Do Daily Levels Of Anxiety And Depression Predict The Amount Of Binge Watching Behaviour On The Same Day?

-

An ESM Study Investigating The Associations Between Depression, Anxiety And Binge Watching Behaviour

Bachelor Thesis

Robert Preissler (1972324)

28.07.2020, Enschede

University of Twente

Department of Psychology, Health & Technology

First Supervisor: Dr. P. ten Klooster

Second Supervisor: Dr. M. Noordzij

Abstract

Objective

With lockdowns and social distancing as major efforts to fight the emergent coronavirus, media consumption is on the rise, especially on video-on-demand platforms such as Netflix. This increasing consumption can foster an already popular behaviour: binge watching. To date, little is known about the temporal associations between mental health and binge-watching, especially with respect to feelings of anxiety and depression.

Methods

To analyze these temporal associations, this study uses the experience sampling method (ESM) design. A convenience sample of 38 participants made daily recordings of how much video-on-demand media they consumed, concerning both hours and episodes over a period of two weeks. Also, to analyze state feelings of anxiety and depression, participants were prompted to fill in the Patient-Health-Questionnaire-4 (PHQ-4), a short-form questionnaire for the two mental health constructs. Moreover, participants filled in longer versions of the PHQ-4 at the beginning and end of the study for trait assessments of depression and anxiety. After the two week period, participant's levels of binge watching behaviour and mental health symptoms were compared and analyzed with repeated measures linear mixed model analysis.

Results

The data showed no significant results regarding the major research questions, namely whether anxiety or depression significantly predict binge watching behaviour in any way. However, trait anxiety scores in the final assessment and binge watching behaviour during the study period were significantly, although weakly, associated (B-Estimate = 0.02; p-value = 0.049).

Conclusion

The current study showed that daily depression and anxiety were not associated with binge-watching behavior on the same day. As many different definitions of binge watching exist, comparisons to other studies are difficult. In the future, the definitions of binge watching should be consolidated to improve comparisons and research should focus on different target groups and longer time periods.

An ESM Study Investigating the Associations between Depression, Anxiety and Binge Watching Behaviour

At the end of 2019, Netflix reported an estimated total of 167 million worldwide subscribers (Watson, 2020). With the onset of the Corona-Crisis in the western world around the beginning of 2020 and increasingly strict lockdown measures by the government, the number of subscribers is skyrocketing (Schumacher, 2020). According to a report from BBC News (2020), almost 16 million people created accounts between January and March 2020, which is twice the number of subscriptions in the beginning of last year.

When more and more people get an account and are locked down at home, streaming hours might also increase tremendously. Then one activity may become increasingly prevalent: binge watching. Binge watching is a term coined by popular media and describes a form of excessive media consumption pertaining mostly to television series, especially when watched on online, video-on-demand streaming platforms such as Netflix or Hulu (Steinbach, 2018). Despite being described as an excessive behaviour, there are countless articles that actually give readers lists such as "the 53 best Netflix series to binge watch" (Wired, 2020), so apparently binge watching seems to be advocated as a widely accepted and even sought after activity. But users should be cautious. As Tefertiller and Maxwell (2018) found, binge watching is still a fairly new concept and there is a strong need of published research regarding the potential psychological and health implications of binge watching.

Despite the popularity of binge watching both as a behavior and as a topic of research, in the literature into binge watching there are countless different definitions of binge watching used and they all focus on different specific aspects (Flayelle, 2020). There is currently no conclusive agreement in the scientific community on a single and clear definition for binge watching behaviour. This ambiguous nature of binge watching definitions has led to much criticism of the concept as a whole (Perks, 2014; Pittman & Sheehan, 2015). Therefore, this study aggregated the most common aspects that previous research used to operationalize binge watching. Based on a review of the literature, for the purpose of this study, binge watching will be defined as watching at least two episodes or one hour of the same show on video-on-demand streaming platforms (Flayelle, 2020; Mikos, 2016; Steiner & Xu, 2018).

Much scientific research has concerned itself with the potentially detrimental effects of binge watching behavior and, as the name 'binge' watching already indicates, a review of the literature reveals a potential connection between binge watching and addiction

(Heatherton & Baumeister, 1991; Leon, Carroll, Chernyk, & Finn, 1985 as cited in Steinbach, 2018). 'Binging' is commonly associated with behaviours of similar definitions, such as binge drinking or binge eating. This addictive characteristic is also facilitated by the fact that streaming platforms like Netflix are available on-demand, all the time and everywhere (Steinbach, 2018). Also, similar binge behaviours have been associated with social anxiety and depression (Brechan & Kvalem, 2015; Tokunaga & Rains, 2010). Tukachinsky and Eyal (2018) carried out a study where they found that individuals with more depressive symptoms engaged in more intense binge watching behaviour and theorized that this might be due to the higher regulation deficiency that those individuals exhibit. Some studies conclude that due to worse levels of mental health people tend to binge watch more, others conclude that binge watching is used to regulate negative emotions. Still, while many previous studies have pointed to a connection between binge watching and mental health problems like depression, currently there is no conclusive evidence how this relationship exactly works (Wheeler, 2015).

According to Pinel and Barnes (2014), depression is a normal reaction, that can occur in many situations of regular life. Examples for those triggering events might be the loss of a loved one, or loss of physical or mental health. Despite the ordinary nature of depression, some people have a tendency towards stronger symptoms. Yet, for this study the focus is not on clinical depression as it is classified in the DSM-5 (Truschel, n.d.). Rather, the more common levels of depression and its symptoms such as anhedonia are considered. For the purpose of this study depression is operationalized by two main common aspects. Firstly, anhedonia, which means to have little interest or pleasure in doing things. Secondly, dysphoria, which describes feeling down, depressed, or hopeless (Kroenke, et. al., 2009).

The second mental health issue that is examined in this study is anxiety. In general, anxiety describes feelings of fear that persist despite the absence of any immediate threat. This can lead to a disruption of normal behaviour, as people avoid certain situations to avoid those dreadful emotions (Pinel & Barnes, 2014). While most other studies in the field of binge watching focus specifically on attachment anxiety, this study aims at the more general type of anxiety, in order to advance the research in new directions (Flayelle, 2020; Wheeler, 2015). Still, a number of those studies are considered to introduce the potential relationship between anxiety and binge watching behaviour in more depth. Moreover, this study does not focus on the clinical characteristics of anxiety that are described in the DSM-5, but rather on more frequent and common everyday feelings (Truschel, n.d.). Hence, it examines anxiety as

a state of feeling nervous, anxious or on edge and not being able to stop or control worrying (Kroenke, et. al., 2009).

In contrast to the variety of research on binge watching and depression, the relationship between binge watching and anxiety is less researched. In a recent study, Tefertiller and Maxwell (2018) found that there are some indications that increased levels of anxiety are associated with an increased likelihood of binge watching, yet they describe their results as not conclusive enough, given a lack of statistical significance. Most commonly, the potential connection between anxiety and binge watching behaviour is explained by means of attachment theory (Tokunaga & Rains, 2010). Based on the idea that viewers form one-sided relationships with characters in a show, viewers keep coming back for more exposure (Devasagayam, 2014). It is theorized that those para-social relationships function the same way as regular interpersonal relationships, and this connects binge watching to social anxiety. People who have an insecure attachment style might watch more, in order to keep those para-social relationships alive (Tefertiller & Maxwell, 2018). This is supported by other literature, as Wheeler (2015) for example found that attachment anxiety was significantly positively associated with watching back-to-back episode of the same television show. Yet, there are also other studies which found that participants who scored high on anxiety questionnaires tended to score low on binge watching behaviour (Clarke, 2019). While there are already many preliminary insights into how depression and anxiety might affect binge watching behaviour, there is still no clear consensus yet, especially considering the temporal nature of these associations.

For that reason, this study does not use a conventional cross-sectional method to measure depression, anxiety and binge watching behaviour. Instead, the ‘Experience Sampling Method’ (ESM) is utilized. In experience sampling, participants are prompted at random times during a normal day to fill out self-report questionnaires (Larson, & Csikszentmihalyi, 2014). Nowadays, participants are mostly signaled through their mobile phones and fill out the questionnaires online or in a specifically for the study downloaded application. ESM is a sampling technique that collects data in everyday situations and therefore measures behaviors and feelings of people more close to the moment they occur, as opposed to qualitative or quantitative survey studies, which take a retrospective approach (Larson & Csikszentmihalyi, 1983 as cited in Xie, Heddy & Greene, 2019). Furthermore, ESM differs from conventional online surveys through the amount of measurements and participants. While in online survey studies there is usually only one measurement among many participants, ESM studies take the opposite approach. The number of participants is

usually smaller, but there are many measurements over a given time period. While there have been a number of studies that conducted research on the topic of binge watching behaviour, only very few have used ESM as the means of data collection. To illustrate, out of 23 studies that were included in a recent systematic review, not one used an ESM design to study binge watching behaviour (Flayelle, 2020). Except for two studies, where a laboratory experiment was conducted, all other studies used a cross-sectional online survey as a means to gather data. This underlines the need for an intensive longitudinal approach to data collection within the topic of binge watching behaviour, especially when mental health issues are under consideration, as an ESM design allows to consider daily fluctuations of mental health, which has been discounted in the research on the topic until now.

The most important advantage of ESM is that it catches people in the moment (Xie, Heddy & Greene, 2019). This leads to a variety of upsides over more traditional forms of research. In addition, ESM studies are able to measure both behavioural and intrapsychic aspects of daily activity, which makes it a perfect fit for the investigation of such a day-to-day activity as binge watching (Larson, & Csikszentmihalyi, 2014). Also, ESM suits the construct of binge watching well, because it alleviates the possible problem that participants have difficulties remembering their watching time after larger amounts of time, since they are asked on a daily basis (Steinbach, 2018). Finally, as opposed to cross-sectional surveys, ESM allows the study of the temporal association between mental health and binge-watching over time.

In conclusion, experience sampling may provide more insight into the concept of binge watching behaviour and its association with momentary fluctuating feelings of anxiety and depression over time.

Research Questions

To fully utilize the special approach to data collection of this study, two research questions are created that aim at the analysis of the data. Each one focuses on daily fluctuations and changes across participants. Thus, these questions are:

RQ1: How does binge watching behaviour develop over time and across participants?

RQ2: How do state and trait levels of depression and anxiety develop over time and across participants?

In previous research, there has been much debate on whether anxiety and depression might actually be part of the same disorder or share an underlying distress factor. Proponents of this view called it the unitary model (Stavrakaki & Vargo, 1986). Years later, in 2007, the American Psychiatric Association held a conference on the same matter. There, they agreed on the notion that anxiety and depression are actually not the same, but rather two distinct illnesses that simply share a number of features (Goldberg, 2008). Therefore, it was decided to create two separate research questions, one for each construct. Thus, these two research questions are:

RQ3: To what extent do depressive state and trait symptoms predict binge watching behaviour?

RQ4: To what extent do anxiety state and trait symptoms predict binge watching behaviour?

Methods

Design

The experience sampling method (ESM) was chosen as the measurement design for this study. More specifically, the signal-contingent sampling strategy was implemented, so the assessments occurred after participants received a notification on their smartphones at random times during certain time frames of the day (Conner & Lehman, 2012). The study took place between 08.04.2020 (day 0) and 23.04.2018 (day 15), but data collection actually started the day after participants were invited (day 1). The measurement period included the Easter holidays and data were collected during the special circumstances and social lockdown measures caused by the Corona pandemic. Measurements were taken twice each day, once for the watching behaviours and once for the state assessments of anxiety and depression.

The study was designed in this way, to make assessments more frequent and unpredictable, which increases the chance of catching participants in the moment. Yet, participants were given more time to fill in the behavioural assessment, since it is not so much an in-the-moment state, since only the memory of the number of hours and episodes watched on the previous day were of interest. Opposed to this, the state assessment had a much smaller time frame during the evening, to increase the chance that their emotional state before watching was assessed. A study duration of two weeks was chosen, as this is most common for ESM studies that investigate more frequent behaviours such as binge watching, and furthermore to avoid burdening the participants for an extensive period of time without compromising statistical significance of the data (Conner & Lehman, 2012). The registration day (day 0) was included to make sure that all participants started on the same day and to make the transition into the regular assessments easier for them.

Figure 1 shows an outline of the study design, with day 0 as the registration day that was mainly for the participants to sign up and fill out the baseline assessments, so demographics, trait anxiety and depression questionnaires and the informed consent. Days 1 to 14 included the repeated daily data collection and on day 15 the baseline trait assessment was repeated again (Figure 1).

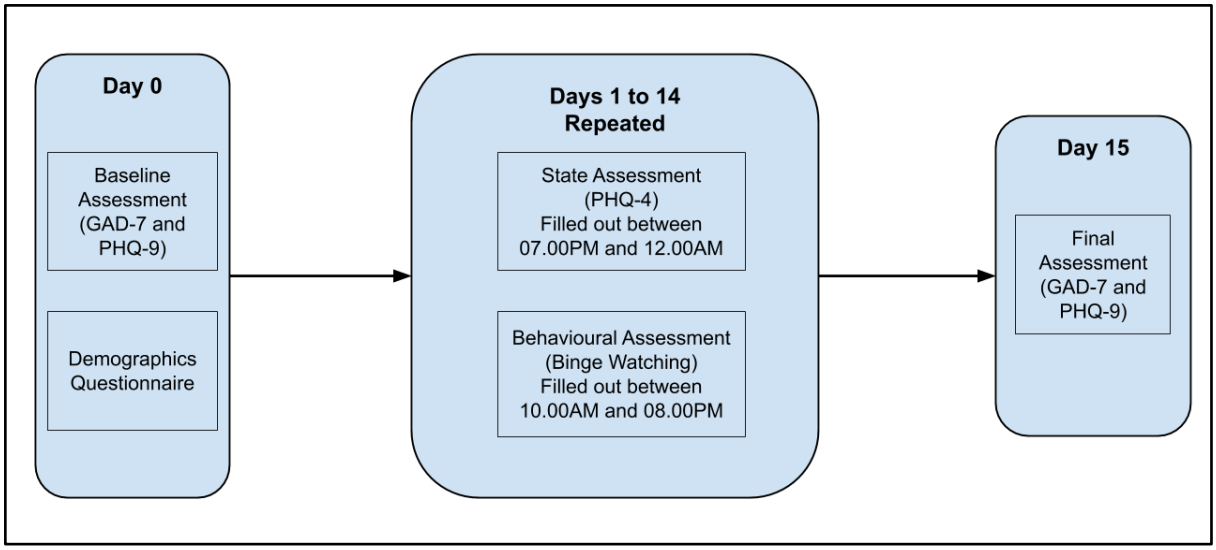


Figure 1. Flowchart of the study design

The questionnaires and assessments were administered with the online tool Ethica Data. This website tool allows the design of research studies without the need for programming expertise. The only requirement is that all participants download the app on their smartphones, they will then receive notifications based on how the app was designed (Ethica Data Services Inc, 2019). Ethica Data allows participants to fill out the assessments very quickly and easily in the app, which makes it easier for them to comply. When filling out the assessments, participants did not have any time restrictions, except for the hours-long time frame that was set in which the questions had to be answered. Also, items were not randomized, so they always came in the same order. Participants could only proceed to the next set of questions when all questions of the previous set were filled in completely, which eliminates the chance of not filling in the assessments correctly or only partly. The Ethics Committee of the Faculty of Behavioural, Management and Social Sciences approved the research. Participants signed an in-app informed consent before participating.

Participants

Given that most of the participants were family members or friends of the researchers, they were recruited either via social media or face-to-face. This route was chosen to ensure that participants would be motivated enough to participate daily over a course of two weeks. It was required for the participants to own a smartphone and to have proficient skills in the English language, in order to participate in this study. Also, only subjects who had access to at least one streaming service were included. Using the guidelines of Conner and Lehman

(2012), of the 42 participants that started the study, those who did not respond to 40% or more of the prompts were deleted from the dataset after completion of data collection. This led to a total of four participants being excluded, due to an average response rate below 60%. Participants overall responded fairly regularly, as the average response rate to the assessments was 81% (SD = 23%, Min = 0%, Max = 100%). The final sample for analysis consisted of 38 participants (*M*age = 23.79, *SD*age = 5.33; 17 Female, 21 Male; Appendix, Table 1).

Procedure

Participants of the study received an email by the research team, with specific instructions on how to participate and use the Ethica Data application. One day after the participants had created an account and signed up for the study following the instructions, they received two daily questionnaires. One in the morning and one in the evening. The first questionnaire asked the participants about their binge-watching behaviour on the previous day, while the second one, the PHQ-4, assessed the participants' anxiety and depression levels at that moment. Additionally, once at the start of the study participants received a questionnaire about their demographics and are asked to fill out longer forms of the two PHQ-4 (Appendix, Figure 1) components: the GAD-7 (Appendix, Figure 2) and the PHQ-9 (Appendix, Figure 3), both with a 2-week recall-period, as part of the baseline trait assessment of anxiety and depression which was repeated again in the end of the study. Filling out the behavioral assessment questionnaire was possible each day from 10.00am to 08.00pm and all participants received notifications on their phones at random times between 10.00am and 10.30am to remind them of this. Similarly, participants received notifications for the evening assessment (the PHQ-4), which they could fill out from 07.00pm to midnight. The reminder for the evening assessment occurred also at a random time, set between 07.00pm and 09.00pm. No compensation was given to the participants.

Measures

A total of three measures for mental health were used in this study. For the baseline and final assessment, the GAD-7 and PHQ-9 to measure trait depression and anxiety levels at the beginning and end of the study were used. Both measures have excellent reliability and validity even in the general population and were very suitable for their purpose (Kroenke & Spitzer, 2002; Spitzer et. al., 2006). The PHQ-9 introduces a total of nine questions, focusing on the prior two weeks. The two areas questions inquire about are anhedonia (6 questions,

e.g. “Little interest or pleasure in doing things”) and dysphoria (3 questions, e.g. “Feeling down, depressed or hopeless”). All statements have to be rated on a four-point scale from “not at all” to “nearly every day”.

Similarly, the GAD-7 poses seven questions that also inquire about the participant’s last two weeks. The two main areas under consideration are restlessness (4 questions, e.g. “trouble relaxing”) and worrying (3 questions, e.g. “feeling nervous, anxious or on edge”) and just as with the PHQ-9, they are rated on four point scale from “not at all” to “nearly every day”. For both tests, scores are added together, and higher scores indicate higher levels of anxiety or depression, respectively. Also, for all four measurements of these questionnaires, so in the beginning and end of study, internal consistency was computed and showed good results (Table 1).

Table 1. Cronbach’s alpha values for the baseline and final measurements

Measurement	Cronbach’s Alpha
Baseline Assessment PHQ-9	0.885
Baseline Assessment GAD-7	0.884
Final Assessment PHQ-9	0.799
Final Assessment GAD-7	0.896

The PHQ-4 was used for the daily measurements. It is an ultra-short questionnaire for depression and anxiety, with good psychometric properties even in the general population (Stanhope, 2016). The short nature of this test allows to incorporate it more easily into the methodology of an ESM design study. The PHQ-4 consists of the two questions from the GAD-2, a measure for anxiety symptoms, and the two questions from the PHQ-2, which measures depressive symptoms. This allows the analysis of each concept individually and together. Since those questionnaires usually inquire about a longer time frame prior to the assessment, the recall period was adapted, so that the questions instead ask about the participant’s past day instead of the usual two weeks.

Finally, the participant’s watching behaviours were measured by a series of questionnaires. Firstly, participants were asked whether or not they had watched video-on-demand content on the previous day. If they answered with a “yes”, then they were asked at

what time they had watched. For this question, participants were offered four possible answers between morning and night. In the end, participants were then asked to type in the number of hours and episodes that they had watched.

Data Analysis

The statistical program for social sciences (SPSS) was used to analyze the data. First, the data from the different questionnaires was downloaded from Ethica Data and transferred from CSV into Excel files, which makes it possible to upload them more easily into SPSS. Next, the datasets for behavioural and state assessment were merged together into a single long format dataset, erroneous data was cleaned and rows with missing input were deleted.

In order to calculate the watching times per time point and participant, descriptive statistics were used to compute the mean values. Participants' anxiety and depression scores were computed by adding together the scores of the different PHQ-4 questions, so that those could be analyzed in detail. Next, a dummy variable was coded to make analysis into binge watching behaviour specifically possible. In line with the findings of the research of the literature, binge-watching was coded as "1" if participants watched at least one hour or two episodes in a row. Everything less constitutes a binge watching score of "0".

To analyze the connection between the different repeated measurements, a series of Linear Mixed Models analyses with an autoregressive covariance structure were conducted, since those types of analyses are able to process nested structures of data with missing values. To gain more insight into how the watching behaviour and mental health scores differ per time point and participant, the latter were used as a fixed factors, while watching behaviour, depression and anxiety scores were the dependent variables. Next, to statistically test the research questions, another series of Linear Mixed Models analyses was conducted and one-day lagged (T-1) scores of depression and anxiety were computed, in order to analyze the connection between mental health scores and watching behaviour as they occurred on the same day. This time, depression and anxiety were set as dependent variables, while the binge watching dummy variable, the number of hours and number of episodes were set as fixed covariates in separate models. Also, a series of Linear Mixed Model Analyses was run with the baseline and final assessment mean scores of anxiety and depression as fixed covariates, to investigate a possible connection between trait-like anxiety and depression and binge watching. Moreover, a Pearson correlation analysis between the means of the baseline and final assessments and the binge watching estimated marginal measurement means per

participant were conducted. Finally, Excel was used to create graphs to visually illustrate the results.

Results

Watching Behaviour Over Time And Participants

The highest amount of watching behaviour occurred on day 6, a Tuesday, with an average of 4.31 episodes (SD = 0.69) and 3.06 hours (SD = 0.43) consumed on that day (Figure 2). Least amounts of watching for both variables occurred on day 9, a Friday, when participants consumed on average 2.18 (SD = 0.65) episodes and 1.67 (SD = 0.43) hours of material from video-on-demand platforms. Neither the number of episodes watched ($p = .153$) nor the number of hours ($p = .152$) were significantly associated with the day of watching. As can be seen from the figure, the two operationalizations of watching behavior, the estimated mean number of episodes watched on a given day and the number of hours watched on the same day, show a very strong positive correlation (Pearson $r = .826$, $p < .001$) with one another over the time points.

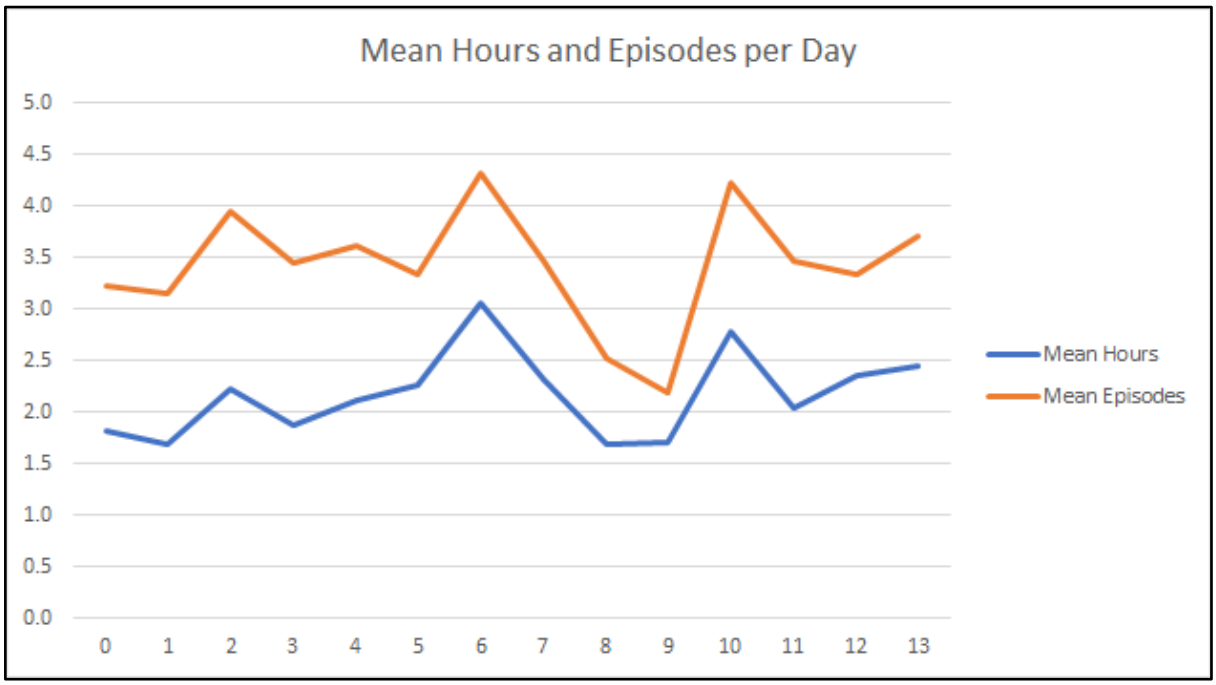


Figure 2. Mean hours and episodes per day

The dummy coded binge watching variable and the day variable also did not show a significant association ($p = .485$) indicating that the proportion of participants that binge watched was not different over time. While the most episodes and hours were watched on day 6, the highest amount of binge-watching occurred on day 10, a Saturday, when half of the participants binge watched (50%, SD = 0.08) (Figure 3). The smallest fraction of binge

watching occurred on days 4 (Sunday) and 8 (Thursday), when only 26.3% (SD = 0.08) of participants met the criteria for binge watching (Figure 3).

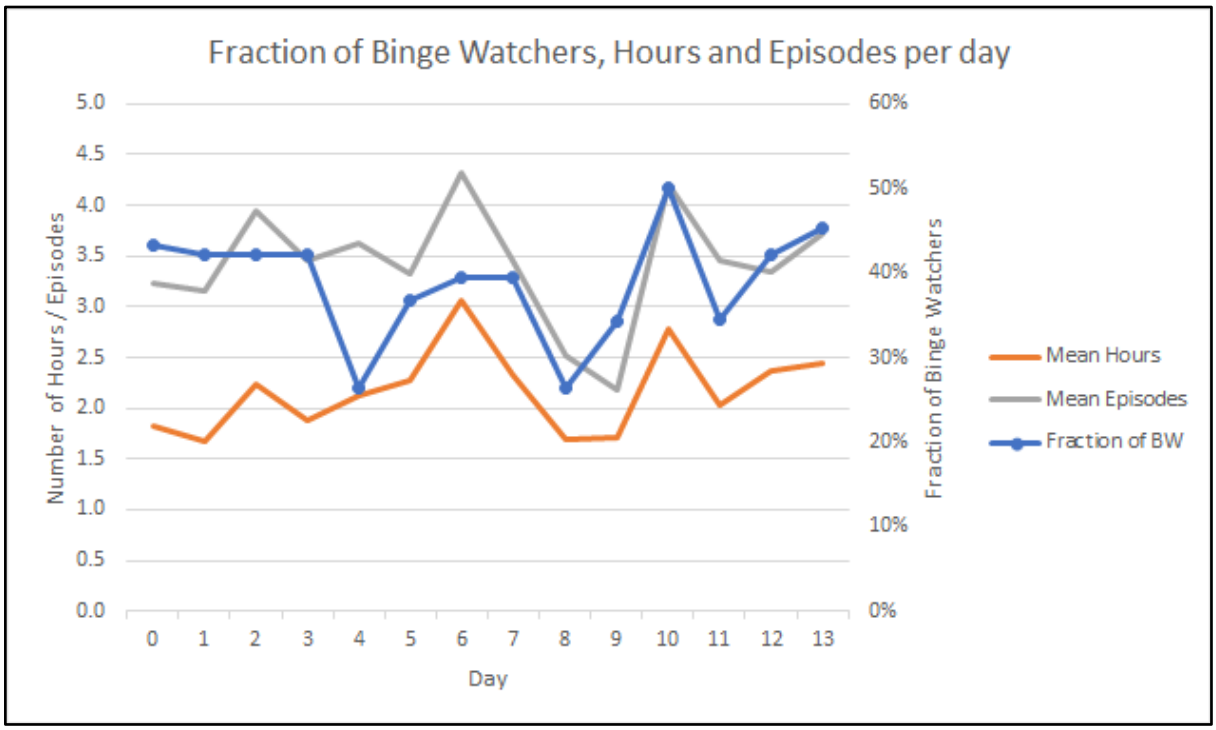


Figure 3. Fraction of participants who were binge watching on a given day

There was considerable variability in the mean watching behaviors of the participants. For instance, Participant #25964 watched on average both the most hours and episodes (Figure 4). This respondent spent on average 9.2 hours (SD = 0.77) per day watching streaming platforms and with that averaged at 13.62 (SD = 0.17) episodes daily. Meanwhile, the least watching behaviour was carried out by participant #25836, who only watched about 27 minutes (SD = 0.76) per day and even less than one episode on average a day (0.71 episodes, SD = 0.15) a day. Both episodes ($p < .01$) and hours ($p < .01$) were significantly related to the participant variable over time, confirming significant differences between participants, and both were strongly intercorrelated, also when analyzed per participant (Pearson $r = .846$, $p < .001$).

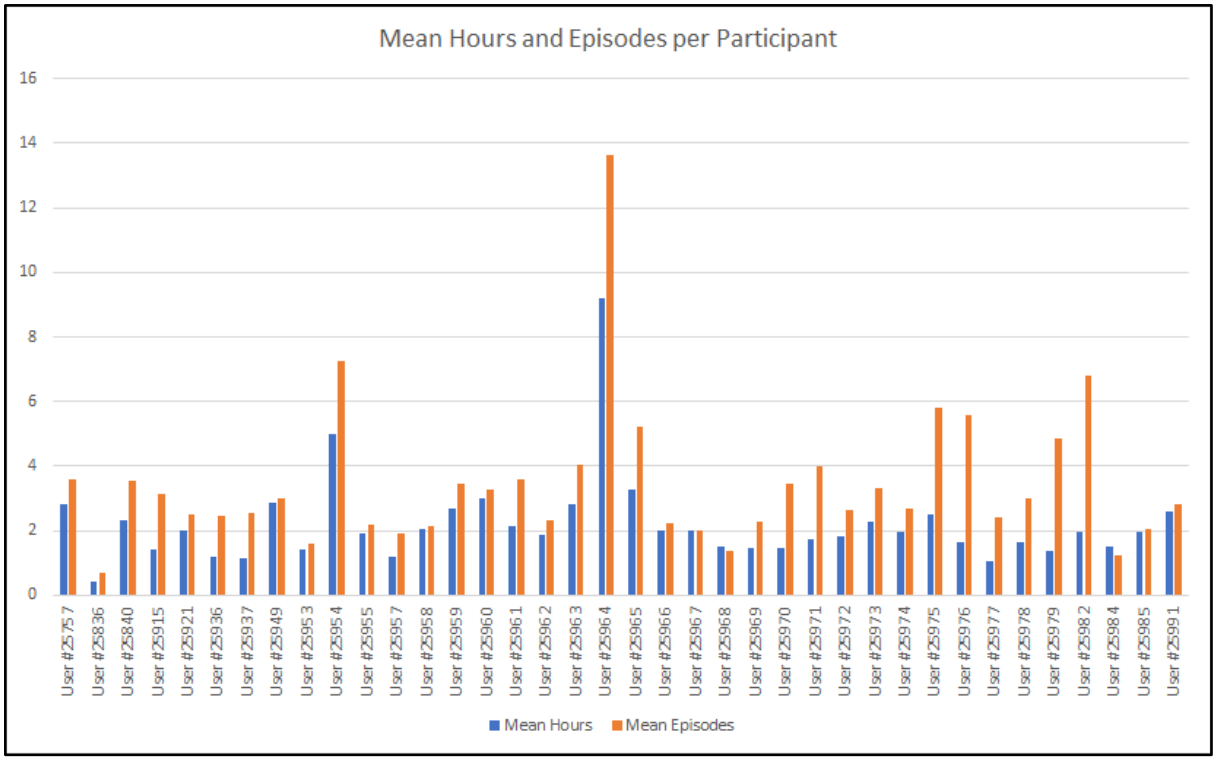


Figure 4. Mean hours and episodes watched in total per participant

Binge watching did show a significant difference (p = .006) between participants. Not only did participant #25836 watch the least number of episodes and hours (Figure 4), this person also did not meet the criteria for binge watching on any occasion (Figure 5). The same is true for participant #25984. Also, the participant with the greatest number of hours watched on average (#25964, Figure 4) engaged in binge watching behaviour on only 43.3% of the days. Meanwhile, participant #25840 met the binge watching criteria 86.6% (SD = 0.17) of the time and thus had the highest overall frequency of binge watching behaviour (Figure 5). Overall, the participants watched on average for 2.2 hours (SD = 1.40), averaging to 3.4 episodes (SD = 2.20) and a mean fraction of binge watching of 39% (SD = 0.24)

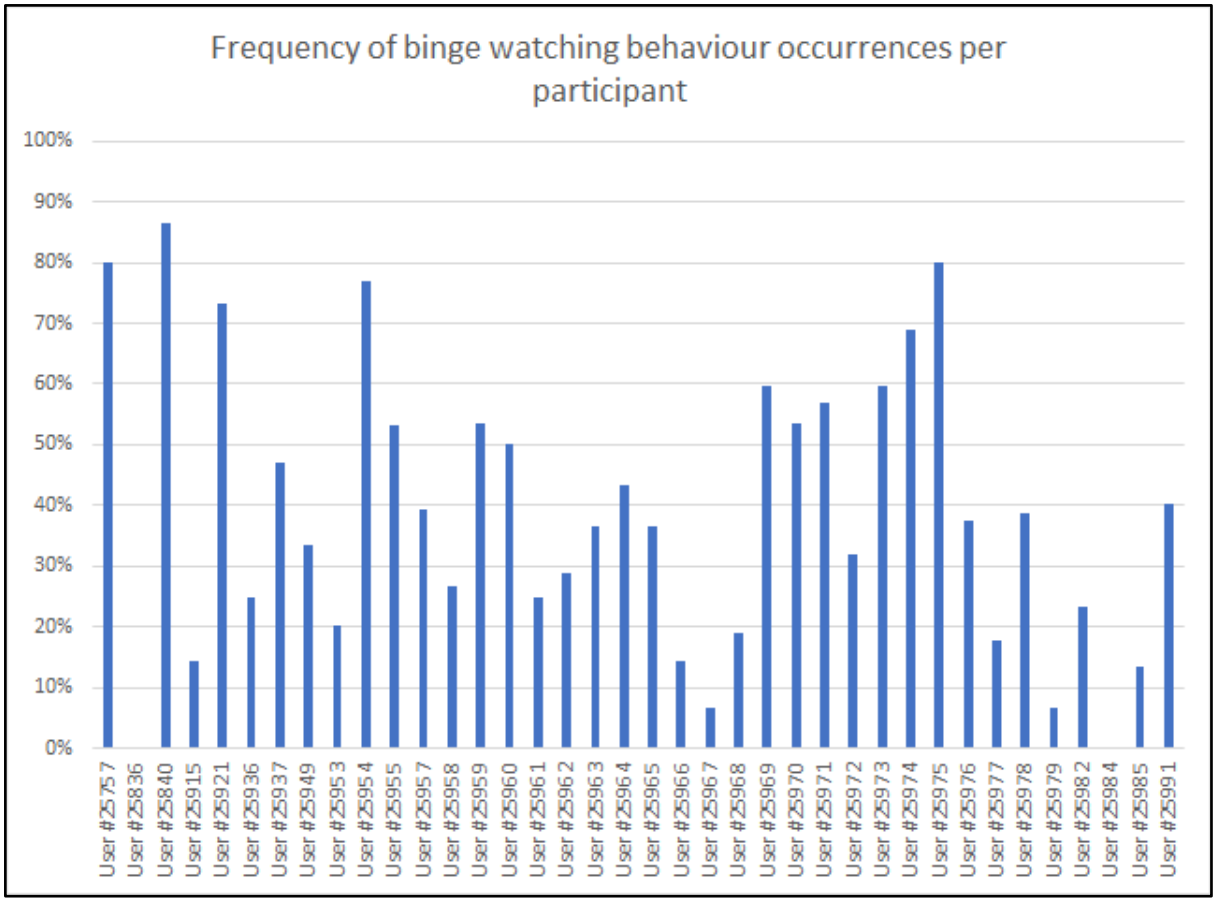


Figure 5. Average frequency of binge watching behaviour occurrences per participant

Mental Health Over Time And Participants

Concerning the measurements of mental health, neither the PHQ-4 depression means ($p = .743$) nor the anxiety means ($p = .557$) were significantly related to the day variable, indicating that at the group level anxiety and depression were fairly stable over time. The highest daily average value for depression reached 1.21 (SD = 0.20) out of 6 and was measured on day 5, a Monday (Figure 6). Moreover, highest anxiety scores were measured on day 3 (Saturday), with a mean of 0.95 (SD = 0.18) out of 5. Lowest scores for depression were found on day 2 (Friday) with 0.70 out of 6. Meanwhile, the lowest scores of anxiety showed on day 9 (Friday) with 0.48 out of 5 (Figure 6). In terms of correlation between the estimated marginal means, anxiety and depression means show a strong correlation over time points (Pearson $r = .513$), although this did not reach statistical significance ($p = .061$).

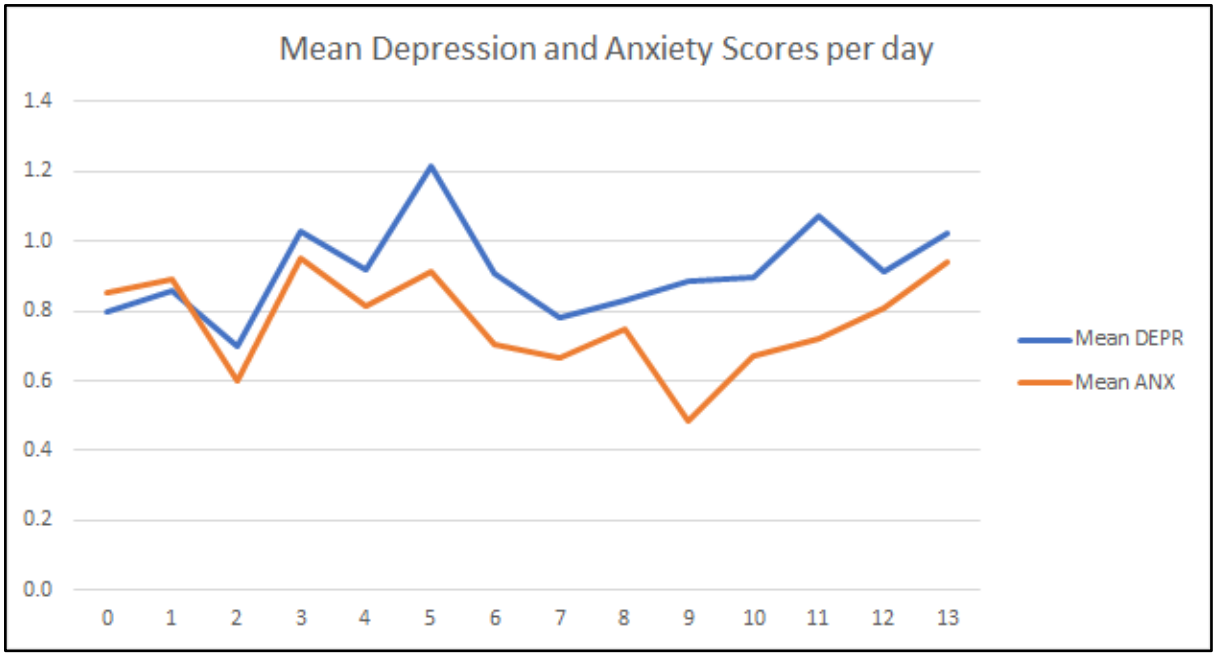


Figure 6. Mean depression and anxiety scores per day

Next, the mental health measurements are analyzed from a participant perspective. Both depression ($p < .001$) and anxiety means ($p < .001$) were significantly related to the participant variable. Moreover, estimated mean depression and anxiety scores showed a strong intercorrelation over participants (Pearson $r = .740$) that was significant ($p < .001$). Participant #25957 had both the highest depression scores, averaging 3.15 (SD = 0.33) overall, and the highest anxiety scores with a mean score of 2.42 (SD = 0.26) (Figure 7). A total of three participants reported depression scores that average to zero (Figure 7). The same was true for the anxiety scores, where three different participants also averaged at zero (Figure 7). On average, participants scored 0.92 (SD = 0.83) out of 6 on the depression scale and 0.76 (SD = 0.71) out of 5 in the anxiety scale.

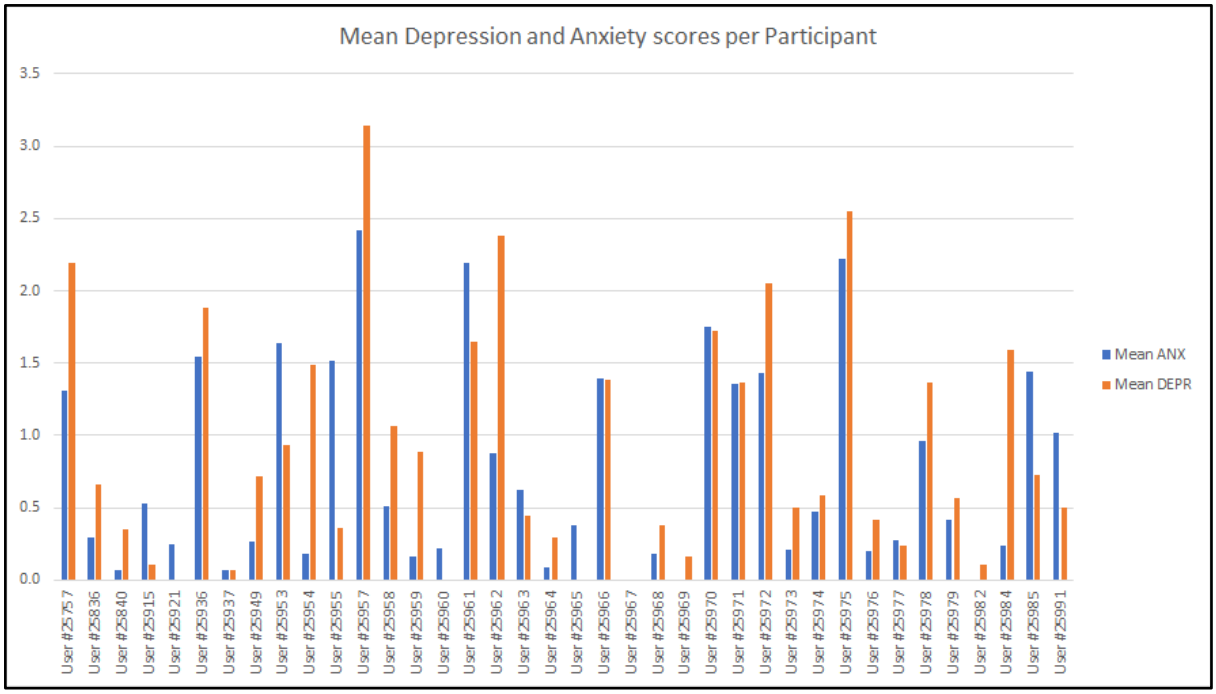


Figure 7. Anxiety and depression scores per participant

Associations Between Depression And Binge Watching

Linear Mixed Models over all time points were used to test the associations between mental health (depression and anxiety) and binge watching. As table 2 and Figures 8 and 9 illustrate, there was no significant connection between any of the three measures of binge watching behaviour and participant's depression scores.

Table 2. Linear mixed model outcome for depression scores with watching behaviours as covariates

Measurement of Binge-Watching	B-Estimate (p-value)
Binge Watching (yes vs. no)	0.041 (.822)
Hours watched	0.057 (.518)
Episodes watched	-0.029 (.525)

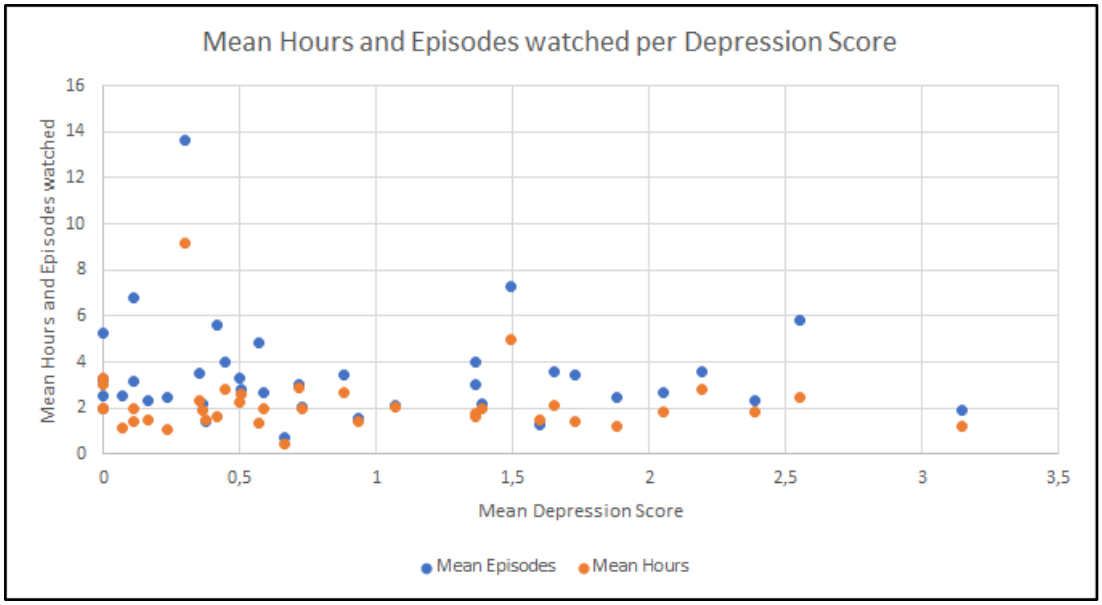


Figure 8. Mean hours and episodes per depression score

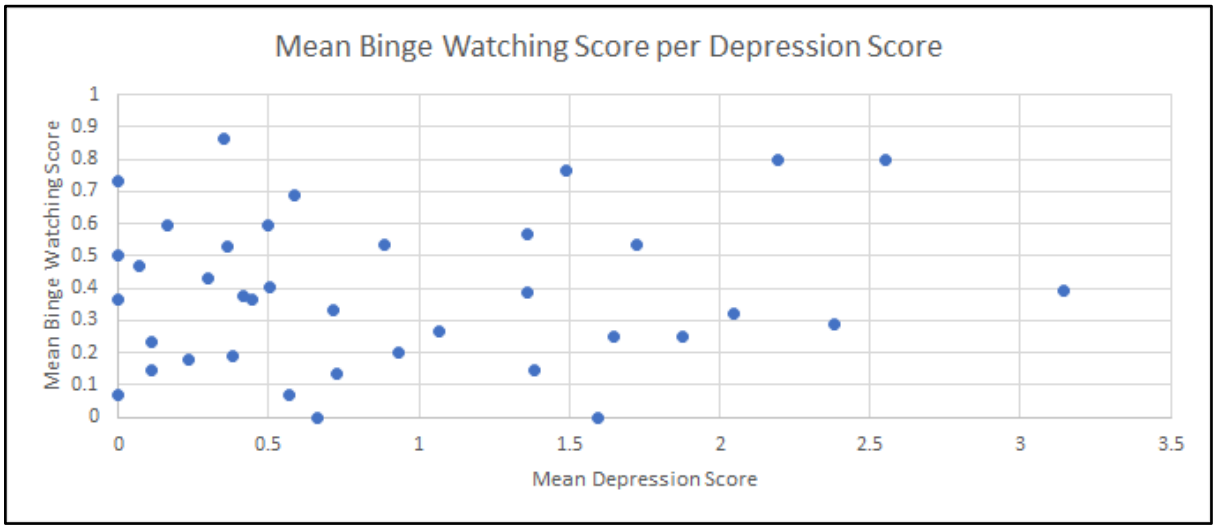


Figure 9. Mean binge watching score per depression score

Next, the baseline and final assessment were considered as potential predictors of watching behavior. Those were meant to serve as additional sources of insight into the connection between trait-like depression scores and binge watching behaviour. As can be seen in table 3, there were no significant connections between the two additional trait measurements of depression and any of the three binge watching variables.

Table 3. Linear mixed model analysis of depression with baseline and final assessment means as covariates

Measurement of Binge-Watching	B-estimate (p-value)	Covariate
Binge Watching (yes vs. no)	0.01 (.559)	Baseline Depression
Hours watched	-0.03 (.534)	Baseline Depression
Episodes watched	-0.02 (.852)	Baseline Depression
Binge Watching (yes vs. no)	0.02 (.172)	Final Depression
Hours watched	-0.05 (.412)	Final Depression
Episodes watched	-0.09 (.436)	Final Depression

Similarly, the analysis of simple correlation between the means showed no significant connections (Table 4). Since the normality assumption for the Pearson correlation was not met, instead the nonparametric Spearman correlation was used.

Table 4. Spearman's rho correlations for baseline and final assessment depression scores

Measurement of Binge-Watching	Spearman's rho with Baseline Depression Scores (p-value)	Spearman's rho with Final Depression Scores (p-value)
Binge Watching (yes vs. no)	0.09 (.591)	0.15 (.427)
Hours watched	-0.04 (.800)	-0.07 (.710)
Episodes watched	0.07 (.706)	-0.08 (.676)

As the data clearly shows no significant connection between depression scores in the PHQ-4 and binge watching behaviours, both measured as states over the two weeks and the trait retrospective measurement at baseline and the end of the study, the findings consistently showed no significant effect of depressive symptoms on binge watching behaviour.

Associations Between Anxiety And Binge Watching

Parallel to the analysis of depression scores, Linear Mixed Models over all time points were used to compare means across all time points. As table 5 illustrates, further supported by figures 10 and 11, there was no significant linear connection between any of the three

measures of binge watching behaviour and participant's anxiety scores, even though p-values were somewhat smaller than for the depression scores.

Table 5. Linear mixed model outcome for anxiety scores with watching behaviours as covariates

Measurement of Binge-Watching	B-Estimate (p-value)
Binge Watching (yes vs. no)	-0.237 (.130)
Hours watched	-0.068 (.359)
Episodes watched	0.056 (.129)

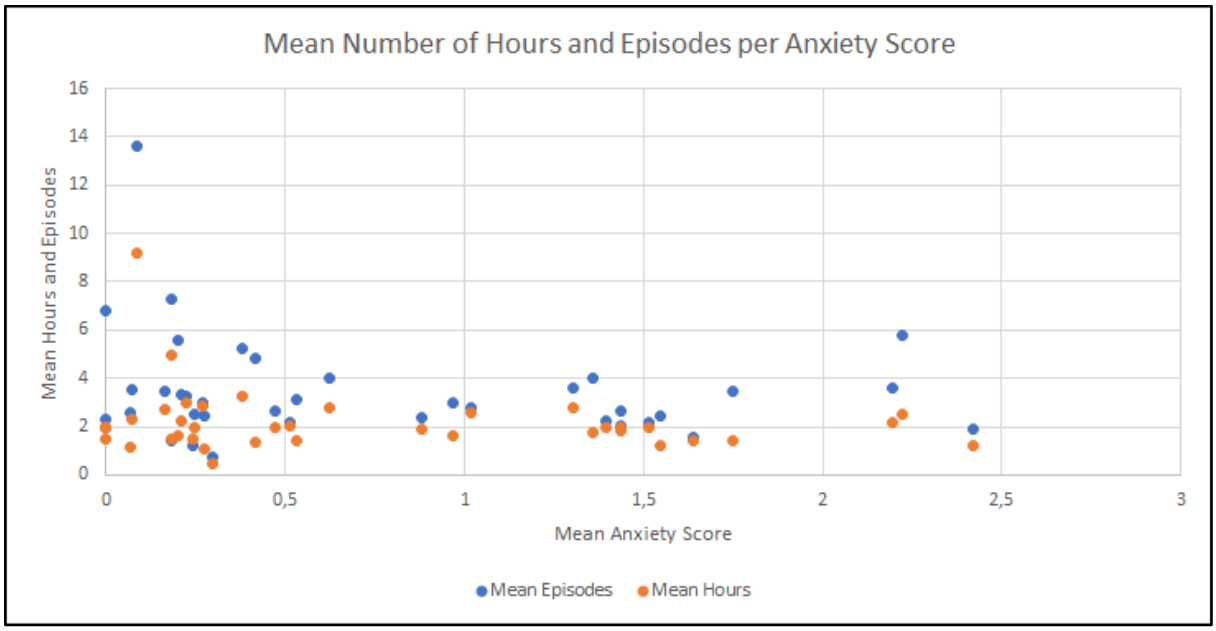


Figure 10. Mean hours and episodes per anxiety score

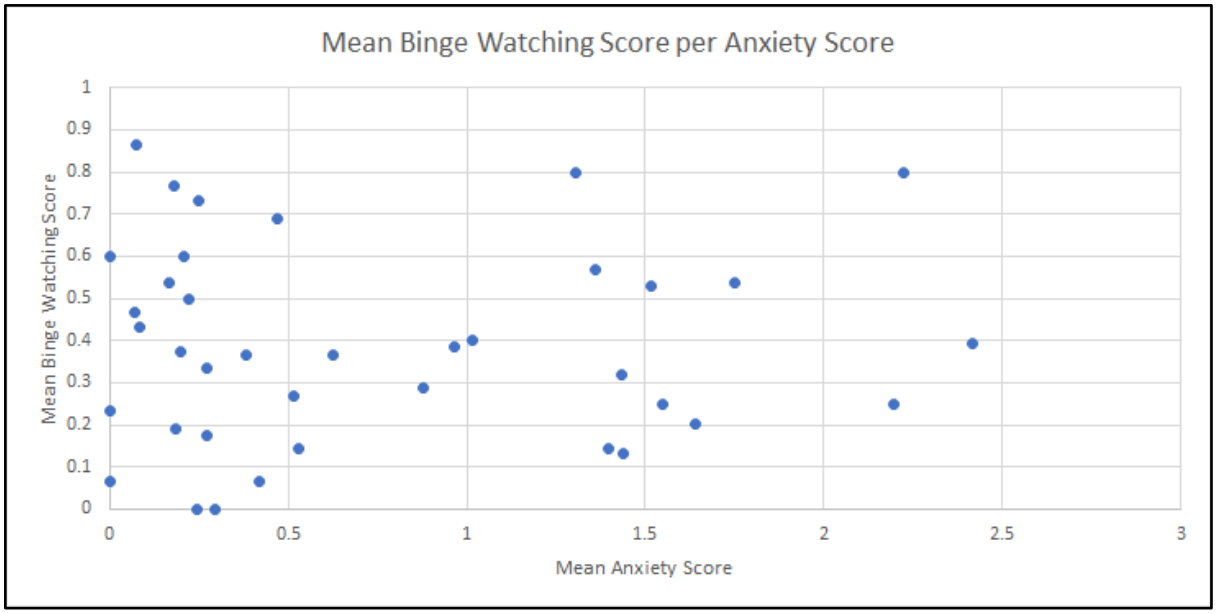


Figure 11. Mean binge watching score per anxiety score

In addition, the baseline and final assessments of trait anxiety were examined. Those were meant to serve as additional sources of insight into the connection between anxiety scores and binge watching behaviour. As the Linear Mixed Model analysis showed, there was only one significant connection between the constructs (Table 6). Binge watching was significantly positively associated with the mean anxiety scores of participants in the final assessment ($p = 0.049$). So, for every percentage increase in binge watching the participant's final anxiety score increased by 0.02 out of 7 (Table 6).

Table 6. Linear mixed model analysis of anxiety with baseline and final assessment means as covariates

Measurement of Binge-Watching	B-estimate (p-value)	Covariate
Binge Watching (yes vs. no)	0.02 (.517)	Baseline Anxiety
Hours watched	0.02 (.083)	Baseline Anxiety
Episodes watched	-0.09 (.644)	Baseline Anxiety
Binge Watching (yes vs. no)	0.02 (.049)	Final Anxiety
Hours watched	-0.05 (.388)	Final Anxiety

Episodes watched	-0.09 (.419)	Final Anxiety

The analysis of correlation between the means showed no significant connections (Table 7).

Table 7. Spearman's rho correlations for baseline and final assessment anxiety scores

Measurement of Binge-Watching	Spearman's rho with Baseline Anxiety Scores (p-value)	Spearman's rho with Final Anxiety Scores (p-value)
Binge Watching (yes vs. no)	0.22 (.215)	0.19 (.322)
Hours watched	-0.05 (.761)	-0.19 (.311)
Episodes watched	0.02 (.925)	-0.09 (.653)

The data clearly shows no consistent significant connection between anxiety scores in the PHQ-4 and binge watching behaviours, both measured over the two weeks and the retrospective measurement at baseline and the end of the study. The findings consistently showed no significant effect of anxiety symptoms on binge watching behaviour, with the exception of a weak but significant association between binge watching and state anxiety levels in the final assessment.

Discussion

This study aimed to analyze the connection between depression, anxiety and binge watching behaviours over time in a two-week experience sampling method (ESM) study. Overall, this study did not find a statistically significant association between participant's daily levels of state depression and state anxiety on binge watching behaviour on the same day. Yet, one significant finding was identified, as participant's binge watching propensity appeared to be a weak predictor of trait anxiety at the end of the study.

Meaning Of The Results And Connection To Previous Studies

For the purpose of this study, binge watching was operationalized along three different variables. Firstly, the number of hours spent watching video-on-demand platforms, secondly the number of episodes and thirdly a dummy-coded binge watching variable, so watching at least two episodes or one hour of the same show on video-on-demand streaming platforms (Flayelle, 2020; Mikos, 2016; Steiner & Xu, 2018). As it turned out, there was not a single significant relationship between any aspect of daily binge watching behavior and state depression scores, as measured with the PHQ-4. Therefore, the findings indicate that depressive symptoms are no significant predictor of binge watching behaviour. Similar to feeling of depression, there was not a single significant association between any of the constructs for state anxiety. Thus, state anxiety appears to be no significant predictor of binge watching behaviour as well. In conclusion, the overall finding of this report can be summarized to the fact that no significant linear connection was found between daily depression and anxiety levels and participant's binge watching behaviours.

The results found in this study reflect the current state of the research body on binge watching behaviour and mental health, especially depression and anxiety. Except for one finding related to trait anxiety, which will be discussed separately, there was no significant result concerning the connection between mental health and binge watching. This lack of significant results would suggest that there is no connection between the constructs, or at least not in the current sample in the studied the time frame. Conflicting results have been found previously by other studies. Tefertiller and Maxwell (2018), concluded that there might be a connection between mental health and binge watching, yet their results were statistically insignificant. While Clarke (2019) found that the higher participant's anxiety scores were, the lower were their binge watching scores, Wheeler (2015) identified a significantly positive connection between attachment anxiety and binge watching. Regarding depression and binge watching, there are some studies that found significant results, while others did not. Also,

different studies focused on different directions of associations, so the impact of mental health on binge watching and the other way around, which complicates comparisons. Accordingly, Wheeler (2015) concluded in her study that the connection between depression and binge watching is not conclusively analyzed.

This notion by Wheeler is well illustrated by the fact that nearly all results in the current study turned out to be non-significant, while one analysis yielded a significant outcome. The connection between participant's binge watching propensity and their score in the final anxiety assessment was significant. While this association between binge watching and trait anxiety has rather small strength, it still illustrates that there indeed might be a connection between mental health, especially anxiety, and binge watching behaviour after all. Yet, it has to be noted that this result is possibly caused by chance due to the high number of tests conducted over the course of the study. Moreover, this finding supplies an argument against the criticism for the multitude of different definitions of binge watching including doubts about the very existence of the construct, as it shows that the construct of binge watching really does make a difference, since there were no significant connections between merely the number of hours or episodes watched and mental health. Then again, this study focused on one specific operationalization of binge watching and perhaps changing the number of hours or number of episodes would change this outcome.

The participants showed high variation in watching behaviour across the sample, which is why a closer look at the individual level is necessary. While one participant spent on average 9 hours per day with video-on-demand services, another watched on average less than half an hour daily. Similar observations are possible for binge watching as a separate behaviour. One participant met the criteria a total of 86% of the days, while a few participants never met the criteria even once. Furthermore, the overall average of binge watching frequency across all days is 39%. This fairly high average underlines that binge watching is not only a real behaviour, but also that it is pretty common across the sample, which has been concluded similarly by other studies, as both Steinbach (2018) and Flayelle (2017) found comparable results. As this further underlines the high prevalence of binge watching, perhaps more research into the definition of binge watching is necessary. As was discussed in the introduction, currently many different definitions of binge watching exist, mostly with different cut-off scores for hours or episodes. Given that binge watching is already such a prevalent behaviour, its negative connotation should be debated. The behaviour appears to be quite common and many studies find no significant connection to mental health, so perhaps watching one hour or two episodes is not the correct way to define binge watching as a

scientific construct. Future studies should therefore raise the cut-off scores both for hours and episodes, given that the present study and many previous studies found it to be such a normal behaviour without implications for participant's state of mental health. Another possibility is that the consequences of binge watching for one's mental health are less short-term in nature, which calls for a more longitudinal study of the long-term consequences.

Regarding the analysis of the outcomes of mental health tests in the sample, on average respondents reported both fairly low anxiety and depression scores. However, the results showed high variability across participants regarding mental health states, yet fairly stable values within participants. These rather low scores are surprising, as a very recent study showed that the corona situation yielded an increase in both anxiety and depression scores across the population compared to prior measurements (Pieh et. al., 2020). That study also used the PHQ-9 and GAD-7 to investigate current levels of anxiety and depression in a representative sample, with much higher results. The average depression score was 6.2 out of 9, the average anxiety score 5.8 out of 7 (Pieh et. al., 2020). This period of time might have been special in terms of both binge watching behaviour and mental health, caused by social distancing and lockdowns. Yet apparently, the impact on the mental health of the current sample of this study seems rather small. This might be due to the characteristics of the sample. Given that most of the participants are university students, they may have a much more flexible schedule and less likely financial worries, compared to those that lost their occupation due to the virus.

Limitations And Directions For Future Research

Since this study used Experience Sampling, a rather novel methodology of in the field of binge watching behaviour, a couple of limitations caused in part by that approach should be noted. As the study took place over a period of more than two weeks, a number of participants understandably lacked the motivation to complete the final assessment questionnaires (PHQ-9 and GAD-7), which leads to a decrease in statistical power. Furthermore, as some previous ESM studies have identified, there is the possible problem of measurement reactivity (Simpson et. al., 2005; Steinbach, 2018). This concept refers to the impact of the measurement on the actual behaviour of the participants. As they were asked to note down their daily watching behaviour, respondents might have become self-aware and changed their usual watching behaviour, possibly watching less to reflect socially more acceptable behaviour. Moreover, there are some obstacles that can occur when one uses experience sampling to study any kind of behaviour. ESM can be very labor-intensive for

both the participants of the study and the researchers themselves (Xie, Heddy & Greene, 2019). For the researchers it can be very time-consuming to program the survey packages and notification schedule. Also, simple mistakes in the programming of the questionnaires or the signaling schedule can hamper the process of data collection.

This study used a convenience sample of friends and family of the researchers, which inherently has a few downsides. Given that the participants were all known to the researchers, they might have been particularly motivated to display socially desirable behaviour, compared to their usual watching behaviour that is not monitored. Also, as most members of the sample were from the direct social environment of the researchers, it can be inferred that they are likely to stem from a similar background. So, the sample may not be very representative of the age group in general. Furthermore, given that most of the participants are university students with very flexible schedules, further research should look at implications of binge watching on people with more rigid schedules and social roles and responsibilities, as this possibly impacts watching behaviour. Moreover, those individuals of the general age group with potentially less self-control could potentially be more at risk, which should be also be investigated in further research. Also, as this study produced a small, yet significant result on the connection between binge watching and trait anxiety, further research should be conducted into that association. Moreover, considering that the coronavirus might have impacted the data and subsequent results of this study, since people might have watched a lot more than they usually would have, there is a need to repeat similar data collection and analyses without disturbance by the lockdown and social distancing. It might also be of interest to analyze how the development of the virus and the resulting measures impacted the watching behaviour of the population. In general, more studies that investigate binge watching could use ESM instead of the more common cross-sectional study design.

Conclusion

As this study did not identify any significant connection between binge watching and mental health, perhaps future studies should lay a different focus. Binge watching is already a very common behaviour, yet due to the many different definitions of it, comparisons are difficult. Perhaps the term ‘binge’ watching is already misleading, as this behaviour does not show the clear negative consequences of other binging behaviours. Consequent research should find a way to unite the definitions, so that future studies can better identify possible mental health implications.

References

BBC News. (2020). Netflix gets 16 million new sign-ups thanks to lockdown. Retrieved from https://www.bbc.com/news/business-52376022

Clarke, K. L. (2019). Multivariate Relationships of Binge Watching-Drinking-Eating With Depression, Anxiety, and Stress in College Students. *Walden Dissertations and Doctoral Studies*. 6883. Retrieved from: https://scholarworks.waldenu.edu/dissertations/6883

Conner, T. S., & Lehman, B. J. (2012). Getting started: Launching a study in daily life.

Ethica Data Services Inc (2019). About. Retrieved from https://ethicadata.com/about

Flayelle, M., Maurage, P., Di Lorenzo, K. R., Vögele, C., Gainsbury, S. M., & Billieux, J. (2020). Binge-watching: What do we know so far? A first systematic review of the evidence. *Current Addiction Reports, 7*(1), 44-60.

Goldberg, D. (2008). Towards DSM-V: the relationship between generalized anxiety disorder and major depressive episode. *Psychological Medicine, 38*(11), 1671-1675.

Larson, R., & Csikszentmihalyi, M. (2014). The experience sampling method. In *Flow and the foundations of positive psychology* (pp. 21-34). Springer, Dordrecht.

Löwe, B., Wahl, I., Rose, M., Spitzer, C., Glaesmer, H., Wingenfeld, K., Schneider, A. & Brähler, E. (2010). A 4-item measure of depression and anxiety: validation and standardization of the Patient Health Questionnaire-4 (PHQ-4) in the general population. Journal of affective disorders, 122(1-2), 86-95.

Mikos, L. (2016). Digital media platforms and the use of TV content: Binge watching and video-on-demand in Germany. *Media and Communication, 4*(3), 154-161.

Mittell, J. (2015). *Complex TV: The poetics of contemporary television storytelling*. NYU Press.

Pieh, C., Budimir, S., & Probst, T. (2020). The effect of age, gender, income, work, and physical activity on mental health during coronavirus disease (COVID-19) lockdown in Austria. *Journal of Psychosomatic Research*, 110186.

Pinel, J. P. J., Barnes, S. J. (2014). Biopsychology of psychiatric disorders. In: J. P. J. Pinel & S. J. Barnes (Eds.), Introduction to biopsychology (pp.470-494). England: Pearson. ISBN 9781292058917

Pittman, M., & Sheehan, K. (2015). Sprinting a media marathon: Uses and gratifications of binge-watching television through Netflix. First Monday, 20(10). Retrieved from: https://firstmonday.org/article/view/6138/4999

Schumacher, E. (2020) Coronavirus: What are Germany's updated lockdown measures?. Retrieved from https://www.dw.com/en/coronavirus-what-are-germanys-updated-lockdown-measures/a-53139313

Stanhope, J. (2016). Patient Health Questionnaire-4. Occupational Medicine, 66(9), 760-761.

Stavrakaki, C., & Vargo, B. (1986). The relationship of anxiety and depression: a review of the literature. *The British Journal of Psychiatry*, *149*(1), 7-16.

Steinbach, E. M. (2018). Binge-Watching and its Impact on Learning Behaviour and Important Daily Life Activities among University Students: A Study using Ecological Momentary Assessment (Bachelor's thesis, University of Twente).

Steiner, E., & Xu, K. (2020). Binge-watching motivates change: Uses and gratifications of streaming video viewers challenge traditional TV research. *Convergence*, *26*(1), 82-101.

Tefertiller, A. C., & Maxwell, L. C. (2018). Depression, emotional states, and the experience of binge-watching narrative television. Atlantic Journal of Communication, 26(5), 278-290.

Tukachinsky R. & Eyal K. (2018) The Psychology of Marathon Television Viewing: Antecedents and Viewer Involvement, Mass Communication and Society, 21:3, 275-295, doi: 10.1080/15205436.2017.1422765

Watson, A. (2020). Netflix subscribers count in the U.S. Retrieved from https://www.statista.com/statistics/250937/quarterly-number-of-netflix-streaming-subscribers-in-the-us/

Wheeler, K. S. (2015). The relationships between television viewing behaviors, attachment, loneliness, depression, and psychological well-being.

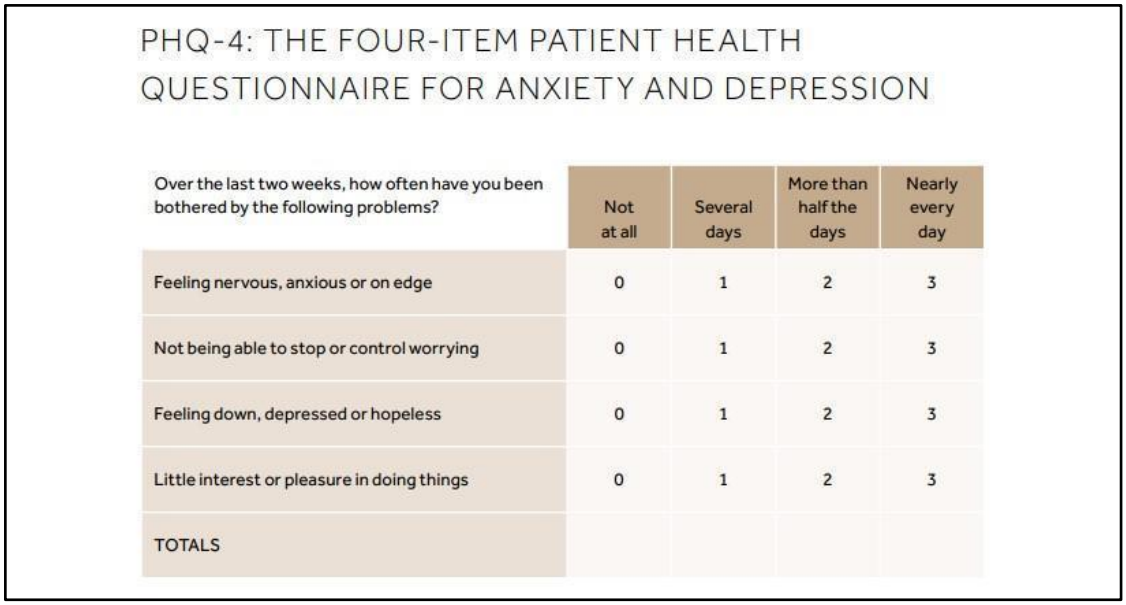
Wired. (2020). 53 of the best Netflix series to binge watch. Retrieved from https://www.wired.co.uk/article/best-shows-netflix

Xie, K., Heddy, B. C., & Greene, B. A. (2019). Affordances of using mobile technology to support experience-sampling method in examining college students & engagement. Computers & Education, 128, 183-198.

Appendix

Table 1. Participant characteristics

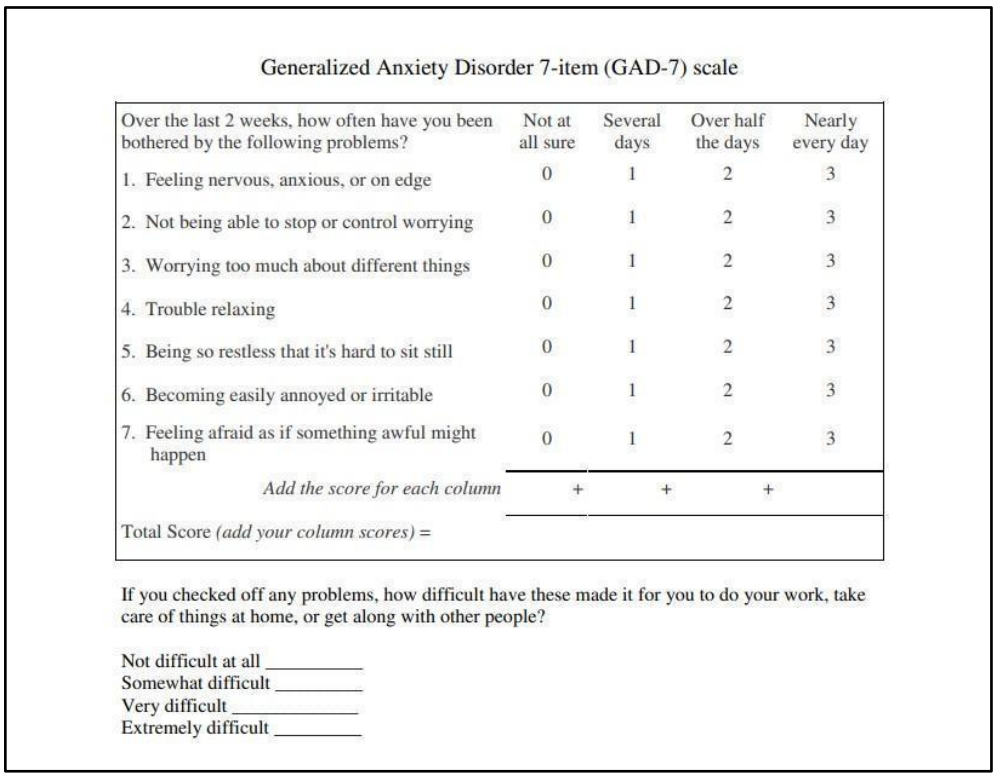
Characteristic	Total Number	Percentage (%)
Gender		
Male	21	55.3
Female	17	44.7
Age (years)		
Mean (SD)	23.79 (5.3)	
18 - 24	31	81.6
27 - 51	7	18.4
Nationality		
German	35	92.1
Dutch	1	2.6
Other European	2	5.3
Occupation		
Students	22	57.9
Employed full-time	9	23.7
Employed part-time	1	2.6
Apprentice	3	7.9
Pupil	1	2.6
Other	2	5.6



PHQ-4: THE FOUR-ITEM PATIENT HEALTH QUESTIONNAIRE FOR ANXIETY AND DEPRESSION

Over the last two weeks, how often have you been bothered by the following problems?	Not at all	Several days	More than half the days	Nearly every day
Feeling nervous, anxious or on edge	0	1	2	3
Not being able to stop or control worrying	0	1	2	3
Feeling down, depressed or hopeless	0	1	2	3
Little interest or pleasure in doing things	0	1	2	3
TOTALS				

Figure 1. PHQ-4



Generalized Anxiety Disorder 7-item (GAD-7) scale

Over the last 2 weeks, how often have you been bothered by the following problems?	Not at all sure	Several days	Over half the days	Nearly every day
1. Feeling nervous, anxious, or on edge	0	1	2	3
2. Not being able to stop or control worrying	0	1	2	3
3. Worrying too much about different things	0	1	2	3
4. Trouble relaxing	0	1	2	3
5. Being so restless that it's hard to sit still	0	1	2	3
6. Becoming easily annoyed or irritable	0	1	2	3
7. Feeling afraid as if something awful might happen	0	1	2	3
Add the score for each column	+	+	+	
Total Score *(add your column scores)* =				

If you checked off any problems, how difficult have these made it for you to do your work, take care of things at home, or get along with other people?

Not difficult at all _____
Somewhat difficult _____
Very difficult _____
Extremely difficult _____

Figure 2. GAD-7

PATIENT HEALTH QUESTIONNAIRE (PHQ-9)

NAME:_____ **DATE:**_____

Over the last *2 weeks,* how often have you been bothered by any of the following problems?
(use "✓" to indicate your answer)

	Not at all	Several days	More than half the days	Nearly every day
1. Little interest or pleasure in doing things	0	1	2	3
2. Feeling down, depressed, or hopeless	0	1	2	3
3. Trouble falling or staying asleep, or sleeping too much	0	1	2	3
4. Feeling tired or having little energy	0	1	2	3
5. Poor appetite or overeating	0	1	2	3
6. Feeling bad about yourself—or that you are a failure or have let yourself or your family down	0	1	2	3
7. Trouble concentrating on things, such as reading the newspaper or watching television	0	1	2	3
8. Moving or speaking so slowly that other people could have noticed. Or the opposite — being so figety or restless that you have been moving around a lot more than usual	0	1	2	3
9. Thoughts that you would be better off dead, or of hurting yourself	0	1	2	3
	add columns		+	+

(Healthcare professional: For interpretation of TOTAL, please refer to accompanying scoring card). TOTAL:

10. If you checked off *any problems,* how *difficult* have these problems made it for you to do your work, take care of things at home, or get along with other people?

- Not difficult at all _____
- Somewhat difficult _____
- Very difficult _____
- Extremely difficult _____

Figure 3. PHQ-9