Social media use patterns and self-harm among secondary school students in England

Lucas Sempé[1,2,*] Mina Fazel[2] Galit Geulayov[2]
Emma Soneson[2]

[1] International Initiative for Impact Evaluation (3ie), United Kingdom.
[2] Department of Psychiatry, Oxford University, Oxford, United Kingdom.

[*] Correspondence: Lucas Sempé <lsempe@3ieimpact.org>

Abstract

Background. Self-harm in adolescence is a growing public health problem, yet most research treats social media as a single, undifferentiated exposure — typically screen time — overlooking differences across platforms, account settings, and parental involvement. We examined whether distinct dimensions of social media use are associated with self-harm among secondary school students in England.

Methods. We analysed data from 24,909 students (Years 7–13, ages 11–19) who participated in the 2023 OxWell Student Survey, a school-based cross-sectional survey across four English regions. Self-harm was classified using a combined approach incorporating structured self-report questions and expert panel review of free-text responses. We examined four dimensions of social media engagement: daily time spent, publicly available accounts, platform-specific use (16 platforms), and parental mediation strategies (9 behaviours). We fitted multilevel logistic regression models with random intercepts for schools, adjusting for gender, year group, ethnicity, care status, socioeconomic deprivation, and neurodivergence. We report adjusted odds ratios, predicted probabilities, and E-values for unmeasured confounding.

Findings. Each additional hour of daily social media use was associated with increased odds of self-harm (adjusted OR 1.19, 95% CI 1.17–1.21). Maintaining a publicly available account was associated with 2.02-fold higher odds (aOR 2.02, 95% CI 1.85–2.21). Substantial platform-level heterogeneity was observed: Omegle.X2350fTicked (Yes) use was associated with the highest odds (aOR 2.8) and a predicted probability of 35%, while X2350m showed the lowest (~17%). Among parental mediation strategies, supportive parental behaviours were

associated with approximately 3–34% lower odds of self-harm, whereas restrictive approaches showed marginally elevated odds (aOR 1.5); however, the cross-sectional design precludes causal interpretation.

Interpretation. Platform type and parental engagement style show stronger associations with self-harm than screen time alone. These findings support platform-aware prevention strategies and suggest that supportive, rather than restrictive, parental involvement merits further investigation in longitudinal designs.

Funding. *[To be completed]*

Research in Context

Evidence before this study. Existing systematic reviews and umbrella reviews have found modest associations between social media use and depressive symptoms, but evidence linking social media directly to self-harm is sparse and inconsistent. Nearly all studies measure social media as undifferentiated screen time, without distinguishing platforms, account configurations, or whether parents are involved. To our knowledge, no large study has examined platform-specific associations, account privacy, and parental mediation simultaneously using multilevel models that account for school-level clustering.
Added value of this study. Using data from over 24,000 secondary school students across four English regions, this study disaggregates the association between social media and self-harm across four dimensions: time spent, platform choice, account publicity, and parental mediation. Anonymous and visually oriented platforms were associated with the highest odds of self-harm; supportive parental engagement was associated with meaningfully lower odds, in contrast with restrictive approaches. Multilevel models with weakly informative priors and E-values are used throughout.
Implications of all the available evidence. Aggregate screen time captures part of the association between social media and self-harm, but platform type, account visibility, and parental engagement each contribute additional information for risk stratification. Prevention strategies that move beyond time-based restrictions to address platform-specific risks and the quality of parental engagement may be more effective.

Introduction

Self-harm among adolescents — including non-suicidal self-injury and suicidal acts — is a growing public health problem. Incidence has risen across the United Kingdom and other Western countries in recent decades (Morgan et al. 2017; Griffin et al. 2018), making the identification of modifiable risk and protective factors a priority.

Social media use has attracted particular attention as a potential risk factor, given how central online platforms have become to adolescent life. Research has linked time on social media to

depressive symptoms (Cunningham, Hudson, and Harkness 2021), social isolation (O'Day and Heimberg 2021), and cyberbullying victimisation (Giumetti and Kowalski 2022; Hamm et al. 2015) — all established correlates of self-harm. An umbrella review of the evidence concluded that while associations between social media use and mental health outcomes are generally small, they are consistent across studies (Valkenburg, Meier, and Beyens 2022). Yet how adolescents use social media likely matters as much as how much time they spend on it. The platforms they prefer, whether their profiles are public, and how parents engage with their online activity may each shape risk in ways that overall screen time does not capture (Odgers and Jensen 2020).

A small but growing body of work has begun to consider platform-specific effects. Different platforms afford different types of interaction: some prioritise anonymous, public exchanges with strangers, while others facilitate private communication with known peers. These structural differences — or "affordances" — may create distinct risk environments. Similarly, whether an adolescent maintains a publicly visible account, rather than a private one, may affect exposure to unsolicited contact, cyberbullying, and harmful content (Mabaso et al. 2024).

Parental engagement with adolescents' online activity represents another dimension that has received limited attention. The parental mediation literature distinguishes between restrictive approaches (limiting access or time) and active or supportive approaches (discussing online experiences, helping navigate distressing content). Evidence from media research suggests that active mediation is more consistently associated with positive outcomes than restriction, but few studies have examined these strategies in relation to self-harm specifically.

Despite a growing literature, several gaps remain. Most studies treat social media as a single variable — typically screen time — without distinguishing between platforms, account configurations, or parental involvement. Few have examined parental mediation in relation to self-harm specifically, and fewer still have used multilevel models to account for clustering within schools.

This paper addresses these gaps using data from the 2023 OxWell Student Survey. We examine four dimensions of social media engagement — daily time online, public account ownership, platform-specific use, and parental mediation — in relation to self-harm among secondary school students across four English regions. For each dimension, we fit multilevel logistic regression models adjusting for gender, year group, ethnicity, care status, deprivation, and neurodivergence, and report odds ratios and predicted probabilities.

Methods

Study design and participants

The OxWell Student Survey is a repeated cross-sectional survey of students sampled from schools across four regions in England, as described in the study protocol (Mansfield et al.

2021). Schools were recruited through partnerships with local authorities, multi-academy trusts, and clinical commissioning groups; within participating schools, all students in the target year groups were eligible to take part. The survey collects data on a broad range of topics relating to mental health and well-being, life experiences, and health-related behaviours. The data analysed in this paper were collected from students in school Years 7 to 13 (ages approximately 11 to 19) during 2023. Participation was voluntary, and no monetary incentives were offered. The study received ethical approval from the University of Oxford Medical Sciences Interdivisional Research Ethics Committee (Ref: R67020/RE001). Informed consent was obtained from school headteachers, and all participants provided assent. Parents were informed in advance and could opt their child out of participation.

The full OxWell 2023 survey was completed by 37,952 students. We excluded respondents who could not be classified on the self-harm outcome (those answering "Prefer not to say", "Not sure what this means", or with missing gateway responses), yielding an analytic sample of 24,909 students from 177 schools. Model-specific sample sizes vary depending on item-level missingness in the exposure variables (see individual model notes). All models used complete-case analysis; the extent of item-level missingness for each exposure is reported in the descriptive tables.

Measures

Recognising the complexity of identifying individuals who self-harm through self-report methods, we employed a combined classification approach. Respondents were classified as having self-harmed if they (a) answered "Yes" to a gateway question asking whether they had ever deliberately self-harmed, followed by an affirmative response to at least one of two subsequent items specifying the method of self-harm (self-injury or overdose), or (b) were classified as self-harmers by an expert panel — comprising researchers with clinical and epidemiological expertise — who independently reviewed free-text responses, following the methodology proposed by Geulayov et al. (2022). Discrepancies were resolved through consensus discussion.

We examined four dimensions of social media engagement. First, respondents reported the number of hours they typically spent per day on social networking sites or forums (e.g., Instagram, TikTok, Discord), with response options ranging from "0 hrs" to "8 hrs or more"; this was modelled as a continuous numeric variable. Second, respondents indicated whether they had a publicly available social media account (i.e., one that does not require a friend request to follow). Third, respondents indicated which platforms they used from a list of 16; a separate model was fitted for each platform, with use (Yes/No) as the predictor. Fourth, respondents answered a series of questions about their parents' or guardians' behaviour with respect to their internet use, developed by Skripkauskaite and Fazel (2022) and the OxWell Study Team; a separate model was fitted for each of nine parental behaviours.

All models were adjusted for a common set of sociodemographic covariates. Gender was classified as Boy, Girl, Gender Diverse (GD), or Gender Non-Disclosing (GND), following Soneson,

O'Leary, and Fazel (2024). Ethnicity was based on self-identification following UK Office for National Statistics categories ("Developing Admin-Based Ethnicity Statistics for England and Wales" 2023). Care status recorded whether the respondent was, or had been, a child in care, looked after, or fostered. Socioeconomic deprivation was measured using a composite index derived from principal component analysis (PCA; first component retained) of items originally developed for the Wales Young People's Survey on Child and Family Poverty 2019 and subsequently adapted for the OxWell survey, covering indicators such as food bank use, cold or damp housing, inability to afford school expenses, and going to bed hungry (individual items are listed in Table 6); the resulting index was standardised to have a mean of zero and unit variance. Neurodivergence was defined as self-reported perception of being dyslexic, dyspraxic, autistic, or having ADHD. We note that this composite measure conflates conditions with distinct aetiologies and behavioural profiles; disaggregated analyses were not feasible given the sample sizes available for individual conditions.

Statistical analysis

We fitted multilevel logistic regression models with random intercepts for schools using the `glmmTMB` package in R, accounting for the clustering of students within schools. Weakly informative priors (Normal(0, 3)) were placed on fixed-effect coefficients to stabilise estimation. For each model, we report adjusted odds ratios and predicted probabilities. Bivariate associations between self-harm and social media use variables are also presented graphically.

Separate models were fitted for each of the 16 platforms and nine parental mediation behaviours — rather than entering all indicators simultaneously — to avoid multicollinearity arising from overlapping platform use and correlated parental behaviours. Because this strategy entails 25 separate hypothesis tests, the family-wise error rate is inflated; findings for individual platforms or parental behaviours should therefore be interpreted in the context of the overall pattern rather than on the basis of any single p-value.

To assess the robustness of key findings to unmeasured confounding, we computed E-values (VanderWeele and Ding 2017), as described in the sensitivity analyses below. Gender-stratified models were fitted to assess whether associations differed between girls and boys. This study has been pre-registered (Sempé et al. 2024).

Results

Table 1 presents the characteristics of the analytic sample, stratified by self-harm classification. The sample comprised 24,909 respondents from 177 schools who could be classified as having or not having self-harmed using the combined classification described above.

Self-harm prevalence differed across sociodemographic groups. Girls and students identifying as gender diverse or gender non-disclosing reported higher rates than boys. White students reported higher rates of self-harm than students from other ethnic backgrounds. Respondents

classified as having self-harmed had higher mean deprivation index scores than those who had not, indicating greater socioeconomic disadvantage. The difference in neurodivergence was particularly marked: 0% of self-harmers considered themselves neurodivergent (dyslexic, dyspraxic, autistic, or having ADHD), compared with 0% of non-self-harmers — a difference of approximately 0 percentage points.

Social media use patterns also differed between groups. Among self-harmers, 15% reported spending eight or more hours per day on social networking sites or forums, and 41% maintained a publicly available social media account. The corresponding figure for public accounts among non-self-harmers was 25%, suggesting a substantial gap in account visibility between the two groups even before adjusting for confounders.

Table 1: Descriptive characteristics by self-harm classification

Characteristic	**N**	**Yes** N = 4,479	**No** N = 20,430	**(Missing)** N = 13,043
Have you ever deliberately self-harmed?	33,737			
Yes		4,479 (100%)	0 (0%)	1,707 (19%)
No		0 (0%)	20,430 (100%)	0 (0%)
Not sure what this means		0 (0%)	0 (0%)	883 (10%)
Prefer not to say		0 (0%)	0 (0%)	4,164 (47%)
Skipped by respondent		0 (0%)	0 (0%)	2,074 (23%)
Missing		0	0	4,215
Year group	37,952			
Y05		214 (4.8%)	2,198 (11%)	1,977 (15%)
Y06		249 (5.6%)	2,317 (11%)	1,726 (13%)
Y07		579 (13%)	3,354 (16%)	2,347 (18%)
Y08		659 (15%)	3,219 (16%)	1,861 (14%)
Y09		728 (16%)	3,103 (15%)	1,630 (12%)
Y10		620 (14%)	2,227 (11%)	1,256 (9.6%)
Y11		686 (15%)	2,041 (10.0%)	1,085 (8.3%)
Y12		449 (10%)	1,192 (5.8%)	744 (5.7%)
Y13		295 (6.6%)	779 (3.8%)	417 (3.2%)
Gender	37,569			
Boy		1,225 (28%)	10,274 (51%)	5,829 (45%)
Girl		2,775 (63%)	9,574 (47%)	6,268 (49%)
Gender Diverse (GD)		197 (4.5%)	70 (0.3%)	143 (1.1%)
Gender Non-Disclosing (GND)		221 (5.0%)	370 (1.8%)	623 (4.8%)
Missing		61	142	180
Ethnic background	37,946			
Asian/Asian British (aggregated)		451 (10%)	3,156 (15%)	1,682 (13%)
Black/Black British/African/Caribbean (aggregated)		161 (3.6%)	991 (4.9%)	580 (4.4%)
Mixed/Multiple Ethnic Groups (aggregated)		298 (6.7%)	1,015 (5.0%)	753 (5.8%)
Other ethnic group		129 (2.9%)	849 (4.2%)	674 (5.2%)
White (aggregated)		2,711 (61%)	10,850 (53%)	6,715 (52%)
Skipped by respondent		729 (16%)	3,569 (17%)	2,633 (20%)
Missing		0	0	6
Child in care, looked after, or fostered?	37,938			
Yes		159 (3.5%)	646 (3.2%)	534 (4.1%)

(continued)

Characteristic	N	Yes N = 4,479	No N = 20,430	(Missing) N = 13,043
No		3,482 (78%)	14,026 (69%)	7,650 (59%)
I don't know what this means		230 (5.1%)	985 (4.8%)	793 (6.1%)
Not now, but I used to be in care		80 (1.8%)	126 (0.6%)	128 (1.0%)
Prefer not to say		61 (1.4%)	106 (0.5%)	174 (1.3%)
Skipped by respondent		467 (10%)	4,541 (22%)	3,750 (29%)
Missing		0	0	14
Neurodivergent (dyslexic/dyspraxic, autistic, ADHD)	37,938			
Yes		1,405 (31%)	2,225 (11%)	2,156 (17%)
No		1,469 (33%)	10,822 (53%)	4,498 (35%)
Not sure		1,061 (24%)	2,631 (13%)	2,313 (18%)
Prefer not to say		73 (1.6%)	197 (1.0%)	299 (2.3%)
Skipped by respondent		471 (11%)	4,555 (22%)	3,763 (29%)
Missing		0	0	14
Deprivation index	36,292	5.73 (3.44)	5.44 (3.39)	5.34 (3.45)
Missing		351	763	546
Hours per day on social networking sites or forums	31,531			
0 hrs		203 (4.9%)	1,478 (7.7%)	332 (4.0%)
30 mins		185 (4.5%)	1,458 (7.6%)	324 (3.9%)
1 hr		220 (5.3%)	1,558 (8.2%)	333 (4.0%)
1 hr 30 mins		184 (4.4%)	1,129 (5.9%)	283 (3.4%)
2 hrs		359 (8.6%)	1,987 (10%)	398 (4.8%)
3 hrs		467 (11%)	1,920 (10%)	466 (5.6%)
4 hrs		426 (10%)	1,458 (7.6%)	437 (5.3%)
5 hrs		380 (9.1%)	1,016 (5.3%)	344 (4.1%)
6 hrs		297 (7.1%)	660 (3.5%)	255 (3.1%)
7 hrs		141 (3.4%)	319 (1.7%)	136 (1.6%)
8 hrs or more		652 (16%)	896 (4.7%)	592 (7.1%)
Skipped by respondent		643 (15%)	5,201 (27%)	4,394 (53%)
Missing		322	1,350	4,749
Publicly available social media account	31,259			
Yes		1,833 (44%)	5,033 (27%)	1,726 (21%)
No		1,655 (40%)	8,576 (45%)	2,099 (25%)
Skipped by respondent		632 (15%)	5,289 (28%)	4,416 (54%)
Missing		359	1,532	4,802

[1] n (%); Mean (SD)

Before fitting adjusted models, we examined bivariate associations between self-harm classification and each social media exposure. Figure 1 shows the distribution of daily social media time, stratified by self-harm classification. Among respondents classified as having self-harmed, 67% reported spending more than two hours per day on social networking sites or forums, compared with 45% of those who had not — a difference of 22 percentage points. At the upper end of the distribution, 19% of self-harmers reported eight or more hours of daily use, indicating a rightward shift in the usage distribution relative to non-self-harmers.

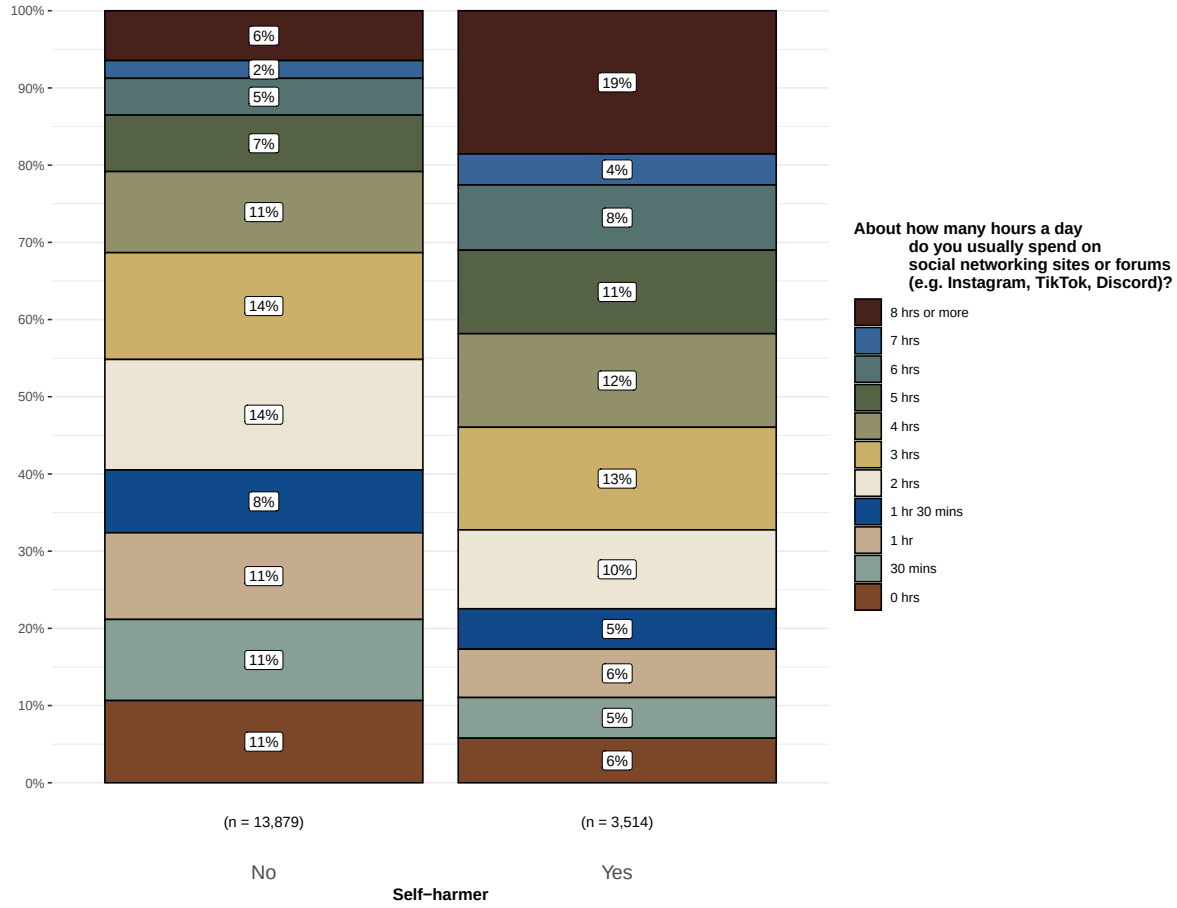

Figure 1: Distribution of daily social media use by self-harm classification

A similar pattern was apparent for account visibility. Figure 2 displays the bivariate association between public account ownership and self-harm classification. 63% of non-self-harmers reported not having a publicly available account, compared with 47% of self-harmers. In other words, self-harmers were more likely to maintain a profile that is visible to anyone without requiring a follow request, a pattern that persisted in the adjusted models presented below.

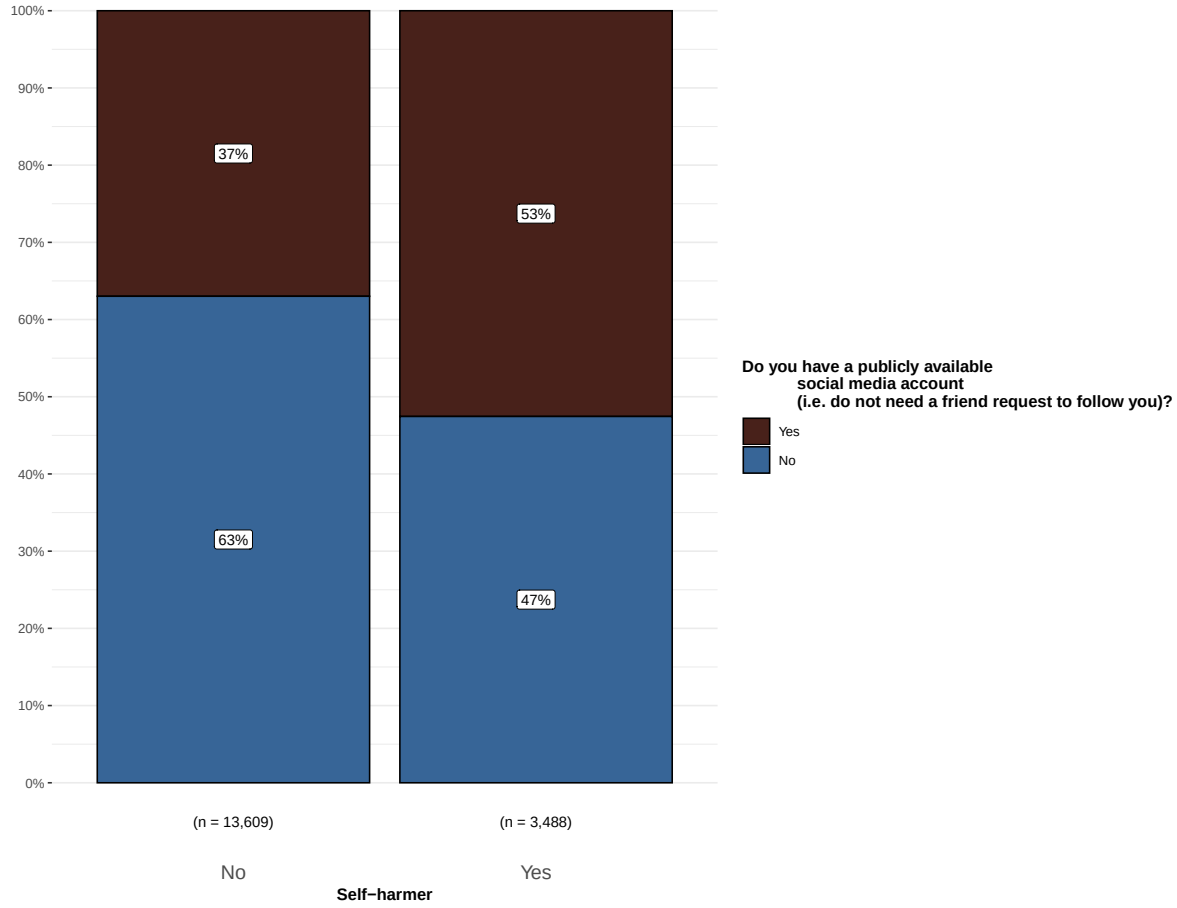

Figure 2: Distribution of public social media account ownership by self-harm classification

We next fitted multilevel logistic regression models to estimate adjusted associations. Table 2 presents adjusted odds ratios from two multilevel logistic regression models, both including random intercepts for schools and adjusted for gender, year group, ethnicity, care status, deprivation, and neurodivergence. In Model (1), daily time spent on social media was entered as a continuous predictor (hours per day). Each additional hour of use was associated with a 19% increase in the odds of self-harm (adjusted OR 1.19, 95% CI 1.17–1.21, $p < .001$). Although the per-hour effect is modest in magnitude, the cumulative difference across the exposure range is substantial, as shown by the predicted probabilities below. In Model (2), maintaining a publicly available social media account was associated with 2.02-fold higher odds of self-harm relative to not having one (adjusted OR 2.02, 95% CI 1.85–2.21, $p < .001$), representing a considerably larger association than daily time alone.

Figure 3 translates these odds ratios into predicted probabilities. The predicted probability of self-harm rose from approximately 19% at three hours of daily use to 36% at eight or more

Table 2: Adjusted odds ratios for daily time on social media and public account ownership

	(1)	(2)
Hours spent on social media (per hour)	1.192***	
	(0.011)	
	(<0.001)	
Has a public social media account		2.024***
		(0.092)
		(<0.001)
Num.Obs.	16 286	15 997

+ p < 0.1, * p < 0.05, ** p < 0.01, *** p < 0.001

Note: Models adjusted for gender, year group, ethnicity, care status, deprivation, and neurodivergence.

hours — an absolute difference of 17 percentage points. For public accounts, the gap was 22% versus 12%, a difference of 10 percentage points that underscores the relevance of account configuration beyond time spent online.

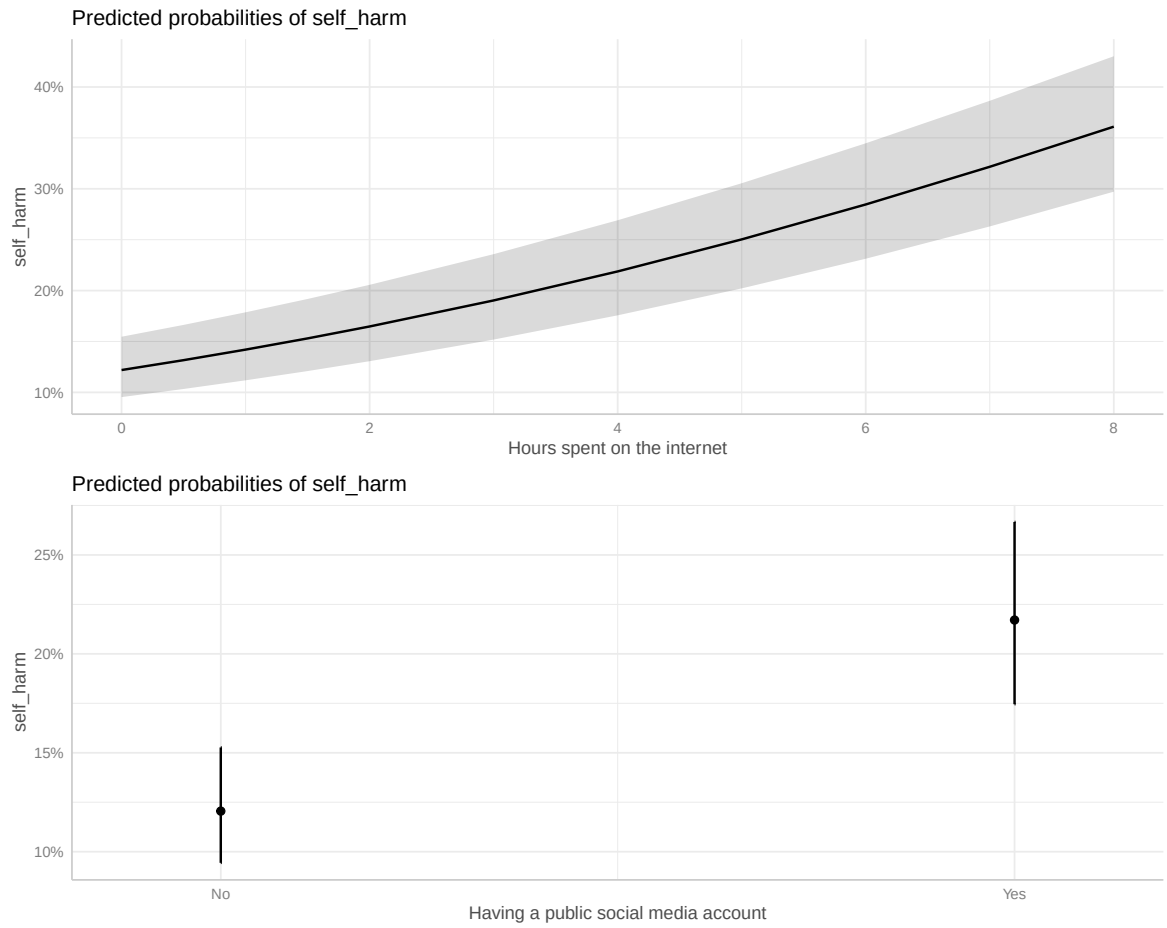

Figure 3: Predicted probabilities of self-harm by time spent on social media and public account ownership

We then examined whether the association with self-harm varied across individual platforms.

Table 3 presents adjusted odds ratios from 16 separate multilevel logistic regressions, each modelling self-harm as a function of use (Yes/No) of a single platform after adjusting for the same set of covariates. The degree of heterogeneity across platforms was considerable. Omegle.X2350fTicked (Yes) — a platform that permits anonymous interactions with strangers — was associated with the highest odds of self-harm (adjusted OR 2.8, $p < .001$). Figure 4 displays the corresponding predicted probabilities, ranging from approximately 35% (X2350f) to 17% (X2350m) — a spread of 18 percentage points that underscores the heterogeneity masked when platform use is treated as interchangeable.

Table 3: Adjusted odds ratios for self-harm by social media platform

	Self-harm (Yes)			Self-harm		
	OR	std.error	pvalue	OR	std.error	pvalue
Snapchat	1.293	0.064	<0.001			
TikTok	1.527	0.078	<0.001			
Instagram	1.395	0.067	<0.001			
Discord	1.771	0.092	<0.001			
Reddit	1.542	0.105	<0.001			
Omegle	2.833	0.252	<0.001			
BeReal	1.615	0.075	<0.001			
Pinterest	1.955	0.092	<0.001			
Telegram	1.893	0.252	<0.001			
Facebook	1.628	0.093	<0.001			
Twitter	1.595	0.084	<0.001			
Signal	1.753	0.295	<0.001			
YouTube				0.971	0.053	0.585
WhatsApp				0.889	0.049	0.034
None of these	1.141	0.209	0.473			
Other	1.107	0.072	0.118			

Note: Each row represents a separate model. All models adjusted for gender, year group, ethnicity, care status, deprivation, and neurodivergence.
Num. Obs.: 20041

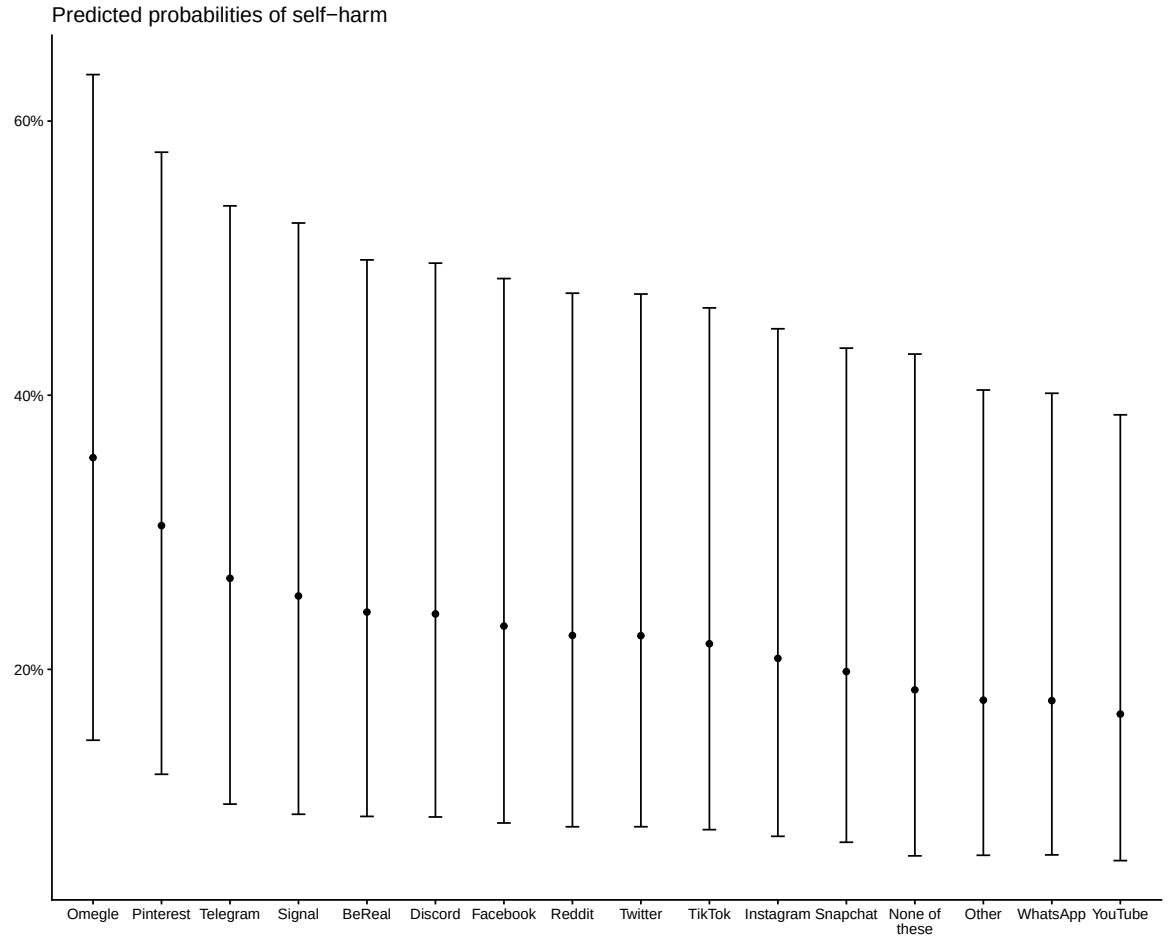

Figure 4: Predicted probabilities of self-harm by social media platform. Error bars represent 95% confidence intervals.

Table 4: Adjusted odds ratios for self-harm by parental mediation strategy

	Self-harm (Yes)			Self-harm		
	OR	std.error	pvalue	OR	std.error	pvalue
Encourage you to explore and learn things on the internet	0.702	0.036	<0.001			
Restrict your internet use	1.356	0.071	<0.001			
Talk to you about what you do online	0.972	0.046	0.548			
Log in to access your online social media accounts (e.g. Facebook, Twitter, Instagram)	1.252	0.094	0.003			
Explain why some websites are appropriate or inappropriate	0.908	0.042	0.040			
Help you when something bothers you on the internet	0.664	0.033	<0.001			
Monitor your internet use				0.883	0.047	0.019
None of these	1.462	0.078	<0.001			
Don't know	0.787	0.048	<0.001			

Note: Each row represents a separate model. All models adjusted for gender, year group, ethnicity, care status, deprivation, and neurodivergence.
Num. Obs. = 20190

Finally, we examined whether parental mediation strategies were associated with self-harm.

Table 4 presents adjusted odds ratios from nine separate multilevel logistic regressions, each modelling the association between a single parental mediation behaviour and self-harm after adjusting for the same covariates used above. The results reveal a clear directional pattern. Restrictive parental behaviour — characterised by limiting or controlling the young person's internet use — was associated with the highest odds of self-harm among all nine behaviours examined (adjusted OR 1.5, $p < 0.01$). By contrast, parental behaviours reflecting active, supportive engagement with children's online experiences were associated with meaningfully lower odds of self-harm (adjusted ORs of 0.66 and 0.7, respectively, both $p < 0.001$). An important caveat applies: the cross-sectional design means that the direction of these associations is ambiguous. Parents may adopt restrictive mediation strategies precisely because their child is already exhibiting signs of distress or self-harm, while supportive engagement may be more feasible when the child is not in crisis. These findings should therefore be interpreted as associations, not as evidence that particular parenting styles cause or prevent self-harm.

Figure 5 converts these odds ratios into predicted probabilities, ranging from approximately 23% for restrictive or disengaged behaviours to 13% for actively supportive ones — an absolute difference of 10 percentage points.

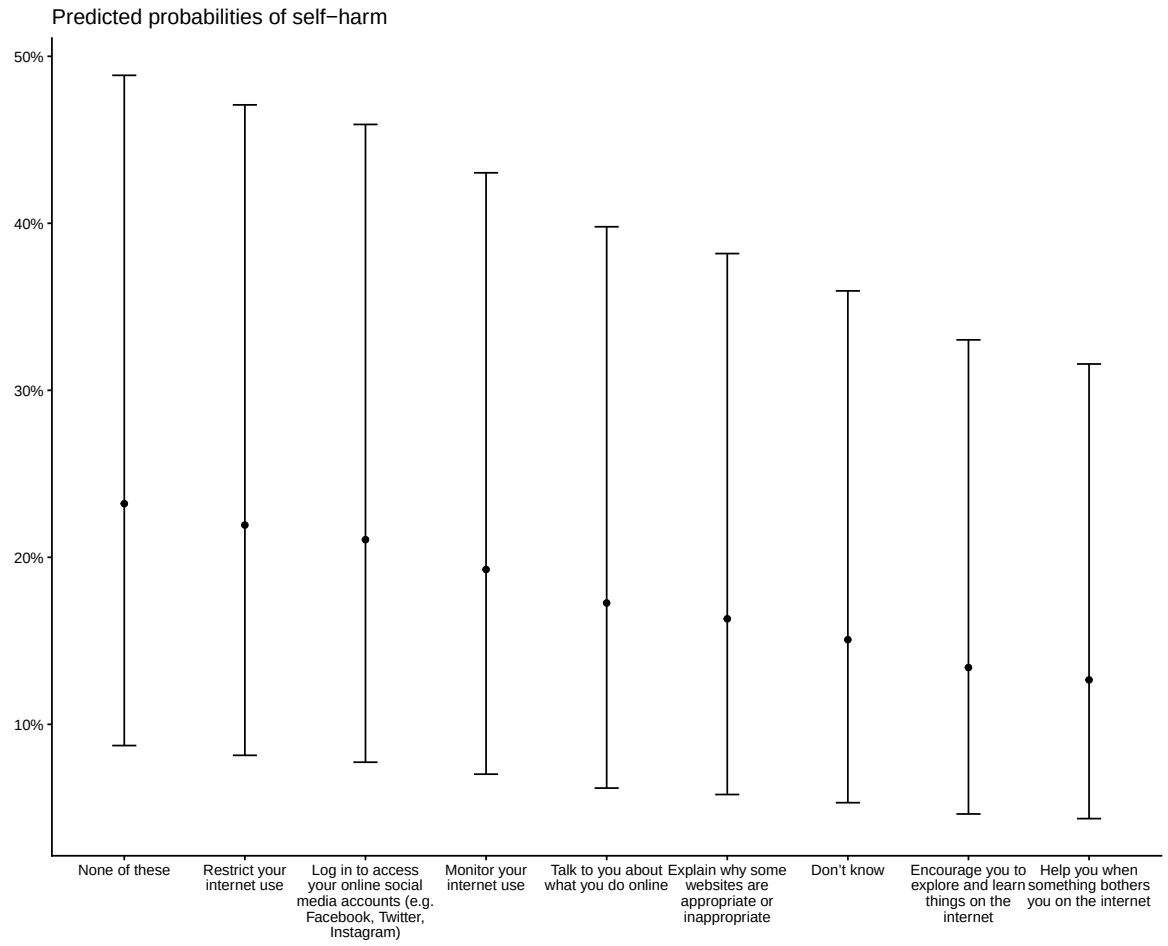

Figure 5: Predicted probabilities of self-harm by parental mediation behaviour. Error bars represent 95% confidence intervals.

We conducted several sensitivity and robustness checks to evaluate these findings.

We first examined the intraclass correlation coefficients (ICCs) to assess the proportion of variance in self-harm attributable to differences between schools. The adjusted ICCs were 0.025 for the daily time model and 0.017 for the public account model. These values confirm that a non-trivial share of the variation in self-harm lies at the school level, justifying the use of multilevel models with random intercepts rather than standard logistic regression, and indicating that school context contributes to differences in self-harm prevalence beyond individual-level characteristics.

To assess the robustness of key findings to unmeasured confounding, we computed E-values (VanderWeele and Ding 2017) — the minimum strength of association that an unmeasured confounder would need to have with both the exposure and the outcome to fully explain away

Table 5: Gender-stratified odds ratios for daily social media time

	Girls	Boys
Hours spent on social media (per hour)	1.210*** [1.183, 1.239]	1.144*** [1.106, 1.182]
Num.Obs.	8362	7322

+ p < 0.1, * p < 0.05, ** p < 0.01, *** p < 0.001

Note: Models adjusted for year group, ethnicity, care status, deprivation, and neurodivergence.

the observed association. For the public account model (aOR 2.02), the E-value is 3.5, meaning an unmeasured confounder would need to be associated with both public account ownership and self-harm by at least this factor to nullify the finding. For Omegle.X2350fTicked (Yes) (aOR 2.8), the E-value is 5. These values suggest that the observed associations are moderately robust to unmeasured confounding, though residual confounding cannot be excluded.

To examine whether the association between social media time and self-harm differed by gender, we fitted separate models for girls and boys, excluding gender from the covariate set. Table 5 presents the results. The direction and magnitude of the association were consistent across genders: each additional hour of daily use was associated with higher odds of self-harm in both girls (OR 1.21) and boys (OR 1.14). Confidence intervals were wider for boys, as expected given the lower prevalence of self-harm in this group. The consistency of the association across genders suggests that the relationship between social media time and self-harm is not driven by, or specific to, one gender group.

Discussion

The central finding of this study is that social media's association with self-harm is not uniform: it depends on which dimension of engagement is examined. Aggregate time spent online, platform choice, account configuration, and parental involvement each yielded distinct patterns, reinforcing the argument that undifferentiated measures of screen time are insufficient for understanding the relationship between digital technology and adolescent mental health (Odgers and Jensen 2020; Valkenburg, Meier, and Beyens 2022). Below, we consider what each dimension contributes to this picture, how the findings relate to the existing literature, and what they imply for policy and practice.

Daily time on social media showed a modest dose–response association with self-harm (adjusted OR 1.19 per additional hour), consistent with previous evidence linking screen time to psychological distress (Cunningham, Hudson, and Harkness 2021). The per-hour effect is small, in keeping with the generally weak effect sizes reported in umbrella reviews of social media and mental health (Valkenburg, Meier, and Beyens 2022), but the cumulative difference is not trivial: predicted probabilities rose from approximately 19% at three hours per day to 36% at eight or more hours. However, time spent online is a blunt summary of what

is, in practice, a heterogeneous set of activities. The more informative findings emerged when we disaggregated engagement into the specific platforms young people use and the conditions under which they use them.

Platform-level analyses revealed considerable heterogeneity in the strength of association with self-harm. Omegle.X2350fTicked (Yes) — a platform characterised by anonymous, public interactions with strangers — was associated with the highest odds (adjusted OR 2.8), whereas platforms oriented toward private, one-to-one communication with known contacts, such as X2350m, showed the weakest associations. The predicted probability of self-harm for users of the platform with the strongest association was approximately 35%, compared with 17% for the lowest — spanning approximately 18 percentage points. This pattern aligns with the concept of platform affordances: the structural features of a platform — whether it encourages anonymity, public visibility, or parasocial engagement with strangers — shape the types of interaction that occur on it, and these in turn appear to carry different levels of risk (Odgers and Jensen 2020). These differences are masked when social media is treated as a single, homogeneous exposure.

The public account finding offers a complementary perspective. Maintaining a publicly visible profile — one that anyone can view or follow without requesting permission — was associated with 2.02-fold higher odds of self-harm, a larger association than daily time alone. This extends recent work by Mabaso et al. (2024), who reported similar associations between public accounts and poorer mental health outcomes using OxWell data but without accounting for school-level clustering. Our multilevel estimates confirm and strengthen this finding. Public accounts may expose adolescents to unsolicited contact, unwanted attention, and harmful content from unknown users — mechanisms that parallel the affordance-based explanation for platform differences described above. Together, the platform and account-visibility findings suggest that it is the degree of openness to strangers, rather than the mere fact of being online, that is most consistently associated with self-harm in these data. As with all exposures examined here, reverse causation remains plausible: adolescents who self-harm may, for example, seek out anonymous platforms or maintain public accounts as part of help-seeking or identity expression.

Parental mediation strategies showed the clearest contrast of any dimension examined. Restrictive approaches — limiting or controlling the young person's internet use — were associated with marginally elevated odds of self-harm, while supportive forms of engagement — helping with distressing content, encouraging independent online exploration — were associated with approximately 3–34% lower odds. This pattern is broadly consistent with the developmental literature on authoritative versus authoritarian parenting and with media-specific research reporting that active mediation outperforms restriction for adolescent well-being. One plausible mechanism is relational: supportive mediation may foster trust and disclosure, making it more likely that a young person will seek help when encountering harmful content or interactions, whereas restriction without dialogue may reduce parental visibility into their child's online life without addressing the underlying risk. However, reverse causation is a particularly acute concern for this dimension: parents may adopt restrictive strategies in direct response to a

child who is already self-harming or in distress, while supportive engagement may be easier to maintain when the child is not in crisis. We therefore emphasise that these associations should not be read as evidence that restrictive parenting causes self-harm, or that supportive mediation prevents it. Longitudinal data are essential to disentangle the direction of effect.

These findings have several tentative implications for policy and clinical practice, subject to the caveats above. First, platform-level heterogeneity argues against blanket restrictions on social media use. Regulatory attention could be more usefully directed at anonymous, stranger-facing platforms — which were associated with the highest odds of self-harm — rather than at social media as a category. Second, current parent-facing guidance from schools and NHS services tends to emphasise screen-time limits, yet our data suggest that the quality of parental engagement may matter more than imposed time boundaries. Guidance could encourage parents to stay informed about their children's online activity, help them navigate distressing content, and create an environment in which young people feel comfortable disclosing negative experiences — rather than defaulting to restriction. Third, in clinical settings, routine psychosocial assessment of adolescents presenting with self-harm could usefully include questions about which platforms they use and whether their accounts are public, alongside standard questions about duration of use. Very high daily use (eight or more hours) may serve as a useful screening flag, but it should not be the sole indicator of concern.

This study has several strengths. The sample is large and school-based ($n > 24{,}000$), providing sufficient statistical power to estimate platform-specific associations. Self-harm was classified using a validated combined approach incorporating both structured self-report items and expert-reviewed free-text responses, which is more sensitive than single-item measures. Multilevel modelling accounted for the non-independence of students within schools. The simultaneous examination of four dimensions of social media engagement within a single dataset allows direct comparison of their relative strength of association. E-values provide transparent assessment of robustness to unmeasured confounding.

Several limitations should be acknowledged. First, and most importantly, we were unable to adjust for pre-existing depression or anxiety. Both conditions are strongly associated with heavy social media use and with self-harm, making them potent confounders. The OxWell survey includes measures of emotional well-being, but these were assessed concurrently with the exposure and outcome and therefore could not serve as baseline covariates without introducing collider bias. The E-values we report are moderate, meaning that a confounder of the magnitude of depression could plausibly attenuate or explain the observed associations; until longitudinal studies with baseline mental health measures are available, the present estimates should be regarded as upper bounds of the true effect. Second, the cross-sectional design precludes causal inference more broadly: we cannot determine whether particular patterns of social media use precede, follow, or co-occur with self-harm. This concern is especially acute for parental mediation, where restrictive approaches may be a response to, rather than a cause of, a child's distress. Third, both exposure and outcome rely on self-report, introducing the possibility of recall bias and social desirability effects.

Several additional methodological caveats apply. Fitting separate models for 16 platforms and nine parental behaviours without formal correction for multiple comparisons increases the risk of false-positive findings; individual platform-level results should be interpreted in the context of the overall pattern rather than in isolation. The composite neurodivergence variable conflates conditions (dyslexia, dyspraxia, autism, ADHD) with distinct aetiology and risk profiles; future work with larger samples should examine these separately. Some rarely used platforms may have been underpowered to detect meaningful associations. The sample covers four English regions and may not generalise to other national or cultural contexts. Finally, the social media landscape evolves rapidly; platform-specific findings may have limited durability as new platforms emerge and existing ones change their features.

In conclusion, different dimensions of social media engagement carry distinct associations with self-harm among adolescents. Platform type, account visibility, and parental involvement each provide information that time spent online alone does not capture. These findings argue against treating social media as a uniform exposure and support the development of platform-aware, affordance-informed prevention strategies. The association between supportive parental engagement and lower odds of self-harm — and the absence of a similar pattern for restrictive approaches — is a finding that warrants testing in longitudinal studies and school-based intervention programmes before informing clinical guidance.

References

Cunningham, Simone, Chloe C. Hudson, and Kate Harkness. 2021. "Social Media and Depression Symptoms: A Meta-Analysis." *Research on Child and Adolescent Psychopathology* 49 (2): 241–53. https://doi.org/10.1007/s10802-020-00715-7.

"Developing Admin-Based Ethnicity Statistics for England and Wales." 2023. https://www.ons.gov.uk/peoplepopulationandcommunity/culturalidentity/ethnicity/articles/developingadminbasedethnicitystatisticsforenglandandwales/2020#ethnic-breakdown.

Geulayov, Galit, Rohan Borschmann, Karen L. Mansfield, Keith Hawton, Paul Moran, and Mina Fazel. 2022. "Utilization and Acceptability of Formal and Informal Support for Adolescents Following Self-Harm Before and During the First COVID-19 Lockdown: Results from a Large-Scale English Schools Survey." *Frontiers in Psychiatry* 13 (June). https://doi.org/10.3389/fpsyt.2022.881248.

Giumetti, Gary W., and Robin M. Kowalski. 2022. "Cyberbullying via Social Media and Well-Being." *Current Opinion in Psychology* 45 (June): 101314. https://doi.org/10.1016/j.copsyc.2022.101314.

Griffin, Eve, Elaine McMahon, Fiona McNicholas, Paul Corcoran, Ivan J. Perry, and Ella Arensman. 2018. "Increasing Rates of Self-Harm Among Children, Adolescents and Young Adults: A 10-Year National Registry Study 2007–2016." *Social Psychiatry and Psychiatric Epidemiology* 53 (7): 663–71. https://doi.org/10.1007/s00127-018-1522-1.

Hamm, Michele P., Amanda S. Newton, Annabritt Chisholm, Jocelyn Shulhan, Andrea Milne, Purnima Sundar, Heather Ennis, Shannon D. Scott, and Lisa Hartling. 2015. "Prevalence

and Effect of Cyberbullying on Children and Young People: A Scoping Review of Social Media Studies." *JAMA Pediatrics* 169 (8): 770–77. https://doi.org/10.1001/jamapediatrics.2015.0944.

Mabaso, Wakithi Siza, Sascha Hein, Gabriela Pavarini, The OxWell Study Team, and Mina Fazel. 2024. "Exploring the Relationship Between Public Social Media Accounts, Adolescent Mental Health, and Parental Guidance in England: Large Cross-Sectional School Survey Study." *Journal of Medical Internet Research* 26 (December): e57154. https://doi.org/10.2196/57154.

Mansfield, Karen Laura, Stephen Puntis, Emma Soneson, Andrea Cipriani, Galit Geulayov, and Mina Fazel. 2021. "Study Protocol: The OxWell School Survey Investigating Social, Emotional and Behavioural Factors Associated with Mental Health and Well-Being." *BMJ Open* 11 (12): e052717. https://doi.org/10.1136/bmjopen-2021-052717.

Morgan, Catharine, Roger T. Webb, Matthew J. Carr, Evangelos Kontopantelis, Jonathan Green, Carolyn A. Chew-Graham, Nav Kapur, and Darren M. Ashcroft. 2017. "Incidence, Clinical Management, and Mortality Risk Following Self Harm Among Children and Adolescents: Cohort Study in Primary Care." *BMJ* 359 (October): j4351. https://doi.org/10.1136/bmj.j4351.

O'Day, Emily B., and Richard G. Heimberg. 2021. "Social Media Use, Social Anxiety, and Loneliness: A Systematic Review." *Computers in Human Behavior Reports* 3 (January): 100070. https://doi.org/10.1016/j.chbr.2021.100070.

Odgers, Candice L., and Michaeline R. Jensen. 2020. "Annual Research Review: Adolescent Mental Health in the Digital Age: Facts, Fears, and Future Directions." *Journal of Child Psychology and Psychiatry* 61 (3): 336–48. https://doi.org/10.1111/jcpp.13190.

Sempé, Lucas, Mina Fazel, Emma Soneson, and Simona Skripkauskaite. 2024. "Exposure to Self-Harm Content Online and Self-Harm Behaviour in Secondary School Pupils," April. https://doi.org/10.17605/OSF.IO/5UXMJ.

Skripkauskaite, Simona, and Mina Fazel. 2022. "Time Spent Gaming, Device Type, Addiction Scores, and Well-Being of Adolescent English Gamers in the 2021 OxWell Survey: Latent Profile Analysis." *JMIR Pediatrics and Parenting* 5 (4): e41480. https://doi.org/10.2196/41480.

Soneson, Emma, Sarah O'Leary, and Mina Fazel. 2024. "Trans and Gender Diverse Identities in Adolescent Health Research: Making the Most of Imperfect Data." *BMJ Mental Health* 27 (1): e301150. https://doi.org/10.1136/bmjment-2024-301150.

Valkenburg, Patti M., Adrian Meier, and Ine Beyens. 2022. "Social Media Use and Its Impact on Adolescent Mental Health: An Umbrella Review of the Evidence." *Current Opinion in Psychology* 44 (April): 58–68. https://doi.org/10.1016/j.copsyc.2021.08.017.

VanderWeele, Tyler J., and Peng Ding. 2017. "Sensitivity Analysis in Observational Research: Introducing the E-Value." *Annals of Internal Medicine* 167 (4): 268–74. https://doi.org/10.7326/M16-2607.

Appendix

Table 6 provides descriptive statistics for the full analytic sample, including the individual items from which the deprivation index was derived. The sample was broadly distributed across year groups, with the largest groups in Years 7 and 8. The gender composition was approximately 50% girls, 47% boys, 1.1% gender diverse, and 2.4% gender non-disclosing. The ethnic breakdown was predominantly White (0%), followed by Asian/Asian British (0%), Mixed/Multiple Ethnic Groups (0%), Black/Black British (0%), and Other (0%).

Regarding care status, 3.2% of respondents reported being currently in care, looked after, or fostered. One in five respondents considered themselves neurodivergent (0%). In terms of social media use, 49.6% of respondents spent three or more hours per day on social networking sites, and 40% maintained publicly available social media accounts.

Table 6: Descriptive statistics for the full analytic sample

Characteristic	**N = 24,909**
Self-harm classification	4,479 / 24,909 (18%)
Year group	
Y05	2,412 / 24,909 (9.7%)
Y06	2,566 / 24,909 (10%)
Y07	3,933 / 24,909 (16%)
Y08	3,878 / 24,909 (16%)
Y09	3,831 / 24,909 (15%)
Y10	2,847 / 24,909 (11%)
Y11	2,727 / 24,909 (11%)
Y12	1,641 / 24,909 (6.6%)
Y13	1,074 / 24,909 (4.3%)
Gender	
Boy	11,499 / 24,706 (47%)
Girl	12,349 / 24,706 (50%)
Gender Diverse (GD)	267 / 24,706 (1.1%)
Gender Non-Disclosing (GND)	591 / 24,706 (2.4%)
Missing	203
Ethnic background	
Asian/Asian British (aggregated)	3,607 / 24,909 (14%)
Black/Black British/African/Caribbean (aggregated)	1,152 / 24,909 (4.6%)
Mixed/Multiple Ethnic Groups (aggregated)	1,313 / 24,909 (5.3%)
Other ethnic group	978 / 24,909 (3.9%)
White (aggregated)	13,561 / 24,909 (54%)
Skipped by respondent	4,298 / 24,909 (17%)
Child in care, looked after, or fostered?	
Yes	805 / 24,909 (3.2%)
No	17,508 / 24,909 (70%)
I don't know what this means	1,215 / 24,909 (4.9%)
Not now, but I used to be in care	206 / 24,909 (0.8%)

(continued)

Characteristic	N = 24,909
Prefer not to say	167 / 24,909 (0.7%)
Skipped by respondent	5,008 / 24,909 (20%)
Neurodivergent	
Yes	3,630 / 24,909 (15%)
No	12,291 / 24,909 (49%)
Not sure	3,692 / 24,909 (15%)
Prefer not to say	270 / 24,909 (1.1%)
Skipped by respondent	5,026 / 24,909 (20%)
Deprivation index	5.49 (3.40)
Missing	1,114
Worry about not having enough money for family needs	
Never or hardly ever	17,054 / 24,909 (68%)
Often	1,654 / 24,909 (6.6%)
Some of the time	5,859 / 24,909 (24%)
Skipped by respondent	342 / 24,909 (1.4%)
Family uses food banks	
Never or hardly ever	22,878 / 24,909 (92%)
Often	310 / 24,909 (1.2%)
Some of the time	1,151 / 24,909 (4.6%)
Skipped by respondent	570 / 24,909 (2.3%)
House is cold and/or damp	
Never or hardly ever	22,162 / 24,909 (89%)
Often	462 / 24,909 (1.9%)
Some of the time	1,876 / 24,909 (7.5%)
Skipped by respondent	409 / 24,909 (1.6%)
Unable to afford uniform, equipment, or school trips	
Never or hardly ever	22,527 / 24,909 (90%)
Often	620 / 24,909 (2.5%)
Some of the time	1,412 / 24,909 (5.7%)
Skipped by respondent	350 / 24,909 (1.4%)
Unable to afford to eat at school	
Never or hardly ever	22,845 / 24,909 (92%)
Often	532 / 24,909 (2.1%)
Some of the time	1,153 / 24,909 (4.6%)
Skipped by respondent	379 / 24,909 (1.5%)
Insufficient space at home for homework or relaxation	
Never or hardly ever	21,231 / 24,909 (85%)
Often	965 / 24,909 (3.9%)
Some of the time	2,330 / 24,909 (9.4%)
Skipped by respondent	383 / 24,909 (1.5%)
No or poor internet access at home	
Never or hardly ever	22,563 / 24,909 (91%)
Often	357 / 24,909 (1.4%)
Some of the time	1,598 / 24,909 (6.4%)
Skipped by respondent	391 / 24,909 (1.6%)
Goes to bed hungry due to food insecurity	

(continued)

Characteristic	N = 24,909
Never or hardly ever	23,443 / 24,909 (94%)
Often	247 / 24,909 (1.0%)
Some of the time	804 / 24,909 (3.2%)
Skipped by respondent	415 / 24,909 (1.7%)
Free school meals (% in school)	20 (16)
Missing	1,114
Hours per day on social networking sites or forums	
0 hrs	1,681 / 23,237 (7.2%)
30 mins	1,643 / 23,237 (7.1%)
1 hr	1,778 / 23,237 (7.7%)
1 hr 30 mins	1,313 / 23,237 (5.7%)
2 hrs	2,346 / 23,237 (10%)
3 hrs	2,387 / 23,237 (10%)
4 hrs	1,884 / 23,237 (8.1%)
5 hrs	1,396 / 23,237 (6.0%)
6 hrs	957 / 23,237 (4.1%)
7 hrs	460 / 23,237 (2.0%)
8 hrs or more	1,548 / 23,237 (6.7%)
Skipped by respondent	5,844 / 23,237 (25%)
Missing	1,672
Publicly available social media account	
Yes	6,866 / 23,018 (30%)
No	10,231 / 23,018 (44%)
Skipped by respondent	5,921 / 23,018 (26%)
Missing	1,891

[1] n / N (%); Mean (SD)

Role of the funding source

The funder had no role in study design, data collection, data analysis, data interpretation, or writing of the report. The corresponding author had full access to all data and had final responsibility for the decision to submit for publication.

Data sharing

The OxWell Student Survey data are not publicly available due to the sensitive nature of the content and the age of participants. Access to anonymised data can be requested from the OxWell Study Team subject to a data sharing agreement and ethical approval.