirely by the between-person component ($b = 0.029$, 95% CI [0.012, 0.047], $p = 0.001$), with the within-person, same-night estimate non-significant ($b = -0.003$, 95% CI [-0.007, 0.002], $p = 0.214$): habitually late-night-gaming participants reported worse sleep on average, but nights on which a given participant gamed more than their own average were not reliably followed by worse sleep that night. An acute pre-sleep arousal mechanism predicts precisely the latter within-person, same-night degradation, so the diary pattern sits more comfortably with trait-level processes such as a durable attributional style, stable lifestyle confounding, or selection of more arousable players into habitual late-night play than with acute arousal per se.

A small per-night effect can nonetheless carry weight when it is sustained over time and aggregated across a common behaviour. Restricting sleep by only about an hour can still disturb emotion regulation (Tomaso et al., 2021) and undermine cognitive and behavioural performance (Belenky et al., 2003); if that restriction is maintained for two weeks, losses in alertness and working memory can rival those seen after an entire night without sleep (Van Dongen et al., 2003). Repeated

small shifts in nightly sleep quality from habitual late-night play may therefore matter for mood, vigilance, and cognition over weeks and months, even if any single night looks unremarkable. Benchmarks tailored to sleep outcomes support this interpretation: Panjeh and colleagues' recalibration from 72 effect sizes in 65 sleep-quality intervention trials (Panjeh et al., 2023) maps "small", "medium", and "large" to Cohen's d of roughly 0.18, 0.33, and 0.56 — below the usual 0.2 / 0.5 / 0.8 anchors. Judged against those sleep-specific standards, the cumulative impact of 1–2 hours of late-night gaming falls in the small-to-medium band, whereas the same pattern might look small on the generic Cohen scale.

Population burden adds another layer. De Rosa and colleagues 2024 note that video gaming is now one of the most widespread adult leisure exposures and that sleep outcomes track how intensively and how often people play, so a shallow per-night gradient can still shift the upper tail of poor sleep when layered onto a common behaviour. Public-health importance depends on both per-person effect magnitude and how prevalent the behaviour is and on this combined view, our findings sit alongside a broader literature reporting statistically detectable but per-person modest associations between digital media use and psychological outcomes.

A key gap in the existing literature is the near-complete absence of timing-specific measurement: Kristensen et al. (2021) noted that none of the studies in their systematic review registered the time of day gaming took place. Most adult gaming-and-sleep studies consequently operationalise exposure as aggregate daily volume or as binary/categorical problematic-gaming contrasts, which are likely to dilute any signal concentrated in the pre-sleep window. A parallel gap arises at the device-specific level: even a recent nationally representative US adult survey (Carlson et al., 2026) collapsed pre-bedtime TV and gaming-console use into a single category, an acknowledged limitation given that passively watching TV and actively playing video games imply very different levels of interactivity and likely have differential impacts on sleep. Our telemetry-based, continuous, time-of-day-specific operationalisation of late-night play is deliberately designed to isolate exactly that window, and the fact that we find a small but robust association only for sleep quality is consistent with reviews identifying pre-sleep timing, arousal, and session duration as key moderators of gaming's effect on sleep (De Rosa et al., 2024; Kemp et al., 2021). That the effect is detectable at all under a stringent, preregistered, multiply-imputed longitudinal specification suggests that timing-specific exposures capture signal that aggregate-volume measures may miss.

Limitations

Constraints bound the scope of our findings. First, our telemetry captures console and PC play but excludes mobile platforms, which account for a substantial share of late-night leisure — particularly in-bed use that bypasses "household curfews" and serves as a pre-sleep emotional-regulation tool. Late-night minutes that shifted to phones therefore go unmeasured, likely attenuating observed associations, and we cannot separate gaming-specific effects from the broader bundle of nocturnal screen behaviours (short-form video, messaging, doomscrolling) that share the same devices.

Second, even within the platforms we do observe, session-level logs cannot distinguish active engagement from idle time, an ambiguity that is especially acute at night when consoles may stay on while players drift off or step away. This can inflate exposure and produce spurious late-night "activity" that does not reflect cognitive arousal. Relatedly, session logs do not record game content or in-session intensity, so arousal-inducing genres (competitive, horror, fast-paced action) cannot be separated from more routine play — a moderator that adult-focused reviews identify as plausibly shaping sleep effects (De Rosa et al., 2024; Kemp et al., 2021).

Third, the analytic sample is restricted to adults recruited from UK- and US-based Prolific panels. Younger players have different sleep physiology, school schedules, parental oversight, and motivational profiles, and our estimates should not be extrapolated to pediatric or older populations. Prolific panels themselves are self-selected online workers with reliable broadband, high digital literacy, and tolerance for repeated surveys; eligibility filters around residency, language, and compliance further exclude players in other linguistic and regulatory contexts, as well as shift workers, caregivers, lower-income players, and those engaged in more extreme or stigmatized gaming profiles. Taken together, these constraints mean our findings speak most directly to

digitally engaged young adults in the UK and US who play on major console and PC ecosystems; extension to other regions, life stages, device ecologies should await complementary data sources.

Future Directions

Future work would benefit from combining telemetry with finer-grained metadata on game content, arousal level, and session proximity to bedtime, and from extending logging to mobile platforms so that the full pre-sleep screen bundle can be measured rather than inferred. Linking telemetry to ambulatory physiological measurements (e.g., consumer-grade actigraphy, heart-rate variability, or ambient light sensing) would additionally allow researchers to separate subjective perception of sleep disturbance from objective architecture, which our self-report instruments could not disentangle, and would be the natural way to adjudicate among the three non-displacement candidates the present design leaves on the table – acute pre-sleep arousal (which predicts within-person, same-night HRV elevation, lengthened sleep-onset latency, and reduced slow-wave activity), attribution (which predicts no such physiological signature), and compensatory phase delay (which predicts a within-person delay in actigraphic sleep onset and offset without a duration loss). Such measurements would also index mechanistic pathways our session-level logs cannot – notably evening light exposure, which has been proposed as a contributor to gaming-related sleep disturbance alongside physiological arousal (Hartmann et al., 2019). The rapid growth of consumer wearables (smartwatches, fitness trackers, and rings) now reach a substantial share of the populations from which gaming cohorts are recruited makes participant-initiated data donation an increasingly tractable route to such measurements. Rather than provisioning research-grade devices to every participant, future telemetry studies could invite players to donate sleep and activity records from devices they already wear, mirroring the platform-level donation model used here for play behaviour. This would lower per-participant cost, extend follow-up windows beyond what dedicated instrumentation typically allows, and bring objective sleep estimates into the same within-person, high-frequency frame as the telemetry itself, while raising its own questions about device heterogeneity, missingness, and the selection pressures introduced by wearable ownership.

Second, the observational design of this study limits the strength of causal claims that can be drawn from any single estimate. One natural next step is to embed telemetry-based exposures within a target trial emulation framework (Hernán et al., 2022; Hernán & Robins, 2016), which makes the hypothetical randomized experiment explicit (e.g., a late-night play session versus none on a given weekday). Target trial emulation is particularly well suited to device-use research because the relevant question is rarely “do gamers sleep worse than non-gamers” but rather “does the same person sleep worse on nights they play late than on nights they do not.” High-frequency telemetry naturally supports such within-person contrasts, and pairing it with g-methods or sequential exchangeability assumptions would allow future work to move from associational gradients toward more interpretable, policy-relevant causal estimates while keeping each individual as their own control.

The telemetry–survey linkage protocol piloted in Open Play is, in principle, portable: with appropriate ethical safeguards, assent procedures, and clinical validation, an analogous continuous-telemetry approach could be used to study adolescents and problematic gamers, populations in which the timing-specific effects of late-night play are most hotly contested and where aggregate self-report measures are especially vulnerable to recall bias. Extending the present methodology to those populations rather than extrapolating our estimates to them is, in our view, the most productive next step.

Data, Code, and Materials Availability

The parent Open Play dataset (Ballou et al., 2025a), including the Nintendo, Xbox, and Steam telemetry and the intake, biweekly, and daily survey tables used here, is openly released at <https://github.com/digital-wellbeing/open-play>. All analytical code for this study is openly released under MIT (code) and CC BY 4.0 (text) at <https://github.com/digital-wellbeing/platform-study-rr-sleep>.

References

A1 Appendix

A1.1 Panel H1 Sensitivity Analysis: Complete-Case

Table A1. Panel H1 Effects of Late-Night Gaming – Complete-Case (Non-Imputed Data)

| | H1a: Sleep Quality | H1b: Sleep Duration | H1c: Daytime Sleepiness | H1d: Wellbeing |
|--|-------------------------|-------------------------|-------------------------|----------------------|
| Daily LN gaming (per 10 min/day, monthly) | 0.06 [0.03, 0.08]*** | -0.00 [-0.02, 0.01] | 0.02 [-0.02, 0.06] | |
| Daily LN gaming (per 10 min/day, biweekly) | | | | -0.02 [-0.05, 0.01] |
| Age (scaled) | 0.24 [-0.12, 0.60] | -0.45 [-0.62, -0.28]*** | -0.80 [-1.36, -0.24]** | 0.07 [-0.58, 0.71] |
| BMI (scaled) | 0.23 [0.08, 0.39]** | -0.09 [-0.16, -0.01]* | 0.24 [-0.00, 0.48]+ | -0.20 [-0.48, 0.07] |
| SES (scaled) | -0.26 [-0.42, -0.11]*** | -0.10 [-0.17, -0.03]** | 0.13 [-0.11, 0.38] | 0.94 [0.67, 1.21]*** |
| Region: US | -0.23 [-0.50, 0.05] | 0.07 [-0.06, 0.20] | 0.29 [-0.13, 0.72] | 0.17 [-0.31, 0.66] |
| Day: Weekend | 0.04 [-0.16, 0.24] | 0.02 [-0.08, 0.11] | 0.00 [-0.30, 0.31] | -0.01 [-0.27, 0.24] |
| SD (Intercept Participant) | 2.01 | 0.96 | 3.12 | 4.23 |
| SD (Residual) | | 0.66 | 2.16 | 2.81 |
| SD (Intercept Gender) | | | 0.46 | 1.20 |
| N Obs | 2482 | 2482 | 2482 | 5704 |
| N Participants | 1102 | 1102 | 1102 | 1469 |
| ICC | 0.80 | 0.68 | 0.68 | 0.71 |

+ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

LN = late-night. Confidence intervals shown in brackets.

ICC = Intraclass Correlation Coefficient (adjusted).

A1.2 Panel H2 Sensitivity Analysis: Complete-Case*Table A2.* Panel H2 Chronotype Moderation – Complete-Case (Non-Imputed Data)

| | H2a: Sleep Quality | H2b: Sleep Duration | H2c: Daytime Sleepiness | H2d: Wellbeing |
|--|---------------------|-------------------------|-------------------------|-----------------------|
| Daily LN gaming (per 10 min/day, monthly) | 0.07 [0.02, 0.12]** | -0.02 [-0.04, 0.01] | 0.04 [-0.02, 0.10] | |
| Daily LN gaming (per 10 min/day, biweekly) | | | | 0.03 [-0.01, 0.07] |
| Chronotype (h, centered) | 0.02 [-0.04, 0.08] | -0.02 [-0.05, 0.01] | 0.02 [-0.07, 0.11] | -0.05 [-0.15, 0.04] |
| LN gaming × Chronotype (h, monthly) | -0.01 [-0.02, 0.00] | 0.00 [-0.00, 0.01]+ | -0.01 [-0.03, 0.00] | |
| LN gaming × Chronotype (h, biweekly) | | | | -0.01 [-0.02, -0.00]* |
| Age (scaled) | 0.01 [-0.44, 0.46] | -0.43 [-0.67, -0.19]*** | -1.18 [-1.90, -0.46]** | 0.16 [-0.69, 1.01] |
| BMI (scaled) | 0.04 [-0.15, 0.24] | -0.06 [-0.15, 0.04] | 0.19 [-0.09, 0.48] | -0.04 [-0.37, 0.30] |
| SES (scaled) | -0.11 [-0.32, 0.09] | -0.11 [-0.21, -0.01]* | 0.22 [-0.08, 0.52] | 1.05 [0.71, 1.40]*** |
| Region: US | -0.16 [-0.52, 0.19] | -0.02 [-0.20, 0.16] | 0.51 [-0.02, 1.05]+ | 0.26 [-0.36, 0.88] |
| Day: Weekend | -0.01 [-0.27, 0.25] | 0.05 [-0.08, 0.17] | -0.02 [-0.39, 0.35] | -0.08 [-0.40, 0.24] |
| SD (Intercept Participant) | 4.51 | 1.02 | 3.03 | 4.12 |
| SD (Intercept Gender) | | 0.04 | 0.33 | 0.85 |
| SD (Residual) | | 0.68 | 2.03 | 2.79 |
| N Obs | 1520 | 1520 | 1520 | 3462 |
| N Participants | 673 | 673 | 673 | 860 |
| ICC | 0.95 | 0.69 | 0.69 | 0.69 |

+ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

LN = late-night. Confidence intervals shown in brackets.

ICC = Intraclass Correlation Coefficient (adjusted).

A1.3 Panel: Wave-Level Missingness

We summarised the extent of missingness for the key self-report outcomes across each survey wave in the raw data before imputation. The *Observations* column reports the number of participants in a wave who completed at least one of the listed measures; the percentages in each row are calculated relative to that wave-specific participant count.

Table A3. Wave-level missingness for key self-report measures (pre-imputation data)

| Wave | Observations | Sleep quality (PSQI item 6) | Sleep duration (hours) | Daytime sleepi- ness (ESS) | Wellbeing (SWEMWBS) |
|------|--------------|--------------------------------|---------------------------|-------------------------------|------------------------|
| 1 | 1578 | | | | 1 (0.1%) |
| 2 | 1578 | 458 (29.0%) | 471 (29.8%) | 472 (29.9%) | 455 (28.8%) |
| 3 | 1578 | | | | 530 (33.6%) |
| 4 | 1578 | 663 (42.0%) | 674 (42.7%) | 669 (42.4%) | 661 (41.9%) |
| 5 | 1578 | | | | 740 (46.9%) |
| 6 | 1578 | 876 (55.5%) | 881 (55.8%) | 888 (56.3%) | 872 (55.3%) |

A1.4 Panel: PSQI Global Score Sensitivity Analysis

This section presents a sensitivity analysis using the PSQI global score as an alternative sleep quality outcome. The PSQI global score is the sum of all 7 PSQI component scores (range 0-21, higher = worse sleep quality), providing a continuous measure compared to the ordinal PSQI item 6 outcome used in the pre-registered H1a hypothesis.

Table A4. Sensitivity Analysis: PSQI Global Score Models (Imputed vs. Original)

| | PSQI Global | | | |
|---|-----------------------|-----------------------|-------------------------|------------------------|
| | Play (Imp.) | Chrono × Play (Imp.) | Play (CC) | Chrono × Play (CC) |
| Daily LN gaming (per 10 min/day, monthly) | 0.09 [0.05, 0.12]*** | 0.09 [0.04, 0.15]*** | 0.04 [0.00, 0.07]* | 0.04 [-0.00, 0.09]+ |
| Age (scaled) | 1.44 [1.04, 1.83]*** | 1.16 [0.63, 1.69]*** | 0.51 [0.07, 0.95]* | 0.15 [-0.44, 0.74] |
| BMI (scaled) | 0.14 [-0.03, 0.31] | 0.01 [-0.20, 0.22] | 0.31 [0.13, 0.50]** | 0.14 [-0.09, 0.37] |
| SES (scaled) | -0.18 [-0.34, -0.01]* | -0.24 [-0.46, -0.02]* | -0.41 [-0.60, -0.22]*** | -0.39 [-0.64, -0.15]** |
| Chronotype (h, centered) | | 0.01 [-0.05, 0.07] | | 0.10 [0.03, 0.18]** |
| LN gaming × Chronotype (h, monthly) | | 0.00 [-0.01, 0.01] | | -0.01 [-0.02, 0.00] |
| Region: US | -0.20 [-0.49, 0.10] | -0.17 [-0.55, 0.22] | -0.38 [-0.71, -0.04]* | -0.42 [-0.86, 0.02]+ |
| Day: Weekend | 0.23 [-0.03, 0.48]+ | 0.24 [-0.09, 0.56] | -0.23 [-0.46, 0.00]+ | -0.19 [-0.48, 0.10] |
| SD (Residual) | | | 1.60 | 1.55 |
| SD (Intercept Participant) | | | 2.47 | 2.52 |
| SD (Intercept Gender) | | | 0.71 | 0.55 |
| N Obs | 4410 | 2580 | 2482 | 1520 |
| N Participants | 1470 | 860 | 1102 | 673 |
| ICC | 0.51 | 0.51 | 0.72 | 0.73 |

+ p < 0.10, * p < 0.05, ** p < 0.01, *** p < 0.001

PSQI global score range: 0-21 (higher = worse sleep quality)

Play = late-night playtime; Chrono = chronotype; Imp. = imputed; CC = complete-case.

Confidence intervals shown in brackets.

ICC = Intraclass Correlation Coefficient (adjusted).

A1.5 Diary: Daily Diary Subsample*Table A5. Daily Diary Subsample Characteristics*

| Characteristic | Total | Analytical |
|--------------------------------------|---------------|---------------|
| A. Sociodemographics | | |
| N participants | 1275 | 1271 |
| N diary entries | 16131 | 15842 |
| Diary entries per person (Mdn, IQR) | 8 (20) | 8 (21) |
| Age | 26.6 (4.9) | 26.6 (4.9) |
| Gender | | |
| Man | 746 (58.5%) | 744 (58.5%) |
| Non-binary or other gender identity | 78 (6.1%) | 78 (6.1%) |
| Woman | 384 (30.1%) | 382 (30.1%) |
| BMI (scaled) | -0.00 (1.00) | 0.00 (1.00) |
| SES (scaled) | -0.00 (1.00) | 0.00 (1.00) |
| B. Chronotype | | |
| No alarm on free days | 657 (77.8%) | 657 (77.8%) |
| MSF sc (HH:MM) | 06:10 (03:22) | 06:10 (03:22) |
| C. Gaming | | |
| LN gaming (min/day, Mdn, IQR) | 0.0 (15.1) | 0.0 (15.0) |
| % entries with any LN gaming | 19.1 (29.3) | 19.2 (29.4) |
| D. Sleep Quality Distribution | | |
| Very poor | 836 (5.3%) | 836 (5.3%) |
| Poor | 2436 (15.4%) | 2436 (15.4%) |
| Fair | 5575 (35.2%) | 5575 (35.2%) |
| Good | 5494 (34.7%) | 5494 (34.7%) |
| Very good | 1501 (9.5%) | 1501 (9.5%) |

Values are M (SD) unless noted. LN = late-night. Sleep quality distribution counts are at the diary-entry level.

A1.6 Diary: Models – Complete-Case*Table A6. Diary H1 and H2 Probit CLMM – Complete-Case (Non-Imputed Data)*

| | H1a: Sleep Quality (direct) | H2a: Sleep Quality (chronotype mod.) |
|---------------------------------------|-----------------------------|--------------------------------------|
| LN gaming within-person (per 10 min) | -0.00 [-0.01, 0.00] | 0.00 [-0.01, 0.01] |
| LN gaming between-person (per 10 min) | 0.03 [0.01, 0.04]** | 0.02 [-0.01, 0.05] |
| Chronotype (h, centered) | | 0.03 [0.01, 0.06]* |
| LN within × Chronotype | | -0.00 [-0.00, -0.00]* |
| LN between × Chronotype | | -0.00 [-0.01, 0.01] |
| Age (scaled) | 0.03 [-0.04, 0.09] | -0.05 [-0.15, 0.05] |
| BMI (scaled) | 0.07 [0.00, 0.13]* | 0.07 [-0.02, 0.17] |
| SES (scaled) | -0.12 [-0.19, -0.06]*** | -0.14 [-0.25, -0.04]** |
| Day: Weekend | -0.00 [-0.04, 0.04] | -0.01 [-0.06, 0.05] |
| N Obs | 14683 | 8392 |
| N Participants | 1132 | 509 |
| ICC | 0.48 | 0.52 |

+ p < 0.10, * p < 0.05, ** p < 0.01, *** p < 0.001

LN = late-night. Confidence intervals shown in brackets.

Complete-case estimates – no imputation applied to the diary outcome.

Cumulative link mixed models (random intercept for participant) on 5-level ordinal sleep quality (positive coefficients = higher probability of worse sleep). Probit link. Late-night gaming per 10 minutes; chronotype in centered hours; age, BMI, SES scaled within the diary subsample. Region excluded: all diary participants are US-only.

Within = daily deviation from person mean; Between = person mean - grand mean.

A1.7 Panel H1 Natural Spline Sensitivity

For each panel outcome (H1a–H1d), we replace the linear gaming predictor with a natural cubic spline (df = 2–6, selected by BIC) to check whether the pre-registered linearity assumption holds. BIC is the primary criterion because its stronger complexity penalty matches the question this check is asking – whether non-linearity is warranted, not whether it yields any in-sample gain; AIC is reported alongside but does not drive df selection. BIC differences are interpreted via [Jones et al. \(2001\)](#)’s rule of thumb on $2|\Delta\text{BIC}|$: 0–2 not worth mentioning, 2–6 positive, 6–10 strong, >10 very strong. Both linear and spline models (Table A7, Figure A1) are fit on the complete-case sample.

Table A7. AIC and BIC comparison for panel H1 linear vs. natural spline models (complete-case data). Best df selected from 2–6 by BIC (primary criterion); AIC values are reported alongside for transparency. ΔAIC = Spline AIC – Linear AIC; ΔBIC = Spline BIC – Linear BIC; negative values indicate improvement over the linear specification. The Verdict column applies the [Jones et al. \(2001\)](#) rule of thumb to $2|\Delta\text{BIC}|$: \approx = not worth mentioning (0–2); + = positive (2–6); ++ = strong (6–10); +++ = very strong (>10), with the favoured model marked L (linear) or S (spline).

| Out-come | Best df | Linear AIC | Spline AIC | Linear BIC | Spline BIC | ΔAIC | ΔBIC | $2 \Delta\text{BIC} $ | Verdict |
|--------------------------|---------|------------|------------|------------|------------|--------------------|--------------------|-----------------------|-------------|
| H1a: Sleep Quality | 2 | 4153.4 | 4147.2 | 4211.5 | 4211.2 | -6.2 | -0.3 | 0.6 | \approx |
| H1b: Sleep Duration | 2 | 6895.4 | 6891.4 | 6947.8 | 6949.6 | -4.0 | 1.8 | 3.6 | 1. L |
| H1c: Day-time Sleepiness | 2 | 12762.5 | 12755.3 | 12820.7 | 12819.3 | -7.2 | -1.4 | 2.8 | 1. S |
| H1d: Wellbeing | 2 | 31144.8 | 31138.4 | 31211.2 | 31211.5 | -6.4 | 0.3 | 0.6 | \approx |

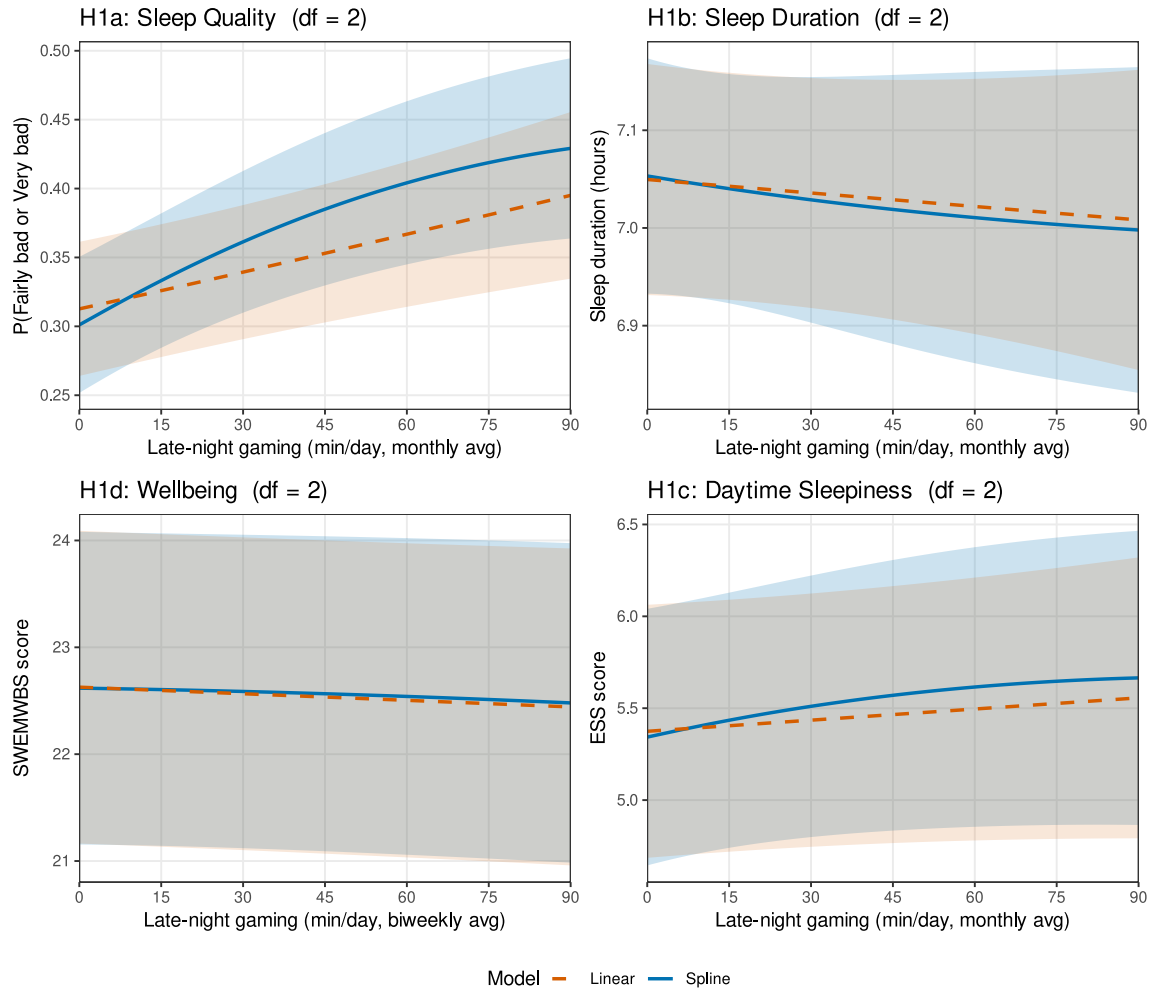


Figure A1. Marginal predicted outcomes as a function of late-night gaming (natural cubic spline, complete-case data). Solid blue line: spline fit with 95% delta-method CI ribbon. Dashed orange line: corresponding linear model fit, with its own 95% delta-method CI ribbon for comparison. X-axis shows average late-night gaming minutes per day on a common 0–90 min scale across panels; curves are drawn over the observed support within that range. All other covariates held at reference values (scaled continuous predictors at zero; isWeekend = 0; reference region). H1a: P(Fairly bad or Very bad sleep quality), marginalised over the participant random intercept via the marginal probit formula. H1b–H1d: predicted means on the original scale. Each panel title shows the selected spline df; AIC and BIC values are reported in Table A7.

A1.8 Panel: Equivalence Testing for All Confirmatory Hypotheses

We conducted frequentist equivalence tests (TOST; [Lakens \(2017\)](#); [Lakens et al. \(2018\)](#)) for every confirmatory hypothesis (H1a–H1d and H2a–H2d), including those that were statistically significant, by checking whether each pooled 90% CI falls inside a Region of Practical Equivalence (ROPE). The ROPE width itself was not preregistered; we adopt Kruschke’s ([Kruschke, 2018](#)) rule-of-thumb $\pm 0.1 \times SD(y)$ as a conventional anchor commonly used in the equivalence-testing literature. Because the focal predictors are raw late-night play-time variables (per 10 min/day) rather than standardized effect sizes, Kruschke’s default ROPE of $\pm 0.1 \times SD(y)$ is rescaled to the coefficient’s native units as $\pm 0.1 \times SD(y) / SD(x)$; for the ordinal probit models (H1a, H2a) the latent residual SD is fixed at 1 by identification, so the ROPE simplifies to $\pm 0.1 / SD(x)$ on the latent scale, which at a category threshold corresponds to a ≈ 3.99 percentage-point shift in cumulative category probability (from $\Phi(0.05) - \Phi(-0.05)$). The TOST rule accepts equivalence when the 90% CI lies entirely inside the ROPE, rejects when it lies entirely outside, and otherwise returns “Undecided”. The natural-unit interpretation of the H1b–H1d response-scale half-widths ($\approx \pm 7$ min of sleep, ± 0.38 ESS points, and ± 0.53 SWEMWBS points respectively) is given in the Discussion alongside the relevant clinical anchors. The same $\pm 0.1 \times SD(y)$ thresholds apply to the H2 interaction coefficients but now describe the change in the late-night gaming slope induced by a one-SD shift in chronotype (MSFsc); this is a stringent benchmark because the product predictor’s SD (≈ 16.53 for late-night minutes \times MSFsc) is substantially larger than that of either factor alone.

Table A8. Equivalence test results for the focal predictor in every confirmatory hypothesis. H1a–H1d test the pooled late-night gaming coefficient; H2a–H2d test the pooled chronotype \times late-night gaming interaction. Statistically significant effects (e.g., H1a) are included for completeness – a significant effect paired with an “Undecided” or “Rejected” equivalence decision indicates that the estimate, although reliably non-zero, is not small enough to be declared practically equivalent to the null under the ROPE.

| Hypothesis | ROPE | 90% CI | Decision | p |
|---|-----------------|-----------------|-----------|--------|
| H1a: Sleep Quality | [−0.028, 0.028] | [0.028, 0.075] | Undecided | < .001 |
| H1b: Sleep Duration | [−0.033, 0.033] | [−0.016, 0.006] | Accepted | 0.456 |
| H1c: Daytime Sleepiness | [−0.107, 0.107] | [−0.026, 0.053] | Accepted | 0.577 |
| H1d: Wellbeing | [−0.134, 0.134] | [−0.033, 0.024] | Accepted | 0.805 |
| H2a: Sleep Quality \times Chronotype | [−0.006, 0.006] | [−0.011, 0.005] | Undecided | 0.487 |
| H2b: Sleep Duration \times Chronotype | [−0.007, 0.007] | [−0.001, 0.007] | Accepted | 0.199 |
| H2c: Daytime Sleepiness \times Chronotype | [−0.022, 0.022] | [−0.021, 0.003] | Accepted | 0.225 |
| H2d: Wellbeing \times Chronotype | [−0.028, 0.028] | [−0.018, 0.000] | Accepted | 0.101 |

ROPE = Region of Practical Equivalence, rescaled to the native units of the raw coefficient: $\pm 0.1 \times SD(y) / SD(x)$ for linear mixed models and $\pm 0.1 / SD(x)$ for the ordinal probit models (H1a, H2a), where the latent residual SD is fixed at 1. $SD(x)$ is the sample SD of the focal predictor (product predictor for interactions).

Decision via TOST rule ([Lakens, 2017](#)): Accepted = 90% CI entirely inside ROPE; Rejected = 90% CI entirely outside ROPE; Undecided = otherwise.

p = Wald p-value for the pooled coefficient (not the TOST equivalence p).

References

- Ballou, N., Földes, T. A., Vuorre, M., Hakman, T., Magnusson, K., & Przybylski, A. (2025a). *Open Play: A longitudinal dataset of multi-platform video game digital trace data and psychological measures*. https://doi.org/10.31234/osf.io/nz96c_v1
- Ballou, N., Hakman, T., Földes, A. T., Vuorre, M., Magnusson, K., & Przybylski, A. K. (2024, October). *Psychological Wellbeing, Sleep, and Video Gaming: Analyses of Comprehensive Digital Traces*. OSF Registries. <https://osf.io/mvngt>
- Ballou, N., Hakman, T., Vuorre, M., Magnusson, K., & Przybylski, A. K. (2025b). How do video games affect mental health? A narrative review of 13 proposed mechanisms. *Technology, Mind, And Behavior*, 6(2), 123–143. <https://doi.org/10.1037/tmb0000152>
- Ballou, N., Vuorre, M., Hakman, T., Magnusson, K., & Przybylski, A. K. (2025c). Perceived value of video games, but not hours played, predicts mental well-being in casual adult Nintendo players. *Royal Society Open Science*, 12(3), 241174. <https://doi.org/10.1098/rsos.241174>
- Belenky, G., Wesensten, N. J., Thorne, D. R., Thomas, M. L., Sing, H. C., Redmond, D. P., Russo, M. B., & Balkin, T. J. (2003). Patterns of Performance Degradation and Restoration During Sleep Restriction and Subsequent Recovery: A Sleep Dose–Response Study. *Journal of Sleep Research*, 12(1), 1–12. <https://doi.org/10.1046/j.1365-2869.2003.00337.x>
- Bruni, O., Sette, S., Fontanesi, L., Baiocco, R., Laghi, F., & Baumgartner, E. (2015). Technology Use and Sleep Quality in Preadolescence and Adolescence. *Journal of Clinical Sleep Medicine*, 11(12), 1433–1441. <https://doi.org/10.5664/jcsm.5282>
- Buysse, D. J., Reynolds, C. F., Monk, T. H., Berman, S. R., & Kupfer, D. J. (1989). The Pittsburgh sleep quality index: A new instrument for psychiatric practice and research. *Psychiatry Research*, 28(2), 193–213. [https://doi.org/10.1016/0165-1781\(89\)90047-4](https://doi.org/10.1016/0165-1781(89)90047-4)
- Cain, N., & Gradisar, M. (2010). Electronic media use and sleep in school-aged children and adolescents: A review. *Sleep Medicine*, 11(8), 735–742. <https://doi.org/10.1016/j.sleep.2010.02.006>
- Carlson, S. E., Sparks, C. D., Dzierzewski, J. M., & Baron, K. G. (2026). Technology Use and Sleep: An Examination of Population Data from the National Sleep Foundation. *Sleep Health*. <https://doi.org/10.1016/j.sleh.2026.03.006>
- Costa, S., & Kuss, D. J. (2019). Current Diagnostic Procedures and Interventions for Gaming Disorders: A Systematic Review. *Frontiers in Psychology*, 10. <https://doi.org/10.3389/fpsyg.2019.00578>
- De Rosa, O., Baker, F. C., Barresi, G., Conte, F., Ficca, G., & Zambotti, M. de. (2024). Video gaming and sleep in adults: A systematic review. *Sleep Medicine*, 124, 91–105. <https://doi.org/10.1016/j.sleep.2024.09.015>
- De Rosa, O., Conte, F., D'Onofrio, P., Malloggi, S., Alterio, A., Rescott, M. L., Giganti, F., & Ficca, G. (2023). Habitual Videogame Playing Does Not Compromise Subjective Sleep Quality and Is Associated with Improved Daytime Functioning. *Brain Sciences*, 13(2), 279. <https://doi.org/10.3390/brainsci13020279>
- Dworak, M., Schierl, T., Bruns, T., & Strüder, H. K. (2007). Impact of singular excessive computer game and television exposure on sleep patterns and memory performance of school-aged children. *Pediatrics*, 120(5), 978–985. <https://doi.org/10.1542/peds.2007-0476>
- Exelmans, L., & Van Den Bulck, J. (2015). Sleep quality is negatively related to video gaming volume in adults. *Journal of Sleep Research*, 24(2), 189–196. <https://doi.org/10.1111/jsr.12255>
- Fernandes, G., Vallim, J., D'Almeida, V., Tufik, S., & Andersen, M. (2023). The effects of social jetlag and sleep variability on sleepiness in a population-based study: The mediating role of sleep debt. *Journal of Sleep Research*, 33. <https://doi.org/10.1111/jsr.14043>
- Gadie, A., Shafto, M. A., Leng, Y., & Kievit, R. (2016). How are age-related differences in sleep quality associated with health outcomes? An epidemiological investigation in a UK cohort of 2406 adults. *BMJ Open*, 7. <https://doi.org/10.1136/bmjopen-2016-014920>
- Gardiner, C. L., Weakley, J., Burke, L. M., Fernandez, F., Johnston, R. D., Leota, J., Russell, S., Munteanu, G., Townshend, A., & Halson, S. L. (2025). Dose and Timing Effects of Caffeine on Subsequent Sleep: A Randomized Clinical Crossover Trial. *Sleep*, 48(4), zsae230. <https://doi.org/10.1093/sleep/zsae230>

- Gumport, N. B., Gasperetti, C. E., Silk, J. S., & Harvey, A. G. (2021). The Impact of Television, Electronic Games, and Social Technology Use on Sleep and Health in Adolescents with an Evening Circadian Preference. *Journal of Youth and Adolescence*, 50(12), 2351–2362. <https://doi.org/10.1007/s10964-021-01429-9>
- Hale, L., & Guan, S. (2015). Screen time and sleep among school-aged children and adolescents: A systematic literature review. *Sleep Medicine Reviews*, 21, 50–58. <https://doi.org/10.1016/j.smrv.2014.07.007>
- Han, X., Zhou, E., & Liu, D. (2024). Electronic Media Use and Sleep Quality: Updated Systematic Review and Meta-Analysis. *Journal of Medical Internet Research*, 26(1), e48356. <https://doi.org/10.2196/48356>
- Hartmann, M., Pelzl, M. A., Kann, P. H., Koehler, U., Betz, M., Hildebrandt, O., & Cassel, W. (2019). The effects of prolonged single night session of videogaming on sleep and declarative memory. *Plos ONE*, 14(11), e224893. <https://doi.org/10.1371/journal.pone.0224893>
- Hernán, M. A., & Robins, J. M. (2016). Using big data to emulate a target trial when a randomized trial is not available. *American Journal of Epidemiology*, 183(8), 758–764. <https://doi.org/10.1093/aje/kwv254>
- Hernán, M. A., Wang, W., & Leaf, D. E. (2022). Target trial emulation: A framework for causal inference from observational data. *JAMA*, 328(24), 2446–2447. <https://doi.org/10.1001/jama.2022.21383>
- Higuchi, S., Motohashi, Y., Liu, Y., & Maeda, A. (2005). Effects of playing a computer game using a bright display on presleep physiological variables, sleep latency, slow wave sleep and REM sleep. *Journal of Sleep Research*, 14(3), 267–273. <https://doi.org/10.1111/j.1365-2869.2005.00463.x>
- Johannes, N., Vuorre, M., & Przybylski, A. K. (2021). Video game play is positively correlated with well-being. *Royal Society Open Science*, 8(2), 202049. <https://doi.org/10.1098/rsos.202049>
- Johns, M. W. (1991). A New Method for Measuring Daytime Sleepiness: The Epworth Sleepiness Scale. *Sleep*, 14(6), 540–545. <https://doi.org/10.1093/sleep/14.6.540>
- Jones, B. L., Nagin, D. S., & Roeder, K. (2001). A SAS Procedure Based on Mixture Models for Estimating Developmental Trajectories. *Sociological Methods & Research*, 29(3), 374–393. <https://doi.org/10.1177/0049124101029003005>
- Kemp, C., Pienaar, P. R., Rosslee, D. T., Lipinska, G., Roden, L. C., & Rae, D. E. (2021). Sleep in Habitual Adult Video Gamers: A Systematic Review. *Frontiers in Neuroscience*, 15, 781351. <https://doi.org/10.3389/fnins.2021.781351>
- King, D. L., & Delfabbro, P. (2009). Understanding and assisting excessive players of video games: A community psychology perspective. *Australian Community Psychologist*, 21(1), 62–74.
- King, D. L., Gradisar, M., Drummond, A., Lovato, N., Wessel, J., Micic, G., Douglas, P., & Delfabbro, P. (2013). The impact of prolonged violent video-gaming on adolescent sleep: An experimental study. *Journal of Sleep Research*, 22(2), 137–143. <https://doi.org/10.1111/j.1365-2869.2012.01060.x>
- Kortesoja, L., Vainikainen, M.-P., Hotulainen, R., & Merikanto, I. (2023). Late-Night Digital Media Use in Relation to Chronotype, Sleep and Tiredness on School Days in Adolescence. *Journal of Youth and Adolescence*, 52(2), 419–433. <https://doi.org/10.1007/s10964-022-01703-4>
- Kristensen, J. H., Pallesen, S., King, D. L., Hysing, M., & Erevik, E. K. (2021). Problematic Gaming and Sleep: A Systematic Review and Meta-Analysis. *Frontiers in Psychiatry*, 12. <https://doi.org/10.3389/fpsy.2021.675237>
- Kruschke, J. K. (2018). Rejecting or Accepting Parameter Values in Bayesian Estimation. *Advances in Methods and Practices in Psychological Science*, 1(2), 270–280. <https://doi.org/10.1177/2515245918771304>
- Lakens, D. (2017). Equivalence Tests: A Practical Primer for t Tests, Correlations, and Meta-Analyses. *Social Psychological and Personality Science*, 8(4), 355–362. <https://doi.org/10.1177/1948550617697177>
- Lakens, D., Scheel, A. M., & Isager, P. M. (2018). Equivalence Testing for Psychological Research: A Tutorial. *Advances in Methods and Practices in Psychological Science*, 1(2), 259–269. <https://doi.org/10.1177/2515245918770963>

- LeBourgeois, M. K., Hale, L., Chang, A.-M., Akacem, L. D., Montgomery-Downs, H. E., & Buxton, O. M. (2017). Digital Media and Sleep in Childhood and Adolescence. *Pediatrics*, *140*(Supplement_2), S92–S96. <https://doi.org/10.1542/peds.2016-1758J>
- Lemola, S., Brand, S., Vogler, N., Perkinson-Gloor, N., Allemand, M., & Grob, A. (2011). Habitual computer game playing at night is related to depressive symptoms. *Personality and Individual Differences*, *51*(2), 117–122. <https://doi.org/10.1016/j.paid.2011.03.024>
- Maheswaran, H., Weich, S., Powell, J., & Stewart-Brown, S. (2012). Evaluating the Responsiveness of the Warwick Edinburgh Mental Well-Being Scale (WEMWBS): Group and Individual Level Analysis. *Health and Quality of Life Outcomes*, *10*, 156. <https://doi.org/10.1186/1477-7525-10-156>
- McCoy, J. G., & Strecker, R. E. (2011). The cognitive cost of sleep lost. *Neurobiology of Learning and Memory*, *96*(4), 564–582. <https://doi.org/10.1016/j.nlm.2011.07.004>
- Muppalla, S. K., Vuppapapati, S., Pulliahgaru, A. R., & Sreenivasulu, H. (2023). Effects of excessive screen time on child development: An updated review and strategies for management. *Cureus*, *15*(6). <https://doi.org/10.7759/cureus.40608>
- Männikkö, N., Ruotsalainen, H., Miettunen, J., Pontes, H. M., & Käriäinen, M. (2020). Problematic Gaming Behaviour and Health-Related Outcomes: A Systematic Review and Meta-Analysis. *Journal of Health Psychology*, *25*(1), 67–81. <https://doi.org/10.1177/1359105317740414>
- Panjeh, S., Nordahl-Hansen, A., & Cogo-Moreira, H. (2023). Establishing new cutoffs for Cohen's *d*: An application using known effect sizes from trials for improving sleep quality on composite mental health. *International Journal of Methods in Psychiatric Research*, *32*(3). <https://doi.org/10.1002/mpr.1969>
- Peracchia, S., & Curcio, G. (2018). Exposure to video games: Effects on sleep and on post-sleep cognitive abilities. A systematic review of experimental evidences. *Sleep Science*, *11*(4), 302–314. <https://doi.org/10.5935/1984-0063.20180046>
- Pirrone, D., Eijnden, R. J. J. M. van den, & Peeters, M. (2024). Why We Can't Stop: The Impact of Rewarding Elements in Videogames on Adolescents' Problematic Gaming Behavior. *Media Psychology*, *27*(3), 379–400. <https://doi.org/10.1080/15213269.2023.2242260>
- Reardon, A., Lushington, K., & Agostini, A. (2023). Adolescent sleep, distress, and technology use: Weekday versus weekend. *Child and Adolescent Mental Health*, *28*(1), 108–116. <https://doi.org/10.1111/camh.12616>
- Robitzsch, A., & Grund, S. (2024). *miceadds: Some Additional Multiple Imputation Functions, Especially for mice*. <https://cran.r-project.org/package=miceadds>
- Roenneberg, T., Wirz-Justice, A., & Mellow, M. (2003). Life between Clocks: Daily Temporal Patterns of Human Chronotypes. *Journal of Biological Rhythms*, *18*(1), 80–90. <https://doi.org/10.1177/0748730402239679>
- Rubin, D. B. (1987). *Multiple Imputation for Nonresponse in Surveys*. John Wiley & Sons. <https://doi.org/10.1002/9780470316696>
- Schrempft, S., Baysson, H., Chessa, A., Lorthe, E., Zaballa, M.-E., Stringhini, S., Guessous, I., & Nehme, M. (2024). Associations between bedtime media use and sleep outcomes in an adult population-based cohort. *Sleep Medicine*, *121*, 226–235. <https://doi.org/10.1016/j.sleep.2024.06.029>
- Simon, E. B., Vallat, R., Barnes, C. M., & Walker, M. P. (2020). Sleep Loss and the Socio-Emotional Brain. *Trends in Cognitive Sciences*, *24*(6), 435–450. <https://doi.org/10.1016/j.tics.2020.02.003>
- Smith, L. J., King, D. L., Richardson, C., Roane, B. M., & Gradisar, M. (2017). Mechanisms influencing older adolescents' bedtimes during videogaming: The roles of game difficulty and flow. *Sleep Medicine*, *39*, 70–76. <https://doi.org/10.1016/j.sleep.2017.09.002>
- Spada, M. M., & Caselli, G. (2017). The Metacognitions about Online Gaming Scale: Development and psychometric properties. *Addictive Behaviors*, *64*, 281–286. <https://doi.org/10.1016/j.addbeh.2015.07.007>
- Tennant, R., Hiller, L., Fishwick, R., Platt, S., Joseph, S., Weich, S., Parkinson, J., Secker, J., & Stewart-Brown, S. (2007). The Warwick-Edinburgh Mental Well-being Scale (WEMWBS): Development and UK validation. *Health and Quality of Life Outcomes*, *5*(1), 63. <https://doi.org/10.1186/1477-7525-5-63>

- Tomaso, C. C., Johnson, A. B., & Nelson, T. D. (2021). The Effect of Sleep Deprivation and Restriction on Mood, Emotion, and Emotion Regulation: Three Meta-Analyses in One. *Sleep*, *44*(6), zsa289. <https://doi.org/10.1093/sleep/zsa289>
- Twenge, J. M. (2019). More Time on Technology, Less Happiness? Associations Between Digital-Media Use and Psychological Well-Being. *Current Directions in Psychological Science*, *28*(4), 372–379. <https://doi.org/10.1177/0963721419838244>
- van Buuren, S. (2018). *Flexible Imputation of Missing Data* (2nd ed.). Chapman, Hall/CRC. <https://doi.org/10.1201/9780429492259>
- van Buuren, S., & Groothuis-Oudshoorn, K. (2011). mice: Multivariate Imputation by Chained Equations in R. *Journal of Statistical Software*, *45*(3), 1–67. <https://doi.org/10.18637/jss.v045.i03>
- Van Dongen, H. P. A., Maislin, G., Mullington, J. M., & Dinges, D. F. (2003). The Cumulative Cost of Additional Wakefulness: Dose–Response Effects on Neurobehavioral Functions and Sleep Physiology from Chronic Sleep Restriction and Total Sleep Deprivation. *Sleep*, *26*(2), 117–126. <https://doi.org/10.1093/sleep/26.2.117>
- Vriend, J. L., Davidson, F. D., Corkum, P. V., Rusak, B., Chambers, C. T., & McLaughlin, E. N. (2013). Manipulating Sleep Duration Alters Emotional Functioning and Cognitive Performance in Children. *Journal of Pediatric Psychology*, *38*(10), 1058–1069. <https://doi.org/10.1093/jpepsy/jst033>
- Vuorre, M., Johannes, N., Magnusson, K., & Przybylski, A. K. (2022). Time spent playing video games is unlikely to impact well-being. *Royal Society Open Science*, *9*(7), 220411. <https://doi.org/10.1098/rsos.220411>
- Weaver, E., Gradisar, M., Dohnt, H., Lovato, N., & Douglas, P. (2010). The Effect of Presleep Video-Game Playing on Adolescent Sleep. *Journal of Clinical Sleep Medicine*, *6*(2), 184–189. <https://doi.org/10.5664/jcsm.27769>
- Wu, S., Peng, M., Zuo, Y., Wang, M., Feng, Z., Wang, H., Sun, M., Gu, J., Wang, X., Han, F., & Dong, X. (2025). Social jetlag and obstructive sleep apnea severity: A retrospective study based on polysomnography. *Sleep Medicine*, *134*, 106722. <https://doi.org/10.1016/j.sleep.2025.106722>
- Yan, L. M., Li, H. J., Fan, Q., Xue, Y. D., & Wang, T. (2024). Chronobiological Perspectives: Association Between Meal Timing and Sleep Quality. *Plos ONE*, *19*(8), e308172. <https://doi.org/10.1371/journal.pone.0308172>
- Zhong, C., Masters, M., Donzella, S. M., Diver, W. R., & Patel, A. V. (2025). Electronic Screen Use and Sleep Duration and Timing in Adults. *JAMA Network Open*, *8*(3), e252493. <https://doi.org/10.1001/jamanetworkopen.2025.2493>