Draft: Exploring Associations Between Web Activities and Depression

Emilia Marchese[1,2,*], **Juhi Kulshrestha**[1,2,*], **and Talayeh Aldavood**[1,2,*]

[1]Aalto University, Departement of Computer Science, Espoo, 02150, Finland
[*]emilia.marchese@aalto.fi, juhi.kulshrestha@aalto.fi, talayeh.aledavood@aalto.fi

ABSTRACT

Depression is one of the leading causes of illness and disability worldwide. Over the past few decades, there has been a surge in our reliance on the internet, advancing the prospect of utilizing our online behavior as a diagnostic tool for identifying depression. At the same time, it raises the question of potential associations between internet use and mental health. Previous research on internet usage for mental health assessment pertains mainly to self-reported measures of internet usage for specific demographic groups. This work explores associations between internet usage behaviour and depression using desktop internet browsing traces from a large hetero heterogeneous population of about 900 individuals. Panel, age group and gender analyses are conducted to observe population and sub-population level associations. It is found that there are marked differences between the associations found for each demographic group. Among the found associations with high statistical significance, the duration spent on message boards and forums showed a positive association with depression in the panel analysis (95% CI (0.008, 0.0026), $P < 0.001$), in the male group analysis (95% CI (0.008,0.026), $P < 0.001$), and in the 50-59 age group analysis (95% CI (0.010, 0.029), $P < 0.001$), while the duration spent on chat and messaging platforms showed a negative association with depression for the female group analysis (95% CI (-0.013,-0.003), $P < 0.001$). These findings underscore that demographic factors shape the relationship between online activity and depression. The found associations emphasize the need for further research into the content consumed on these platforms and how it relates to mental health.

Introduction

Depression is a leading cause of disability worldwide[1], affecting about 5% of individuals globally[2]. Following the COVID-19 pandemic, there has been a global rise in depression incidence of more than 25%[3]. The World Health Organization predicts that depression incidence will only grow in the next decade, making depression the leading cause of illness by 2030[1]. Despite its high prevalence, it is estimated that 50% of the people suffering from depression are not recognized or adequately treated[4]. Our growing reliance on the internet makes it reasonable to consider internet usage patterns as a possible detection tool for mental health conditions. Additionally, the evolving nature of our online interactions raises the question of potential associations between internet use and mental health. Identifying internet browsing behaviors associated with depression can help in developing early detection tools for depression using web browsing information, enabling mental health professionals to provide timely support and potentially mitigate the severity of the condition.

Previous studies have explored associations between internet usage and depression using self-reported measures of internet use. In 2016, Hökby S. et al.[5] aimed to assess whether mental effects of internet use were attributable to the content of the internet use or to the perceived consequences of internet use, such as sleep loss and socialization. They recruited 2286 European adolescents and asked them to answer two surveys 4 months apart. The survey included a depression, stress and anxiety assessment (DASS-42) and internet usage questions for 7 different activities (socializing, gaming, school/work, gambling, newsreading/watching, pornography, targeted searches) and perceived consequences of the activities such as finding friends, sleep loss, learning and others. In their cross-sectional analysis, they found that both the time spent on the internet and on various internet activities were statistically significant predictors, but that the perceived consequences of engaging in those activities were more important predictors. Only gaming, gambling and targeted searches had mental health effects that were not fully accounted for by perceived consequences.

A 3-wave study[6] of duration of three years involving 27507 people in England aged 50 or older aimed to explore the relationship between internet use and mental health in older adults. The adults were asked to answer questions about their socioeconomic status and depression (CES-D). Internet use was assessed with a questionnaire on the time spent on communication, entertainment, information access, finances, and ecommerce. Data was collected through computer-assisted personal interviews, self-completion questionnaires and nurse assessments. In their longitudinal analysis, they noted that using the internet for communication purpose, specifically email use, was associated with better mental health and that using the internet for information access, specifically job searching, was associated with worse mental health.

Instead of relying on self-assessed measurements of internet use, Katikalapudi R. et al.[7] made one of the first attempts to monitor internet usage data and relate it to mental health. They used data from the Missouri S&T campus CiscoNetflow network to explore associations between depression and internet usage for college students. They monitored the internet package flows of 216 college students for the duration of 45 days, and asked them to respond to a one-time depression assessment survey (CES-D). They used distribution difference tests to show that students with depressive symptoms had higher average packets per flow, higher remote file octects, higher email usage and higher entropy in the flow duration. They speculate that the higher average packets per flow might be an indication of streaming and gaming, and the the higher entropy of duration an indication of frequent switching between tasks.

Several other studies have delved into the associations between various internet activities and mental health, yielding conflicting findings. A non-comprehensive summary of these associations is provided in Table 1.

Table 1. Internet activities and depression: a non-comprehensive list of internet activities that have been associated with depression in the existing literature.

Feature	Source
Email	Email usage as a form of communication was negatively associated with depression in the older population[6] and in studies with college students[8].
Social-networking	People with mild to severe depression had higher time spent on social networking apps[9]
Communication and instant messaging use	In the older population, using the internet for communication purposes was protective of depressive symptoms[6]. Using the internet to communicate with friends and family was associated with declines in depression[10]. Communication with ones social circle only has been sometimes negatively associated with depression[11] while other studies have found instant messaging to be positively related with depressive symptoms[12]
Games	For adolescents, gaming was a significant predictor of mental health[5 13]
Gambling	For adolescents, gambling was a significant predictor of mental health[5]
Targeted searches	For adolescents and the older population, targeted searches were a significant predictor of mental health[5 6]
Shopping	Shopping disorder and online shopping have been associated with depression[14]
Job related	For the older population, the frequency of job related targeted searches were significant predictors of mental health[14]
Vaguebooking on boards and socials	Vaguebooking, the practice of making a vague post on social media, was predictive of suicidal ideation in adolescents, which is a depression symptom[15]
Streaming	Higher packets per flow were indicative of streaming and positively correlated with depression for college students[13]
Adult content	High female pornography use has been associated with more depressive symptoms for US college students[16]. For male adults, the relationship between watching adult content hinges on the congruence between one's moral beliefs about pornography and thier viewing practices. Depressed men likely view higher levels of pornography as a coping aid, especially when they do not view it as immoral[17]. Another study found that viewing sexual material one does not deem as pornographic was related to higher levels of depressive symptoms[18].
Health	Depressed individuals may experience excessive worry regarding health[19]. Moreover, previous studies have shown that there are gender differences in health information searches[20 21 22], with women being more likely to search for health information, and that health related internet use is associated positively with depression[10 22].

The existing literature on depression and internet use pertains mostly to specific socio-demographic groups, such as adolescents[5] and college students[7] or the elderly population[6], and includes few repeated measurements often based on self-reported measures of internet use. Studies that rely on self-reported measurements of internet use are prone false reporting and to recall-bias[23]. The results from these studies for similar internet activities differ significantly depending on the analysed population. This study aims to contribute to the research on internet use and depression by observing repeated measurements of internet use from desktop devices and depression scores from a large population of 900 German individuals aged 18 or older. The measurements of internet use are collected using raw web browsing traces from desktop devices, offering a more objective alternative over self-reported measures. Moreover, the study is conducted for four months, with depression scores assessed monthly. Recognizing that there are differences between demographic groups and their level of interaction with the internet, the analysis is performed for the entire population, by gender and by age-group separately. The internet activities that have shown existing associations in previous studies (see Table 1) are included in the study. Additionally, the total duration spent browsing in the morning, afternoon, evening and night respectively is included to observe whether the internet use at specific times of the day might be associated with depression.

The depression severity score is assessed monthly using the PHQ-9[24] scale, from July 2024 to October 2024. The PHQ-9 scale aims to assess the severity of depression symptoms in the preceding two weeks. For each wave, the duration spent for the time of day (Morning, Afternoon, Evening, Night), and the duration spent on each internet activity (shopping, chat and messaging, social networking, health, search engines and portals, games, gambling, email, streaming, message boards and forums, job-related, adult, education) is calculated from the web browsing traces of the two-weeks period preceding the survey. Hierarchical mixed effects models with panelist random effects are used to observe statistically significant associations between the internet usage features and the depression scores. For each analysis, age and season (summer, autumn) fixed effects are included as controls. For the panel and age group analysis, the gender fixed effect is included as control.

Results

Partecipants

A final dataset of 2840 entries of internet use features from 953 participants is used. The count of participant by demographic group is reported in Table 3. Not all participants have observations for all the waves included in this study, due to participants dropping out of the study or not passing the selection criteria (see "Methods" section). Figure 1 shows the distribution of PHQ-9 depression scores over the four waves of the study.

Table 2. Selected panelists included in the analyses: for each wave, the panelists are included in the analyses if they pass the selection criteria. A panelist may have missing observations in one or more wave due to dropping out from the study or not passing the selection criteria.

Wave	18–29	30–39	40–49	50–59	60+	Male	Female	Panel
1: July	68	148	218	250	212	463	433	896
2: August	54	125	194	230	195	427	371	798
3: September	43	85	153	162	146	316	273	589
4: October	37	83	137	155	145	303	254	557
Total observations	202	441	702	797	698	1509	1331	2840

Table 3. Panelists by demographic group: counts of selected panelists by demographic group. Panelists may have missing observations for one or more waves due to dropping-out from the study or not passing the selection criteria.

Age Group	Male	Female	Panel
18-29	22	50	72
30-39	77	86	163
40-49	138	92	230
50-59	143	122	265
60+	115	108	223
Total	495	458	953

Panel

The panel level results are shown in Table 7. Age, gender, and season fixed effects are included in the model to account for these effects on the PHQ-9 scores. The analysis reveals that age, gender, season, the duration spent on message boards and forums and the duration spent on adult content are statistically significantly associated with depression severity. Specifically, identifying as a female (95% CI (0.624,1.920), $P < 0.001$) and the time spent of message boards and forums (95% CI (0.008, 0.0026), $P < 0.001$) showed a positive association with PHQ-9 score. Age (95% CI ($0.116, -0.063$), $P < 0.001$), season (95%

Table 5. Study adherence by baseline depression severity: for each PHQ-9 depression threshold, the count of panelist at baseline (Wave 1) is reported. The adherance columns show the percentage of those panelists having observations in the consequent waves.

Thresholds	Partecipants Wave 1	Adherance Wave 2	Adherance Wave 3	Adherance Wave 4
Normal: PHQ-9 < 5	422	86%	67%	62%
Mild : $5 \leq$PHQ-9 < 10	270	80%	60%	55%
Moderate: $10 \leq$ PHQ-9 < 15	124	82%	60%	54%
Moderately Severe: $15 \leq$ PHQ-9 < 20	61	83%	49%	54%
Severe: $20 \leq$ PHQ-9 ≤ 27	19	84%	52%	52%

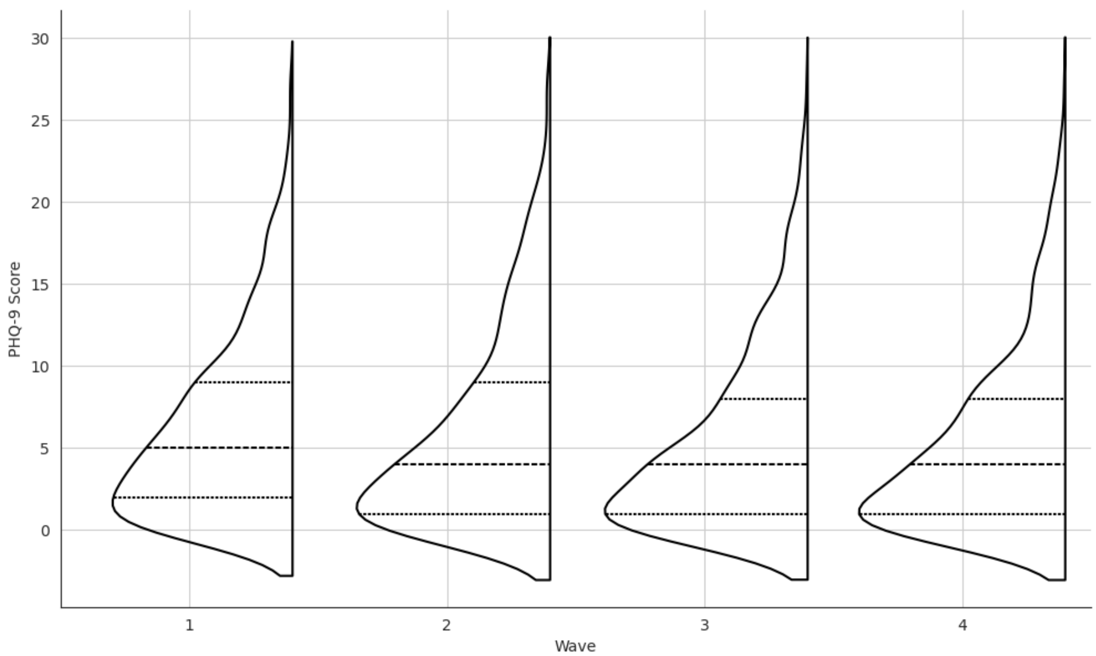

Figure 1. PHQ-9 score distribution across waves. For each wave, the score distribution is from all the panelists included in the analysis for that specific wave, according to the panelist selection criteria.

CI ($-0.592,-0.182$), P < 0.001) and the duration spent on adult content (95% CI (-0.002,0.000), P < 0.05) showed statistically significant negative associations with PHQ-9 scores.

Gender

The results from the gender level models are shown in Table 8. The female model includes observations from the participants identifying as female, and the male model includes observations from the participants identifying as male. The season and age fixed effects are included in the model. Age and season show negative statistically significant associations with PHQ-9 scores for both genders. For females, the time spent on chat and messaging websites has a negative statistically significant association with PHQ-9 scores (95% CI ($-0.013,-0.003$), P < 0.001). For males, the time spent on message boards and forums (95% CI (0.008,-0.026), P < 0.001) and the time spent on job-related websites (95% CI (0.022,-0.170), P < 0.05) show statistically significant positive associations with depression severity scores. On the other hand, the time spent on adult content (95% CI (-0.002,0.000), P < 0.05) shows a small statistically significant negative association with PHQ-9 scores.

Age Group

The age group results are shown in Table 11 for age groups 18–29, 30–39 and 40–49, and in Table 12 for age groups 50–59 and 60+. The analysis for age group 18–29 reveals negative statistically significant associations between email (95% CI (− 0.030, − 0.001), P < 0.05) and message boards and forums (95% CI (− 0.203, − 0.0.026), P < 0.05) duration and PHQ-9 severity scores. No statistically significant associations are found between the internet usage fixed effects and depression scores for age groups 30–39 and 40–49. The analysis for the age group 50–59 reveals a positive statistically significant association with the time spent on message boards and forums (95% CI (0.010, 0.029), P < 0.001) and shopping websites (95% CI (0.000, 0.005), P < 0.05), and a negative statistically significant association with the time spent on chat and messaging platforms (95% CI (− 0.008, − 0.000), P < 0.05). Lastly, the analysis for the 60+ age groups shows a positive statistically significant association with the time spent on the internet in the afternoon (95% CI (0.000, 0.002), P < 0.05).

Discussion

Several internet usage effects have been found to be associated with PHQ-9 depression severity scores. These effects differ by gender and age group. The age fixed effect showed a statistically significant negative association in the panel level analysis, gender analysis, and within the analysis of specific age-groups (18–29, 30–39, 60+). This is consistent with existing psychological findings, which shown that depressive symptoms become less severe with age[19,25]. The statistical significance of gender in the studied population and for specific age groups (40–49, 50–59) is also consistent with previous psychological findings showing that females are more likely to experience depressive symptoms[19,25]. The season control variable was statistically significant in the population model, in the gender models, and for some age groups (40–49, 60 +), showing that the autumn season (wave 3 and 4), compared to the summer season (wave 1 and 2), has a negative association with depression. Existing studies support depressive symptoms worsening during fall and winter: specifically, seasonal depressive symptoms, or seasonal affective disorders (SAD), are more likely to occur in fall and winter[19,26]. As shown by Figure 1, the distributions of PHQ-9 scores do not change drastically between the first two waves (summer season) and the last two waves (autumn season), but there is a small drop in average scores. Table 5 reveals that much higher percentages of people scoring higher levels of depression (Moderately Severe, Severe) on the first wave drop-out from the study after the second wave, compared to people scoring lower levels of depressions (Normal, Mild, Moderate) at baseline. The statistical significance of the seasonal effect could therefore be the result of this differential dropout, instead of the result of seasonal changes.

The duration spent on message boards and forums is a positive predictor in the panel model, in the male model, and in the 50-59 age group model, while it is a negative predictor in the 18–29 age group model. More analysis is needed to identify how this feature relates to depression, specifically by observing the content viewed in the message boards and forums. Previous findings have shown that vague booking on message boards and socials is predictive of suicide ideation in adolescents[15], which is a serious depressive symptom. Online forums and message boards can sometimes be places where negative or distressing content is shared. Exposure to such content, including discussions about personal problems, health issues, or negative world events, can contribute to feelings of depression. Cyberbulling and harassment in some online communities might also be the cause of exacerbated depression. On the other hand, people who are already experiencing depression might turn to online forums and message boards as a way to seek help or find a community of others who understand their struggles. The statistical significance of this internet activity in several of the models highlights the need of more in-depth research. The time spent on chat and messaging was a negative predictor of depression in the female model and in the 50-59 age group model. The significance of this parameter is consistent with previous research showing that using the internet for communication purposes was associated with declines in depression[11,10]. It could be speculated that women might experience health benefits from online communication due to their tendency to seek emotional support and express feelings more openly, which can provide therapeutic relief and strengthen social bonds. The statistical significance of this internet activity for the 50-59 age group is lower, and the confidence intervals are large. Regardless, it could be speculated that people in their 50s might be facing unique stressors, such as career pressures and health concerns, or changes in family dynamics including what's been labelled as the "empty nest syndrome", the feeling of grief that parents experience when their children move out from their home. Empty nesters who are digital technology users tend to show significantly lower loneliness, lower depression symptom severity and higher perceived social support compared to non-digital technology users[27]. It's possible that online communication may help prevent and relieve some of the mental burden of these changes. More research on the scope of these online interactions is needed to further explore its association with depression severity.

The negative association with email use for people age 18-29 is consistent from previous findings from college students[8] and in the older population[6], although the confidence interval of this estimate are large and the statistical significance low. Similar arguments could be made for the positive associations with shopping duration (50-59) and afternoon internet usage duration (60+) in the age group models, and the negative association with job-related content in the male model. Adult content demonstrated a negative association in the panel and male model, also with wide confidence intervals and lower statistical significance. More research is needed to make further inferences about these internet activities and depression. For all models, the Intra Class Correlations (ICC) are high, and the marginal R^2s are low, indicating that the PHQ-9 scores are heavily dependent on the individual panelist effect, and that the fixed effects included are responsible for a small proportion of the total variance in the dependent variable. The substantial contribution of the panelist random effect highlights the significance of individual differences not captured by the fixed effects in explaining the observed variations in the depression scores.

The study has several limitations. The number of panelists by age group, specifically 18–29 and 30–39, is small. Results from the age-group analysis might be less robust. The study also faced a concerning drop of panelists, with panelists having higher depression scores at baseline more likely to drop from the study than panelists scoring lower depression severity at baseline. This results in less observations for panelists having higher depression levels at the start of this study, potentially undermining the found associations for these individuals. This differential dropout is a form of selection bias, which may cause underestimation of the effects and increase variability in the estimates, leading to large confidence intervals as observed for many of the found statistically significant associations. To contrast this, an analysis that takes into account the dropout rates should be conducted, with more robust inference methods. Lastly, this study limits itself to internet usage on desktop devices, showing an incomplete picture of the level of interaction with the internet. Future study should integrate measurements from mobile devices.

Despite these limitations, this study reveals interesting associations. The associations found for the use of message boards and forums and online chat and messaging platforms calls for more attention being paid to these internet activities in relation to depression. The presence of different associations in separate demographic groups shows that there might be differences between how the interaction with these activities relates to the well-being of different demographic segments. Future studies should explore the intention and content behind these internet activities, their relation to depression, and why the effects may differ across demographic groups. This study contributes to the growing body of literature on the relationship between internet use and well-being. Using objective measures of internet use from desktop devices, it provides an alternative to self-reported measures of internet use, offering valuable insights into the distinct ways online activities on desktop devices associate with mental health across various demographic groups.

Methods

Procedure

The study is conducted on an panel of 1066 subjects from Germany. A third party company runs a population-representative panel who can be invited to participate in different studies. The company collects the data through a web-browser plug-in and prompts the panelists to answer the monthly survey via a website or mobile app. The panel distribution follows that of the general German demographic. The desktop web browsing traces are collected continuously for a period of four months (06/2023 - 10/2023), and the panelists are asked to respond to the PHQ-9 questionnaire on a monthly basis. The PHQ-9 is used to assess the depression severity of the respondents in their PHQ-9 interval, which is the two weeks period prior taking the survey. For each survey wave, the panelists who responded to the questionnaire are selected for the analysis if they have at least an observation in their PHQ-9 interval browsing traces and one observation before or in the first day of their PHQ-9 interval. These selection criteria are applied to avoid participants who used the device only to answer the surveys. The number of panelists included in each wave following these criteria is reported in Table 2.

The browsing traces are URL views, including the start time of the observation, the duration spent observing the URL, and a category specific to the domain of the URL. A 3-minutes inactivity time-out is applied to the duration of the URL view in the lack of any mouse movement. The domain categories are collected from the Webschrinker API[28]. The custom category *email* is added using a list of known email sub-domains. Consequent equal URLs from the same entertainment or streaming-media domain are merged together to account for the 3-minutes inactivity time-out. Since domains might have more than one category, the URL view duration is split into equal fragments across its categories. For each PHQ-9 interval, the total duration for the category is the sum of the duration of its fragments.

Analysis

Hierarchical mixed effects with panelist random effects are used to find associations between internet usage features and depression scores. Hierarchical mixed effects models are statistical models that incorporate both fixed effects and random effects. These models are used to analyze data that exhibit nested or hierarchical structures, where observations are grouped into different levels or clusters. Mixed-effects models are particularly useful when there is a need to account for variability at multiple levels. In this longitudinal study, there is variability for each individual panelist, therefore the panelist ID is added as the random effect in the hierarchical models, which translates in practice to linear models with a random intercept for each panelist. Fixed effects are used to model systematic and non-random influences on the dependent variable. They represent factors for which the goal is to estimate specific, population-level, constant effects in regards to the dependent variable. In the context of this study, the features of interest are the generated internet usage features and the aim is to find the associations of these features with the PHQ-9 scores while also accounting for the fixed effect of demographic and season features and individual level variability (panelist random effect). The panelist random effect translates to a random intercept for each panelist. The general equation for the therefore defined as follows, where y in the PHQ-9 score, $y|panelist$ the random intercept effect of the panelist, and $f(x_1,...,x_{20})$ are the fixed effects included in the model:

$$
\begin{aligned}
\text{Population} &:= y|panelist = f(x_1,x_2,x_3,...,x_{20}) = \beta_0 + \beta x_1 + ... + \beta x_{20} \\
\text{Female} &:= y|panelist = f(x_1,x_3,x_4,...,x_{20}) = \beta_0 + \beta x_1 + \beta x_3 + ... + \beta x_{20} \\
\text{Male} &:= y|panelist = f(x_1,x_3,x_4,...,x_{20}) = \beta_0 + \beta x_1 + \beta x_3 + ... + \beta x_{20} \\
18-29 &:= y|panelist = f(x_1,x_2,x_3,...,x_{20}) = \beta_0 + \beta x_1 + ... + \beta x_{20} \\
30\text{–}39 &:= y|panelist = f(x_1,x_2,x_3,...,x_{20}) = \beta_0 + \beta x_1 + ... + \beta x_{20} \\
40\text{–}49 &:= y|panelist = f(x_1,x_2,x_3,...,x_{20}) = \beta_0 + \beta x_1 + ... + \beta x_{20} \\
50\text{–}59 &:= y|panelist = f(x_1,x_2,x_3,...,x_{20}) = \beta_0 + \beta x_1 + ... + \beta x_{20} \\
60+ &:= y|panelist = f(x_1,x_2,x_3,...,x_{20}) = \beta_0 + \beta x_1 + ... + \beta x_{20}
\end{aligned}
$$

Where the fixed effects are:

$$
\begin{aligned}
x_1 := \text{age}, x_2 := \text{gender}, x_3 = \text{season}, x_4 = \text{Night}, x_5 = \text{Morning}, x_6 = \text{Afternoon}, x_7 = \text{Evening}, \\
x_8 = \text{shopping}, x_9 = \text{chat and messaging}, x_{10} := \text{social networking}, x_{11} := \text{health}, \\
x_{12} := \text{search engines and portals}, x_{13} := \text{games}, x_{14} := \text{gambling}, x_{15} := \text{email}, x_{16} := \text{streaming media}, \\
x_{17} := \text{message boards and forums}, x_{18} := \text{job related}, x_{19} := \text{adult}, x_{20} := \text{education}
\end{aligned}
$$

And the features $x_4,..,x_{20}$ refer to the total duration of the internet activity of internet usage at the time of day in the two weeks prior taking the survey. For all models, the highest variance inflation factor (VIF) is 3.3 (afternoon duration). This VIF value is reasonably small, indicating negligible levels of multicollinearity between the fixed effects[29][30]. The models are run using the `lme4` R library[31], which uses the Nakagawa et al.[32] method to compute the fixed effect coefficients. A p-value of 0.05 is used to identify statistically significant coefficients among the features for each model. The Beta coefficients translate to the effect of a unit increase of the feature on the dependent variable, making it challenging to compare the relative importance of the features in the model. For this reason, standardize coefficients are calculated using the Gelman method[33]. The standardized coefficients (std. Beta) are directly comparable and can give more insights about the strength of the features on the dependent variable.

Table 7. Results for the panel analysis predicting PHQ-9 depression scores from the duration spent on internet use activities. The feature gender2 refers to the female gender, Season2 refers to the autumn season.

	Panel			
Predictors	Beta	std. Beta	CI	std. CI
(Intercept)	9.809 ***	-0.055	8.424 – 11.194	-0.137 – 0.026
age	**-0.089** ***	**-0.198**	**-0.116 – -0.063**	**-0.256 – -0.139**
gender2	**1.272** ***	**0.230**	**0.624 – 1.920**	**0.113 – 0.347**
Season2	**-0.387** ***	**-0.070**	**-0.592 – -0.182**	**-0.107 – -0.033**
N: duration	-0.000	-0.035	-0.001 – 0.000	-0.084 – 0.013
M: duration	-0.000	-0.014	-0.001 – 0.000	-0.058 – 0.029
A: duration	0.000	0.048	-0.000 – 0.001	-0.006 – 0.101
E: duration	0.000	0.001	-0.000 – 0.000	-0.053 – 0.054
shopping: duration	0.001	0.021	-0.000 – 0.002	-0.009 – 0.052
chat and messaging: duration	-0.001	-0.019	-0.003 – 0.001	-0.063 – 0.024
social networking: duration	-0.000	-0.025	-0.001 – 0.000	-0.074 – 0.025
health: duration	-0.002	-0.005	-0.013 – 0.008	-0.028 – 0.017
search engines and portals: duration	0.000	0.033	-0.000 – 0.001	-0.018 – 0.083
games: duration	-0.000	-0.011	-0.002 – 0.001	-0.052 – 0.031
gambling: duration	-0.001	-0.013	-0.003 – 0.002	-0.052 – 0.026
email: duration	-0.001	-0.017	-0.003 – 0.001	-0.058 – 0.024
streaming media: duration	0.000	0.017	-0.001 – 0.001	-0.027 – 0.061
message boards and forums: duration	**0.017** ***	**0.060**	**0.008 – 0.026**	**0.029 – 0.091**
job related: duration	0.012	0.006	-0.030 – 0.054	-0.015 – 0.028
adult: duration	**-0.001** *	**-0.039**	**-0.002 – -0.000**	**-0.076 – -0.003**
education: duration	-0.000	-0.011	-0.001 – 0.001	-0.047 – 0.025
Random Effects				
σ^2	6.26			
τ00 panelist	22.80			
ICC	0.78			
N panelist	953			
Observations	2840			
Marginal R^2 / Conditional R^2	0.067 / 0.799			
AIC	15826.420			
log-Likelihood	-7777.411			
* p0.05 ** p0.01 *** p0.001				

Table 9. Results for the gender analysis predicting PHQ-9 depression scores from the duration spent on internet use activities. The feature gender2 refers to the female gender, Season2 refers to the autumn season.

Gender	Male				Female			
Predictors	Beta	std. Beta	CI	std. CI	Beta	std. Beta	CI	std. CI
(Intercept)	9.784 ***	0.043	7.805 – 11.762	-0.040 – 0.126	10.974 ***	0.064	9.186 – 12.761	-0.022 – 0.150
age	**-0.093** ***	-0.195	-0.131 – -0.054	-0.277 – -0.114	**-0.083** ***	-0.195	-0.119 – -0.047	-0.280 – -0.110
Season2	**-0.281** *	-0.052	-0.543 – -0.019	-0.100 – -0.003	**-0.522** **	-0.094	-0.842 – -0.201	-0.152 – -0.036
N: duration	-0.000	-0.029	-0.001 – 0.000	-0.089 – 0.031	-0.001	-0.038	-0.002 – 0.001	-0.127 – 0.052
M: duration	0.000	0.015	-0.000 – 0.001	-0.038 – 0.069	-0.001	-0.051	-0.001 – 0.000	-0.128 – 0.027
A: duration	0.000	0.013	-0.001 – 0.001	-0.059 – 0.084	0.001	0.078	-0.000 – 0.001	-0.006 – 0.161
E: duration	0.000	0.020	-0.000 – 0.001	-0.052 – 0.092	-0.000	-0.035	-0.001 – 0.000	-0.120 – 0.050
shopping: duration	0.000	0.008	-0.001 – 0.002	-0.032 – 0.047	0.002	0.043	-0.000 – 0.005	-0.008 – 0.094
chat and messaging: duration	0.001	0.031	-0.001 – 0.004	-0.030 – 0.093	**-0.008** ***	-0.094	-0.013 – -0.003	-0.149 – -0.039
social networking: duration	-0.000	-0.007	-0.001 – 0.001	-0.069 – 0.054	-0.000	-0.016	-0.001 – 0.001	-0.103 – 0.071
health: duration	-0.002	-0.003	-0.023 – 0.019	-0.034 – 0.028	-0.003	-0.009	-0.015 – 0.009	-0.043 – 0.025
search engines and portals: duration	0.000	0.016	-0.001 – 0.002	-0.038 – 0.071	0.000	0.040	-0.001 – 0.001	-0.058 – 0.138
games: duration	-0.000	-0.014	-0.003 – 0.002	-0.078 – 0.050	0.000	0.000	-0.002 – 0.002	-0.055 – 0.056
gambling: duration	-0.000	-0.006	-0.003 – 0.002	-0.062 – 0.051	-0.005	-0.031	-0.012 – 0.002	-0.074 – 0.011
email: duration	0.000	0.009	-0.002 – 0.003	-0.047 – 0.065	-0.002	-0.039	-0.004 – 0.001	-0.101 – 0.022
streaming media: duration	0.001	0.041	-0.000 – 0.002	-0.018 – 0.100	0.000	0.017	-0.001 – 0.002	-0.059 – 0.093
message boards and forums: duration	**0.017** ***	0.082	0.008 – 0.026	0.041 – 0.123	-0.019	-0.012	-0.082 – 0.043	-0.051 – 0.027
job related: duration	**0.096** *	0.038	0.022 – 0.170	0.009 – 0.068	-0.021	-0.013	-0.073 – 0.031	-0.045 – 0.019
adult: duration	**-0.001** *	-0.054	-0.002 – -0.000	-0.102 – -0.005	-0.003	-0.025	-0.010 – 0.004	-0.080 – 0.030
education: duration	-0.001	-0.033	-0.002 – 0.000	-0.087 – 0.021	0.001	0.024	-0.002 – 0.004	-0.036 – 0.084
Random Effects								
σ^2	5.48				6.95			
τ00	22.65 panelist				22.91 panelist			
ICC	0.81				0.77			
N	495 panelist				458 panelist			
Observations	1509				1331			
Marginal R^2 / Conditional R^2	0.055 / 0.816				0.060 / 0.781			
AIC	8362.702				7647.811			
log-Likelihood	-4052.751				-3701.657			

* p0.05 ** p0.01 *** p0.001

Table 11. Results for the age-group analysis predicting PHQ-9 depression scores from the duration spent on internet use activities. The feature gender2 refers to the female gender, Season2 refers to the autumn season.

Age Group	18-29				30-39				40-49			
Predictors	Beta	std. Beta	CI	std. CI	Beta	std. Beta	CI	std. CI	Beta	std. Beta	CI	std. CI
(Intercept)	22.026 ***	-0.164	10.732 – 33.320	-0.533 – 0.206	18.372 ***	0.053	8.005 – 28.739	-0.157 – 0.263	16.124 **	-0.052	4.979 – 27.269	-0.207 – 0.103
age	**-0.590** **	-0.265	-1.033 – -0.147	-0.463 – -0.066	**-0.296** *	-0.144	-0.589 – -0.003	-0.287 – -0.002	-0.249	-0.117	-0.499 – 0.001	-0.235 – 0.001
gender2	1.032	0.173	-1.617 – 3.680	-0.271 – 0.617	0.123	0.021	-1.549 – 1.796	-0.269 – 0.312	**1.897** **	0.342	0.542 – 3.252	0.098 – 0.587
Season2	0.352	0.059	-0.499 – 1.202	-0.084 – 0.201	-0.394	-0.069	-1.016 – 0.228	-0.177 – 0.040	**-0.612** **	-0.110	-1.052 – -0.172	-0.190 – -0.031
N: duration	-0.003	-0.146	-0.007 – 0.001	-0.331 – 0.039	-0.001	-0.153	-0.004 – 0.001	-0.428 – 0.123	-0.001	-0.067	-0.002 – 0.000	-0.167 – 0.034
M: duration	0.001	0.078	-0.001 – 0.004	-0.081 – 0.237	-0.001	-0.084	-0.003 – 0.001	-0.225 – 0.057	0.001	0.051	-0.000 – 0.001	-0.037 – 0.139
A: duration	-0.000	-0.004	-0.003 – 0.003	-0.246 – 0.239	-0.000	-0.022	-0.002 – 0.001	-0.206 – 0.163	0.000	0.047	-0.001 – 0.002	-0.080 – 0.173
E: duration	-0.000	-0.030	-0.003 – 0.002	-0.228 – 0.168	-0.001	-0.109	-0.003 – 0.001	-0.295 – 0.078	-0.000	-0.006	-0.001 – 0.001	-0.122 – 0.109
shopping: duration	0.003	0.041	-0.006 – 0.012	-0.089 – 0.172	-0.001	-0.013	-0.006 – 0.005	-0.105 – 0.079	0.000	0.010	-0.002 – 0.002	-0.045 – 0.065
chat and messaging: duration	-0.018	-0.086	-0.037 – 0.001	-0.178 – 0.006	-0.007	-0.043	-0.020 – 0.005	-0.115 – 0.030	-0.008	-0.050	-0.020 – 0.004	-0.122 – 0.022
social networking: duration	0.004	0.069	-0.003 – 0.010	-0.055 – 0.193	0.002	0.049	-0.003 – 0.006	-0.091 – 0.190	0.001	0.042	-0.001 – 0.002	-0.064 – 0.147
health: duration	0.043	0.039	-0.042 – 0.128	-0.038 – 0.115	-0.006	-0.008	-0.052 – 0.041	-0.072 – 0.057	-0.009	-0.013	-0.045 – 0.027	-0.068 – 0.042
search engines and portals: duration	0.007	0.081	-0.004 – 0.019	-0.049 – 0.211	0.001	0.248	-0.000 – 0.003	-0.061 – 0.557	-0.000	-0.001	-0.002 – 0.002	-0.061 – 0.059
games: duration	0.005	0.009	-0.064 – 0.075	-0.111 – 0.128	-0.003	-0.070	-0.008 – 0.002	-0.192 – 0.053	-0.001	-0.023	-0.004 – 0.002	-0.094 – 0.047
gambling: duration	-0.246	-0.036	-0.817 – 0.325	-0.121 – 0.048	-0.035	-0.015	-0.188 – 0.118	-0.082 – 0.051	-0.004	-0.039	-0.013 – 0.005	-0.119 – 0.041
email: duration	**-0.016** *	-0.154	-0.030 – -0.001	-0.292 – -0.015	0.006	0.042	-0.009 – 0.020	-0.063 – 0.146	-0.003	-0.048	-0.007 – 0.002	-0.139 – 0.042
streaming media: duration	0.004	0.204	-0.000 – 0.009	-0.018 – 0.426	0.002	0.158	-0.001 – 0.004	-0.063 – 0.379	0.001	0.061	-0.001 – 0.003	-0.059 – 0.181
message boards and forums: duration	**-0.115** *	-0.107	-0.203 – -0.026	-0.190 – -0.024	0.017	0.038	-0.032 – 0.065	-0.073 – 0.149	0.013	0.044	-0.011 – 0.036	-0.040 – 0.127
job related: duration	0.021	0.016	-0.089 – 0.130	-0.068 – 0.099	0.061	0.032	-0.058 – 0.179	-0.031 – 0.096	0.029	0.012	-0.075 – 0.133	-0.031 – 0.055
adult: duration	0.012	0.048	-0.011 – 0.035	-0.042 – 0.138	-0.004	-0.032	-0.016 – 0.007	-0.115 – 0.051	-0.001	-0.044	-0.003 – 0.001	-0.118 – 0.030
education: duration	0.003	0.045	-0.006 – 0.011	-0.105 – 0.195	0.002	0.045	-0.003 – 0.008	-0.063 – 0.153	-0.001	-0.044	-0.003 – 0.001	-0.162 – 0.075
Random Effects												
σ^2	6.71				8.15				7.08			
$\tau00$	22.05 panelist				23.70 panelist				22.43 panelist			
ICC	0.77				0.74				0.76			
N	72 panelist				163 panelist				230 panelist			
Observations	202				441				702			
Marginal R^2 / Conditional R^2	0.134 / 0.798				0.054 / 0.758				0.054 / 0.773			
AIC	1296.794				2726.803				4126.387			
log-Likelihood	-559.196				-1259.923				-1946.883			

* p0.05 ** p0.01 *** p0.001

Table 13. Results for the age-group analysis predicting PHQ-9 depression scores from the duration spent on internet use activities. The feature gender2 refers to the female gender, Season2 refers to the autumn season.

Age Group	50-59				60+			
Predictors	Beta	std. Beta	CI	std. CI	Beta	std. Beta	CI	std. CI
(Intercept)	8.104	-0.054	-3.754 – 19.963	-0.209 – 0.102	16.545 **	-0.070	6.077 – 27.014	-0.239 – 0.099
age	-0.057	-0.030	-0.272 – 0.159	-0.142 – 0.083	**-0.188** *	-0.137	-0.350 – -0.027	-0.254 – -0.019
gender2	**1.258** *	0.232	0.020 – 2.496	0.004 – 0.461	1.141	0.233	-0.042 – 2.323	-0.009 – 0.475
Season2	-0.337	-0.062	-0.682 – 0.008	-0.126 – 0.002	**-0.461** *	-0.094	-0.844 – -0.079	-0.173 – -0.016
N: duration	-0.001	-0.065	-0.003 – 0.000	-0.146 – 0.016	0.001	0.034	-0.001 – 0.002	-0.048 – 0.116
M: duration	0.000	0.011	-0.001 – 0.001	-0.071 – 0.092	-0.001	-0.057	-0.002 – 0.001	-0.168 – 0.055
A: duration	0.000	0.002	-0.001 – 0.001	-0.093 – 0.096	**0.001** *	0.120	0.000 – 0.002	0.012 – 0.228
E: duration	0.000	0.012	-0.001 – 0.001	-0.086 – 0.110	0.000	0.046	-0.000 – 0.001	-0.070 – 0.162
shopping: duration	**0.003** *	0.070	0.000 – 0.005	0.002 – 0.137	0.002	0.033	-0.002 – 0.005	-0.040 – 0.106
chat and messaging: duration	**-0.004** *	-0.054	-0.008 – -0.000	-0.104 – -0.003	0.001	0.039	-0.002 – 0.004	-0.070 – 0.149
social networking: duration	-0.000	-0.019	-0.001 – 0.001	-0.110 – 0.072	-0.001	-0.126	-0.002 – 0.000	-0.254 – 0.001
health: duration	0.004	0.008	-0.016 – 0.024	-0.031 – 0.046	-0.004	-0.016	-0.016 – 0.008	-0.063 – 0.030
search engines and portals: duration	0.000	0.022	-0.001 – 0.002	-0.060 – 0.104	-0.001	-0.038	-0.002 – 0.001	-0.126 – 0.050
games: duration	0.002	0.073	-0.001 – 0.004	-0.020 – 0.166	-0.002	-0.032	-0.005 – 0.002	-0.097 – 0.034
gambling: duration	-0.002	-0.046	-0.006 – 0.002	-0.146 – 0.054	-0.001	-0.022	-0.005 – 0.002	-0.082 – 0.039
email: duration	-0.001	-0.031	-0.004 – 0.002	-0.104 – 0.042	-0.000	-0.005	-0.003 – 0.003	-0.090 – 0.081
streaming media: duration	0.001	0.034	-0.001 – 0.003	-0.047 – 0.114	-0.002	-0.066	-0.005 – 0.000	-0.146 – 0.013
message boards and forums: duration	**0.019** ***	0.104	0.010 – 0.029	0.051 – 0.157	0.024	0.021	-0.051 – 0.099	-0.044 – 0.085
job related: duration	-0.012	-0.007	-0.074 – 0.050	-0.042 – 0.029	0.007	0.003	-0.105 – 0.120	-0.045 – 0.051
adult: duration	-0.001	-0.060	-0.003 – 0.000	-0.135 – 0.015	-0.002	-0.033	-0.005 – 0.001	-0.097 – 0.031
education: duration	-0.000	-0.008	-0.003 – 0.002	-0.078 – 0.062	-0.002	-0.062	-0.006 – 0.001	-0.166 – 0.041
Random Effects								
σ^2	4.74				5.29			
$\tau 00$	23.42 panelist				17.56 panelist			
ICC	0.83				0.77			
N	265 panelist				223 panelist			
Observations	797				698			
Marginal R^2 / Conditional R^2	0.045 / 0.839				0.064 / 0.783			
AIC	4459.298				3910.500			
log-Likelihood	-2107.745				-1836.435			

* p0.05 ** p0.01 *** p0.001

References

1. Liu, Q. *et al.* Changes in the global burden of depression from 1990 to 2017: Findings from the global burden of disease study. **126**, 134–140, DOI: 10.1016/j.jpsychires.2019.08.002.
2. Organization, W. H. Depressive disorder (2023). Accessed: January 21, 2024.
3. Santomauro, D. F. *et al.* Global prevalence and burden of depressive and anxiety disorders in 204 countries and territories in 2020 due to the covid-19 pandemic. *The Lancet* **398**, 1700–1712 (2021).
4. Lotfi, L., Flyckt, L., Krakau, I., Mårtensson, B. & Nilsson, G. H. Undetected depression in primary healthcare: occurrence, severity and co-morbidity in a two-stage procedure of opportunistic screening. *Nord. journal psychiatry* **64**, 421–427, DOI: 10.3109/08039481003786378 (2010).
5. Hökby, S. *et al.* Are mental health effects of internet use attributable to the web-based content or perceived consequences of usage? a longitudinal study of european adolescents. **3**, e31, DOI: 10.2196/mental.5925.
6. Lam, S. S. M., Jivraj, S. & Scholes, S. Exploring the relationship between internet use and mental health among older adults in england: Longitudinal observational study. **22**, e15683, DOI: 10.2196/15683.
7. Katikalapudi, R., Chellappan, S., Montgomery, F., Wunsch, D. & Lutzen, K. Associating internet usage with depressive behavior among college students. **31**, 73–80, DOI: 10.1109/MTS.2012.2225462.
8. Yue, C. *et al.* Automatic depression prediction using internet traffic characteristics on smartphones. **18**, 100137, DOI: 10.1016/j.smhl.2020.100137.
9. Razavi, R., Gharipour, A. & Gharipour, M. Depression screening using mobile phone usage metadata: a machine learning approach. **27**, 522–530, DOI: 10.1093/jamia/ocz221.
10. Bessière, K., Pressman, S., Kiesler, S. & Kraut, R. Effects of internet use on health and depression: a longitudinal study. *J. medical Internet research* **12**, e6 (2010).
11. Bessière, K., Kiesler, S., Kraut, R. & Boneva, B. S. EFFECTS OF INTERNET USE AND SOCIAL RESOURCES ON CHANGES IN DEPRESSION. **11**, 47–70, DOI: 10.1080/13691180701858851.
12. Fan Zhang, Tingshao Zhu, Ang Li, Yilin Li & Xinguo Xu. A survey of web behavior and mental health. In *2011 6th International Conference on Pervasive Computing and Applications*, 189–195, DOI: 10.1109/ICPCA.2011.6106503 (IEEE).
13. Purwandari, B., Wibawa, W. S., Fitriah, N., Christia, M. & Bintari, D. R. Internet addiction and mental health prediction using ensemble learning based on web browsing history. In *Proceedings of the 3rd International Conference on Software Engineering and Information Management*, 155–159, DOI: 10.1145/3378936.3378947 (ACM).
14. Müller, A. *et al.* Online shopping in treatment-seeking patients with buying-shopping disorder. *Compr. Psychiatry* **94**, 152120, DOI: 10.1016/j.comppsych.2019.152120 (2019).
15. Berryman, C., Ferguson, C. J. & Negy, C. Social media use and mental health among young adults. *Psychiatr. quarterly* **89**, 307–314, DOI: 10.1007/s11126-017-9535-6 (2018).
16. Willoughby, B. J., Carroll, J. S., Nelson, L. J. & Padilla-Walker, L. M. Associations between relational sexual behaviour, pornography use, and pornography acceptance among us college students. *Cult. health & sexuality* **16**, 1052–1069 (2014).
17. Perry, S. L. Pornography use and depressive symptoms: Examining the role of moral incongruence. *Soc. Mental Heal.* **8**, 195–213 (2018).
18. Willoughby, B. J., Busby, D. M. & Young-Petersen, B. Understanding associations between personal definitions of pornography, using pornography, and depression. *Sex. Res. Soc. Policy* **16**, 342–356 (2019).
19. American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders* (American Psychiatric Association, Arlington, VA, 2013), 5th edn.
20. Brodie, M. *et al.* Health information, the internet, and the digital divide: Despite recent improvements, americans' access to the internet—and to the growing body of health information there—remains uneven. *Heal. affairs* **19**, 255–265 (2000).
21. Baker, L., Wagner, T. H., Singer, S. & Bundorf, M. K. Use of the internet and e-mail for health care information: results from a national survey. *Jama* **289**, 2400–2406 (2003).
22. Drentea, P., Goldner, M., Cotten, S. & Hale, T. The association among gender, computer use and online health searching, and mental health. *Information, Commun. & Soc.* **11**, 509–525 (2008).

23. Althubaiti, A. Information bias in health research: definition, pitfalls, and adjustment methods. *J. Multidiscip. Healthc.* **9**, 211–217, DOI: 10.2147/JMDH.S104807 (2016). PMID: 27217764, https://www.tandfonline.com/doi/pdf/10.2147/JMDH.S104807.

24. Kroenke, K., Spitzer, R. L. & Williams, J. B. W. The phq-9. *J. Gen. Intern. Medicine* **16**, 606–613, DOI: https://doi.org/10.1046/j.1525-1497.2001.016009606.x (2001). https://onlinelibrary.wiley.com/doi/pdf/10.1046/j.1525-1497.2001.016009606.x.

25. Luppa, M. *et al.* Age-and gender-specific prevalence of depression in latest-life–systematic review and meta-analysis. *J. affective disorders* **136**, 212–221 (2012).

26. Melrose, S. *et al.* Seasonal affective disorder: an overview of assessment and treatment approaches. *Depress. research treatment* **2015** (2015).

27. Murugan, Y., Nagarajan, P., Subrahmanyam, D. & Kattimani, S. Severity of loneliness, depression and perceived social support in adults in the empty nest stage of the family life cycle and the influence of using digital technology. *Asian J. Psychiatry* **76**, 103245 (2022).

28. Webshrinker. Webshrinker simplified categories (2023). Accessed: December 14, 2023.

29. Chatterjee, S. & Hadi, A. S. *Regression analysis by example* (John Wiley & Sons, 2015).

30. Midi, H. & Bagheri, A. Robust multicollinearity diagnostic measure in collinear data set. In *Proceedings of the 4th international conference on applied mathematics, simulation, modeling*, 138–142 (World Scientific and Engineering Academy and Society (WSEAS), 2010).

31. Bates, D., Mächler, M., Bolker, B. & Walker, S. *lme4: Linear Mixed-Effects Models using 'Eigen' and S4* (2022). R package version 1.1-27.

32. Nakagawa, S. & Schielzeth, H. A general and simple method for obtaining r^2 from generalized linear mixed-effects models. *Methods Ecol. Evol.* **4**, 133–142, DOI: 10.1111/j.2041-210x.2012.00261.x (2013).

33. Gelman, A. Scaling regression inputs by dividing by two standard deviations. *Stat. medicine* **27**, 2865–2873 (2008).