

The Relationship between Instagram Engagement and Symptoms of Depression in young adults

Daniel Barth

Faculty of Behavioural, Management, and Social Sciences (BMS)

Department of Psychology, University of Twente

202000384: BSc Thesis Psychology

First supervisor: Martha Kreuzberg

Second supervisor: Dr. J. Piano Simones

July 3rd, 2023

Abstract

Social networking sites (SNS) have become increasingly popular among young adults around the world. While they are beneficial for communicating and connecting people, they have also been found to have negative effects. Based on this, this study aimed to investigate the relationship between Instagram engagement and symptoms of depression in young adults. Furthermore, it was hypothesized that the biological sex of an individual moderates this relationship. The cross-sectional study was carried out via an online questionnaire with a total of 68 participants between the ages of 18-24. Most of them were from Germany and a few were from the Netherlands. Overall, they were evaluated as low to medium users of SNS. When it comes to depression, it was found that the participants were on average mildly depressed. The finding suggested a significant positive relationship between Instagram engagement and symptoms of depression ($p = .001$). However, no significant moderation effect was found with the variable of sex ($p = .40$). In summary, there is a significant and positive relationship between Instagram engagement and symptoms of depression. However, sex was not found to be a moderator of the relationship. Nonetheless, there might still be other factors that might influence this relationship, such as the content which is consumed. Due to this, future research should focus on getting a deeper understanding of what influences this relationship.

Keywords: Instagram, Young adults, Symptoms of Depression, Sex, Cross-sectional Survey, Instagram Engagement

The Relationship between Instagram Engagement and Symptoms of Depression in young adults

Social networking sites (SNS) have become an increasingly important part of our daily lives. Their ability to connect people around the globe and enable them to share their stories, pictures, and videos, reformed the way people interact with each other in today's society. With a total of 1,33 billion active users in 2023, Instagram is one of the most used social media platforms worldwide (Barnhart, 2023). Instagram's ability to easily share pictures, videos, and audio messages paired with its broad availability on smartphones, pcs, and tablets make it an interesting platform for the younger generation (Barnhart, 2023). Especially young adults between the ages of 18-24 use Instagram, with 31% of Instagrams users belonging to that age range. This is simultaneously the largest group of users on Instagram (Barnhart, 2023). This differentiates Instagram from the biggest SNS, Facebook, where only 20% of all users are aged between 18-24 (Barnhart, 2023; We Are Social, & DataReportal, & Meltwater, 2023). However, especially young adults run a particular risk of experiencing negative drawbacks from social media use (Ivie et al. 2020). A common example of such a drawback is symptoms of depression. It is important to note that there is no absolute classification of these symptoms and that they may vary to a certain degree between different individuals. Some people try to adjust their unsatisfactory mood by further engaging in social media use (Woods & Scott, 2016).

Instagram

When Instagram launched in 2010, 25,000 users signed up on the first day. Since then, Instagram has managed to establish itself as one of the most popular SNS (Instagram, n. D.; Barnhart, 2023). It was initially designed to be a photo-sharing website which then developed into a SNS that enables its users to share multiple types of content. Among these types of content are photos, videos, and stories. In addition to this, Instagram also lets you like, comment, and save other people's content as well (Blystone, 2022). These functions provide a novel way of communicating, which is particularly appealing to young adults (Huang & Su, 2018). Due to this, the biggest group of Instagram users are between 18-24 years old, which make up a total of 31% of all users (Barnhart, 2023). One of the reasons why Instagram may appeal to young adults is the fact that it allows people to connect with their friends and family all over the world (Green & Martinez, 2018). Furthermore, Instagram enables an individual to quickly access and spread information to other users, so that they can stay up to date on recent events (Vázquez-Herrero et al., 2019). In general, Instagram has become a multidimensional social media platform, with multiple functions that are especially popular among young adults (Huang & Su, 2018).

However, according to research by Amedie (2015), social media is causing a reduction in social interaction and is used as a preferred way to communicate, over meeting in real life, since SNS enable an individual to communicate with others, at any time. This may appear in the form of choosing new online connections to random people over meeting members of your family in real life (Shahjahan & Chisty, 2014). In addition to this, there are indications that there is a relationship between Instagram and symptoms of depression (Adeyanju et al., 2021).

Depression

Depression is a common mental health disorder that is prevalent throughout the world. Over the past 10 years, rates of depression in young adults have increased significantly, with a prevalence of 1-3% in young adults and children (Thapar et al., 2022). Young people are especially at risk of the influence of depression, as they are still developing in multiple areas of their lives, which makes them more vulnerable to the influences of depression (Thapar et al., 2022). In general, it is hard to precisely pinpoint what depression is exactly since no definition successfully encompasses all of its various facets. Nonetheless, the World Health Organization (2023) describes depression as a common mental health disorder that manifests itself as experiencing reduced enjoyment of certain things over longer periods of time. In addition to this, an individual's general mood is often negatively affected as well (World Health Organization, 2023). The DSM-IV provides a framework for depression with certain symptoms. These symptoms include but are not limited to: "Significant weight loss (or poor appetite) or weight gain" and "Insomnia and hypersomnia" (O'Connor et al., 2009). However, it is important to note that symptoms of depression can appear in a variety of ways and are not limited to the above-mentioned symptoms. When it comes to the prevalence of depression, it is estimated that around 3.8% of the world's population experience depression. Additionally, depression appears to be around 50% more common in women compared to men (World Health Organization, 2023).

Instagram and depression

When looking at the relationship between using Instagram and symptoms of depression, research suggests that it is a rather complex one. This is due to the fact that some findings indicate a potential positive relationship between engaging in SNS and mental well-being, while others indicate a negative influence on mental well-being (Baker & Algorta, 2016). However, most of the examined sources indicate a positive relationship between Instagram engagement and symptoms of depression. According to research by Ivie et al. (2020), actively using SNS such as Instagram is associated with symptoms of depression. It is

important to note, that it is rather the actual use of SNS that is associated with symptoms of depression, than using any sort of digital device in general. Furthermore, the meta-analysis conducted by Ivie et al. (2020) points out that there is a small positive relationship between engagement on Instagram and symptoms of depression. One reason why there is a rise in depression among young adults may be due to the increased engagement of younger people on SNS (Ivie et al., 2020). A study by Lin et al. (2016) confirms the previous findings, which discovered a strong as well as significant relationship between engagement in SNS and depression. There are multiple possibilities as to why engagement on Instagram may lead to symptoms of depression. One reason is that the act of observing new content from people, who are perceived to have more positive qualities, is linked to increased depressive symptoms (Hwang, 2019). Another reason is, that increased Instagram usage may promote comparison to other users, which in turn is partially linked to symptoms of depression (Hwang, 2019). Thus, increasingly using Instagram and comparing oneself to people who are perceived to be better off, promotes the emergence of symptoms of depression. (Hwang, 2019)

Biological sex

Another important factor in the relationship between engagement on Instagram and symptoms of depression is the sex of an individual. While both sexes may become depressed, women are up to 50% more likely to become depressed compared to men (World Health Organization, 2023). The underlying reasons as to why this is the case are multifaceted since the relationship between sex and depression is rather complex. Nonetheless, research by Noble (2005) identified that women's heightened risk of depression might be due to underlying genetic issues such as hormonal instability. This instability encourages changes in mood, which in turn may foster the occurrence of depression.

Another reason why the variable of *sex* is included in this research is that there is a common consensus in the social science domain that the biological sex of an individual influences their cognitive abilities as well as social behaviours (Hyde, 2014). An example of such a difference would be that men and women can differ substantially in their development and maturation (Bale & Epperson, 2016). Based on this, it might be beneficial to include the variable of *sex*, to account for any differences between men and women.

Target group

The potential positive relationship between engagement on SNS like Instagram and symptoms of depression may be especially relevant for young adults (Ivie et al., 2020). The reason for this is that young people are in an area of their life, where a lot of changes occur. These changes can manifest themselves as development in the social areas of their lives,

which means that they are more prone to be influenced in social situations, especially when other young adults from the same age group are involved (Ciranka & Van Den Bos, 2019). Furthermore, over the past decade, there was an increase in individuals affected by depression, which are considered young adults (Thapar et al., 2022). This increase in depression may be related to a general increase in social media use (Ivie et al., 2020). The reason why symptoms of depression in young adults can have serious negative consequences is that depression has a strong positive relationship with suicide (Shorey et al., 2022).

Another reason why especially young adults may be vulnerable to the influence of social media use on symptoms of depression is that their social media engagement, in general, is quite high, with around 84% of adults between the ages of 18 and 29 engaging on some sort of SNS (Atske, 2022). This is further emphasized by Instagram's largest user group being between the ages of 18-24, composing 31% of all users (Barnhart, 2023). Based on young adults' increased vulnerability to symptoms of depression potentially as a consequence of Instagram engagement, they have been chosen to be the target group in the present study (Ivie et al., 2020).

The present study

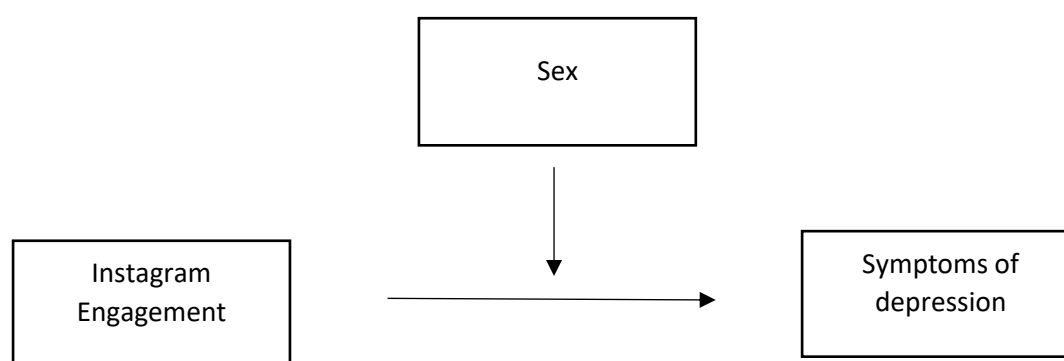
Since it is supposed that engagement of young adults on Instagram is associated with symptoms of depression, this study aims at getting a deeper understanding of this relationship. For this, it is examined whether Instagram engagement influences symptoms of depression in young adults. Instagram is chosen as a SNS for this analysis, as it is a SNS that is primarily navigated by young adults (Barnhart, 2023). While they are the biggest part of Instagram users, they are also especially exposed to symptoms of depression, since they are still developing socially and cognitively (Shorey et al., 2022). In addition to this, women are significantly more at risk of becoming depressed compared to men (World Health Organization, 2023). Based on this, a moderator variable is included in the study to examine whether the sex of an individual influences the relationship between using Instagram and symptoms of depression. To test the influence of the moderator variable, as well as assess the original relationship between Instagram engagement and symptoms of depression, a survey was created and distributed to the participants. There are a total of two research questions (RQs) that are central to this study. The first RQ: *'Is there a positive relationship between Instagram engagement and symptoms of depression in young adults?'* and the second RQ: *"Does the sex of an individual influence the relationship between Instagram engagement and symptoms of depression in young adults?'*. These two RQs result in the following hypotheses:

1. Hypothesis: Young adults with higher rates of Instagram engagement experience more symptoms of depression.
2. Hypothesis: The sex of young adults positively moderates the relationship between their Instagram engagement and symptoms of depression.

The study design is visualized in Figure 1.

Figure 1

Visualization of study design



Methods

Study Design

In order to study the relationship between Instagram engagement and symptoms of depression in young adults and how the variable of sex can moderate this relationship, an online survey has been distributed to the participants. Here, Instagram engagement acts as the independent variable, and symptoms of depression as the dependent.

Participants

This study is based on a sample of 68 participants, all aged between 18-24 years. The only requirements to participate in this study were being over 18 years of age and having and using an Instagram account. Of the participants, 66% are women (45) and 34% are male (23). Furthermore, 82% (56) indicated German as their nationality, 6% (4) stated that they are Dutch and 12% (8) declared that they had a different nationality than German or Dutch. Each participant gave their consent to the data collection at the beginning of the study.

All of the participants were gathered by using convenience sampling or snowball sampling, via the University of Twente platform “SONA”, social media platform posts, or via personal contacts. Additionally, all participants were over the age of 18. The SONA platform allows participants from the University of Twente to gather mandatory research credits by participating in our study. For this study, they received 0.25 SONA credits.

Materials

In this study, the online platform Qualtrics was used to gather the data. Here, a questionnaire was created which included general questions about the demographical data of the participants as well as different scales, measuring Instagram engagement as well as symptoms of depression. To successfully participate in the study, participants needed a laptop, a smartphone, or any other technical device that allows them open and complete the questionnaire. Furthermore, they needed a stable internet connection since this study was entirely online.

Instagram engagement

To measure how much participants are engaging in Instagram; the social networking time use scale (SONTUs) was implemented into the survey. It has five subscales and consists of a total of 29 items. The items represent different scenarios, in which participants are asked to indicate how often they used a certain social media site at that moment, over the past week. That way, the overall engagement on Instagram of a participant is assessed. An example of an item is: “How often do you always use social networking sites like Instagram during the past week when you are relaxing?”. The participant then indicates on a Likert scale from 1-11 how frequently they used a social media site in that scenario. A value of 1 in this case means ‘not applicable to me during the last week’ and a value of 11 indicates ‘I used it more than three times but spent more than 30 minutes each time’ (Olufadi, 2016). To calculate the scores for the individual participants, their scores are coded. The code for this is 1-3 = 1, 4-6 = 2, 7-9 = 3, and 10-11 = 4. Afterwards, the items of the five underlying subscales are scored and are then added together to create the final ‘global’ score. (Olufadi, 2016). A global score of 5-9 indicates that an individual is a low Instagram user. A score between 10-14 as an average user, a score from 15-19 as a high user, and a score above 19 as an extremely high user of Instagram.

The results from Olufadi (2016) suggest that the Sontus has a very high reliability ($\alpha = 0.92$) for the entire scale. The alpha values for the sub-scales ranged from .83 to .91 and are visualized in Table 1. Additionally, the scale possesses a moderately positive convergent validity, with a value between .26 - .43, which is an indication of construct validity.

Table 1*Cronbach's alpha values*

<i>Subscale</i>	<i>Cronbach's alpha</i>
1	.91
2	.89
3	.85
4	.86
5	.83

Symptoms of depression

To measure symptoms of depression, the patient health questionnaire (PHQ-9) was used. It is a self-reported scale that consists of a total of nine items that measure symptoms of depression over the past two weeks. Each of the items is measured on a scale from 0-3, with 0 meaning “*Not at all*” and 3 meaning “*Nearly every day*”. An example of an item from the scale is: “Over the last 2 weeks, how often have you been bothered by feeling tired or having little energy” (Kroenke et al., 2001). The PHQ-9 was specifically chosen to measure symptoms of depression, as it has been utilized in the context of social media before, such as in a study by Namhyeok et al. (2022). To get an indication of a participant's levels of depression, all the values which they indicated for the different items are combined. There are a total of five different severities that the PHQ-9 distinguishes between. The ranking of this is visualized in Table 2.

The reliability of the PHQ-9 was excellent ($\alpha = .89$). Regarding the validity of the scale, it was shown that there was a high correlation between the increasing scores on the PHQ-9 scale and decreasing scores on the SF-20 scale, which measures the quality of life in individuals (Kroenke et al., 2001).

Table 2*PHQ-9 score*

<i>PHQ-9 score</i>	<i>Depression Severity</i>
0-4	Minimal depression
5-9	Mild depression
10-14	Moderate depression
15-19	Moderately severe depression
20-27	Severe depression

Note: Definitions by (Kroenke et al., 2001).

Demographical data

At the beginning of the survey, participants were asked to indicate different demographic data. This included their age, gender, sex, and nationality. Participants were able to choose between ‘male’ and ‘female’ for the segment of sex. Furthermore, they were asked to indicate whether they were ‘German’, ‘Dutch’ or ‘Other’, for their nationality. Subsequently, the variable sex is going to be used to assess whether it moderates the relationship between *Instagram engagement* and *symptoms of depression*.

Procedure

Before the data collection, the research was approved by the ethics committee of the University of Twente. Shortly after, an online study was developed by using “Qualtrics”. “Qualtrics” is a website that lets you create a variety of surveys that can be customized entirely to individual preferences. Afterwards, a link was created which was then distributed to potential participants. They received the link either by participating in the study via the University of Twente program “Sona”, via social media posts, or through direct messages. When participants decided to participate in the study, they were first informed about how their provided data will be treated and that they can withdraw from the study at any given time, without having to provide any reasons. Additionally, contact information was provided for researchers as well as the supervisor were provided to answer any question that might occur. After agreeing to the informed consent form, participants continued to provide demographic data, such as their age and sex. They were then asked to answer the items on different scales measuring symptoms of depression and their Instagram engagement. When the participants completed the questionnaire, they were ultimately thanked for their participation, and that their answers have been recorded.

Data analysis

After the data was conducted via Qualtrics, R (V. 4.3.0) was used to perform various analyses (Appendix B). First, the data was imported into R-studios. Before the analyses, the data were inspected and every participant who did not consent or did not complete the survey was removed from the data set. In addition to this, every column that was not necessary for the analyses was removed (e.g., the duration participants spent completing the study). Before the beginning of the analyses, the values for the SONTUS scale and the PHQ-9 scale were coded and added to specific columns, which represented the final scores of every participant on the corresponding scale. Additionally, the assumptions for normality, homoscedasticity, and linearity were checked. The assumptions for homoscedasticity and linearity were fulfilled,

but the data was not normally distributed. The first analysis performed was the inferential statistics. Furthermore, the data was visualized in different forms (e.g.: histogram).

Subsequently, a Kendall correlational analysis was performed to verify the correlation between the SONTUS scale and the PHQ-9 questionnaire. Since the data were non-normally distributed, Kendall's tau was used for this analysis. Subsequently, a linear model was created to further inspect the relation between the two scales. This was done to check the direction of the relationship and whether it was significant.

To test whether the sex of the participant moderates the relationship between the independent and dependent variable, multiple linear regression was conducted. Here, it was analysed whether sex moderates the relationship between the two variables.

Results

Descriptives

Initially, descriptive analyses were performed to evaluate the results of the SONTUS and the PHQ-9 questionnaire. All the results are showcased in Table 3. The data suggests that the participants can be evaluated as low to average Instagram users ($M = 9.25$, $SD = 2.76$). Furthermore, the levels of depression of the participants can on average be classified as mild depression ($M = 7.6$, $SD = 4.81$).

Table 3

Variable	Mean	Standard Deviation
Engagement on Instagram	9.25	2.76
Symptoms of depression	7.6	4.81

Correlations

In addition, the correlation between the dependent and independent variables was checked. The Kendall correlation analysis indicated a moderate positive correlation between Instagram engagement and symptoms of depression ($\tau = 0.3$, $p = >.001$).

Inferential statistics

In addition to correlation analysis, a moderated regression model was created with *engagement on Instagram* as the independent variable and symptoms of depression as the dependent variable. Furthermore, sex was included as a moderator variable. The results of the moderated regression analysis can be found in Table 4. It was found that there is an effect of Instagram engagement on symptoms of depression ($R^2 = .12$, $p = .003$). Moreover, a moderate positive relationship was found between the two variables ($b = .67$) (Cohen, 1988).

When taking these results into account, the first hypothesis:” Young adults with higher rates of Instagram engagement experience more symptoms of depression.” can be accepted.

Table 4

Moderate regression analysis between Instagram engagement (IV) and symptoms of depression (DV)

	<i>estimate</i>	<i>R²</i>	<i>p</i>	<i>t(DF)</i>
Intercept	1.41			
Instagram engagement	.67	.13	.001	66

Finally, it was expected that sex moderates the relationship between Instagram engagement and symptoms of depression. In order to test this, a moderated regression model was created with sex as a moderator variable, which influences the relationship between the IV and the DV. The results of this indicate that sex does not moderate the relationship between Instagram engagement and symptoms of depression ($p = .40$). The results of this analysis are further visualized in Table 5. However, even when sex was included as a moderator variable, the original relationship between the IV and the DV was still significant. Based on these findings, the second hypothesis:” The sex of young adults positively moderates the relationship between their Instagram engagement and symptoms of depression” needs to be rejected.

Table 5

Multiple linear regression output

Variables	<i>B</i>	<i>SE</i>	<i>t</i>	<i>p</i>
Intercept	1.04			
Instagram engagement	.78	.26	2.9	.005
Sex	1.40	3.82	-.37	.71
Instagram engagement*Sex	-.34	.40	-.84	.40

Discussion

This study aimed to investigate the relationship between *Instagram engagement* and *symptoms of depression* in young adults. In addition to this, it was assessed whether *sex* moderates this relationship. The reason why young adults were chosen as the target group is that they are especially at risk of experiencing symptoms of depression concerning social

media use (Ivie et al., 2020). The results of this study aim to provide a better understanding of the topic of Instagram engagement in regards to symptoms of depression.

Implications

Firstly, when looking at the different results of the inferential statistics, interesting results can be observed. On average, the participants were rated as low to medium users of Instagram. There are multiple possible reasons as to why this is the case. One explanation might be that the participants of this sample might simply use Instagram less than the average user. Since only a total of 68 respondents participated, it is difficult to draw inferences about the general population.

Moreover, it was found that the average participant showed mild levels of depression. When comparing these findings to a study conducted by Mulansky et al. (2022), it becomes apparent that the values differ. The researchers found an average PHQ-9 score of 12.18, which classifies the participants as moderately depressed, for adults between the ages of 18 and 29 years. These findings might be due to the fact that the participants were low users of Instagram, to begin with. Regarding the first hypothesis of the study, it was anticipated that young adults with higher rates of Instagram engagement experience more symptoms of depression. When looking at the results of the Kendall correlation analysis, as well as the results of the moderated regression analysis, it becomes apparent that there is a moderate positive relationship between Instagram engagement and symptoms of depression. These findings are consistent with previous research that indicated a small positive effect of social media use on depression in young adults (Ivie et al., 2020). Furthermore, Mulansky et al. (2022) found a positive relationship between using Instagram and symptoms of depression as well. Based on these findings, there is an indication of a positive relationship between Instagram engagement and symptoms of depression in young adults between the ages of 18 and 24.

In terms of the second hypothesis, it was estimated that the sex of young adults moderates the relationship between their Instagram engagement and symptoms of depression. When considering the results of the moderation analysis, it becomes observable that the moderation effect is not significant. This means that the second hypothesis cannot be confirmed. One possible explanation as to why this is the case would be that there are additional underlying aspects that influence symptoms of depression regarding Instagram use, that influence men and women. Research by Sherlock and Wagstaff (2019) suggests that the type of content which is consumed is crucial. They suggest that women who view content related to fitness and beauty experience a decrease in perceived personal attractiveness.

Subsequently, this decrease fosters the occurrence of symptoms of depression. Furthermore, findings by Parent et al. (2019) estimate that there is a relationship between social media use, toxic masculinity, and depression. It could thus be the case that primarily men who engage in toxic masculinity online are more prone to symptoms of depression. Based on these findings, one possible explanation for why sex does not moderate the relationship between Instagram engagement and symptoms of depression is, that there are other underlying factors, such as the consumed content or online activities, that have not been accounted for in this study.

Strengths and Limitations

One strength of this study is that the PHQ-9 questionnaire, which was used to assess the participant's symptoms of depression is valid and reliable. In addition to this, the PHQ-9 questionnaire has been used by different researchers before, to assess symptoms of depression in the context of social media use (Mulansky et al., 2022; Zhang et al., 2023). Furthermore, the PHQ-9 was found to be applicable to the general population, as opposed to only patients (Martin et al., 2006). Based on this, the PHQ-9 can be classified as a useful tool to assess symptoms of depression in the domain of social media. Another strength of the study is that it solely resolves around Instagram and no other SNS. Since most studies focus on multiple SNS at once, this study discovered new information which can be precisely applied to Instagram. Due to this, more about the complex relationship between Instagram and symptoms of depression was discovered.

However, there are a couple of limitations that become apparent. First, it should be noted that by using a cross-sectional study design, only data at one point in time is collected from the participants (Kesmodel, 2018). Due to this, it might be the case that the increased or decreased symptoms of depression of the participants then are not related to Instagram engagement, but other external factors. More precisely, by using a cross-sectional research design, it becomes difficult to make assumptions, if it is not ensured that the observed effect is stable over time (Kesmodel, 2018).

Secondly, by using a self-report questionnaire, it cannot be ruled out that some participants do not answer the questions truthfully. This may result in additional biases which, in turn, reduce the validity of the investigation. An example of such a bias is the social desirability bias, where participants tend to answer self-report questionnaires in a way that makes them seem more desirable (Van De Mortel, 2008). This not only includes participants deliberately providing false answers, but also potential misunderstanding of items on the questionnaires. Therefore, it could be possible that participants answered the questionnaire to

the best of their ability, yet they still might have understood the questions in a different way than the researcher.

Implications for future research

Regarding future research in this domain, it might prove beneficial to include a study design involving a pre-and post-test, to further collect data at different points in time. This could be done by measuring the participants' Instagram engagement and symptoms of depression at multiple points in time, to see assess potential variations. Furthermore, this might provide a clearer understanding of the relationship between the participant's Instagram engagement and their symptoms of depression. It might also be beneficial to test different factors in future studies, as the presented moderation effect was found to be non-significant, and there might be other factors involved, that moderate the relationship between Instagram engagement and symptoms of depression. By doing this, different confounding variables could be identified, related to the relationship between the two variables.

Lastly, it might prove useful to use a bigger sample with more participants. By doing this, the derived results would be more generalizable since there are more observations that way. Doing this might provide a deeper understanding of the complexity of the relationship between the two variables.

Conclusion

The daily use of SNS has become a part of the daily routine of most people around the world. Although it might have beneficial effects on individuals, there are certainly downsides associated with its use. This study focused on its relationship with symptoms of depression in young adults. The results indicate that there is a moderately positive and significant relation between the two. Despite these findings, it also became apparent that sex was not a significant moderator of this relationship. However, this study provided further information on the complex relationship between social media use and feelings of depression, while also highlighting the need for further research in this field.

References

- Adeyanju, G. C., Solfa, R. P., Tran, T. L., Wohlfarth, S., Büttner, J., Osobajo, O. A. & Otitoju, A. (2021). Behavioral symptoms of mental health disorder such as depression among young people using Instagram: a systematic review. *Springer Link*, 6(1).
<https://doi.org/10.1186/s41231-021-00092-3>
- Amedie, J. (2015). The Impact of Social Media on Society. *Pop Culture Intersections*.
https://scholarcommons.scu.edu/cgi/viewcontent.cgi?article=1002&context=engl_176
- Argoff, C. & McCleane, G. (2009). Pain management secrets. In *Elsevier eBooks*.
<https://ci.nii.ac.jp/ncid/BB0305772X>
- Atske, S. (2022, 11. Mai). *Social Media Use in 2021 | Pew Research Center*. Pew Research Center: Internet, Science & Tech.
[https://www.pewresearch.org/internet/2021/04/07/social-media-use-in-2021/#:~:text=Some%2084%25%20of%20adults%20ages,45%25\)%20report%20doing%20this.](https://www.pewresearch.org/internet/2021/04/07/social-media-use-in-2021/#:~:text=Some%2084%25%20of%20adults%20ages,45%25)%20report%20doing%20this.)
- Baker, D. & Algorta, G. P. (2016). The Relationship Between Online Social Networking and Depression: A Systematic Review of Quantitative Studies. *Cyberpsychology, Behavior, and Social Networking*, 19(11), 638–648.
<https://doi.org/10.1089/cyber.2016.0206>
- Bale, T. L. & Epperson, C. N. (2016). Sex as a Biological Variable: Who, What, When, Why, and How. *Neuropsychopharmacology*, 42(2), 386–396. <https://doi.org/10.1038/npp.2016.215>
- Barnhart, B. (2023). Social media demographics to inform your brand’s strategy in 2023. *Sprout Social*. <https://sproutsocial.com/insights/new-social-media-demographics/>
- Blystone, D. (2022). Instagram: What It Is, Its History, and How the Popular App Works. *Investopedia*. <https://www.investopedia.com/articles/investing/102615/story-instagram-rise-1-photo0sharing-app.asp#toc-history-of-instagram>

- Ciranka, S. & Van Den Bos, W. (2019). Social Influence in Adolescent Decision-Making: A Formal Framework. *Frontiers in Psychology*, *10*. <https://doi.org/10.3389/fpsyg.2019.01915>
- Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences* (2nd ed.). Routledge. <https://doi.org/10.4324/9780203771587>
- Green, D. D. & Martinez, R. (2018). In a World of Social Media: A Case Study Analysis of Instagram. *American research journal of business and management*, *4*(1). <https://doi.org/10.21694/2379-1047.18012>
- Huang, Y. & Su, S. (2018). Motives for Instagram Use and Topics of Interest among Young Adults. *Future Internet*, *10*(8), 77. <https://doi.org/10.3390/fi10080077>
- Hwang, H. S. (2019). Why Social Comparison on Instagram Matters: Its impact on Depression. *Ksii Transactions on Internet and Information Systems*. <https://doi.org/10.3837/tiis.2019.03.029>
- Hyde, J. S. (2014). Gender Similarities and Differences. *Annual Review of Psychology*, *65*(1), 373–398. <https://doi.org/10.1146/annurev-psych-010213-115057>
- Instagram. (n. D.). *Instagram Launches*. accessed 24. May 2023, via <https://about.instagram.com/blog/announcements/instagram-launches>
- Ivie, E. J., Pettitt, A., Moses, L. J. & Allen, N. B. (2020). A meta-analysis of the association between adolescent social media use and depressive symptoms. *Journal of Affective Disorders*, *275*, 165–174. <https://doi.org/10.1016/j.jad.2020.06.014>
- Kesmodel, U. S. (2018). Cross-sectional studies - what are they good for? *Acta Obstetrica et Gynecologica Scandinavica*, *97*(4), 388–393. <https://doi.org/10.1111/aogs.13331>
- Kessler, R. C. (2003). Epidemiology of women and depression. *Journal of Affective Disorders*, *74*(1), 5–13. [https://doi.org/10.1016/s0165-0327\(02\)00426-3](https://doi.org/10.1016/s0165-0327(02)00426-3)

- Kroenke, K., Spitzer, R. L. & Williams, J. B. W. (2001). The PHQ-9. *Journal of General Internal Medicine*, 16(9), 606–613. <https://doi.org/10.1046/j.1525-1497.2001.016009606.x>
- Li, J., Lau, J., Mo, P. K. H., Su, X., Tang, J., Qin, Z. & Gross, D. L. (2017). Insomnia partially mediated the association between problematic Internet use and depression among secondary school students in China. *Journal of behavioral addictions*, 6(4), 554–563. <https://doi.org/10.1556/2006.6.2017.085>
- Lin, L., Sidani, J. E., Shensa, A., Radovic, A., Miller, E., Colditz, J. B., Hoffman, B. J., Giles, L. M. & Primack, B. A. (2016). ASSOCIATION BETWEEN SOCIAL MEDIA USE AND DEPRESSION AMONG U.S. YOUNG ADULTS. *Depression and Anxiety*, 33(4), 323–331. <https://doi.org/10.1002/da.22466>
- Mann, V. A., Sasanuma, S., Sakuma, N. & Masaki, S. (1990). Sex differences in cognitive abilities: A cross-cultural perspective. *Neuropsychologia*, 28(10), 1063–1077. [https://doi.org/10.1016/0028-3932\(90\)90141-a](https://doi.org/10.1016/0028-3932(90)90141-a)
- Martin, A., Rief, W., Klaiberg, A. & Braehler, E. (2006). Validity of the Brief Patient Health Questionnaire Mood Scale (PHQ-9) in the general population. *General Hospital Psychiatry*, 28(1), 71–77. <https://doi.org/10.1016/j.genhosppsych.2005.07.003>
- Mulansky, L., Pryss, R., Cohrdes, C., Baumeister, H. & Beierle, F. (2022a). *Social Media App Usage in Relation with PHQ-9 Depression Scores during the COVID-19 Pandemic*. <https://doi.org/10.1145/3544793.3563411>
- Namhyeok, K., Kim, J., Park, D. H., Ji, S. R. & Kim, J. M. (2022). Analysis of depression in social media texts through the Patient Health Questionnaire-9 and natural language processing. *Digital health*, 8, 205520762211142. <https://doi.org/10.1177/20552076221114204>
- Noble, R. E. (2005). Depression in women. *Metabolism-clinical and Experimental*, 54(5), 49–52. <https://doi.org/10.1016/j.metabol.2005.01.014>

- O'Connor, E., Whitlock, E., Gaynes, B. N. & Beil, T. (2009). Screening for Depression in Adults and Older Adults in Primary Care: An Updated Systematic Review. *Agency for Healthcare Research and Quality (US)*.
<https://europepmc.org/abstract/MED/20722174>
- Olufadi, Y. (2016). Social networking time use scale (SONTUS): A new instrument for measuring the time spent on the social networking sites. *Telematics and Informatics*, 33(2), 452–471. <https://doi.org/10.1016/j.tele.2015.11.002>
- Parent, M. C., Gobble, T. & Rochlen, A. B. (2019). Social media behavior, toxic masculinity, and depression. *Psychology of Men and Masculinity*, 20(3), 277–287. <https://doi.org/10.1037/men0000156>
- Shahjahan, A. T. M. & Chisty, K. U. (2014). Social Media Research and Its Effect on Our Society. *Journal of information and communication convergence engineering*, 8(6), 2009–2013. <https://doi.org/10.5281/zenodo.1093964>
- Sherlock, M. & Wagstaff, D. (2019). Exploring the relationship between frequency of Instagram use, exposure to idealized images, and psychological well-being in women. *Psychology of popular media culture*, 8(4), 482–490.
<https://doi.org/10.1037/ppm0000182>
- Shorey, S., Ng, E. D. & Wong, C. (2022). Global prevalence of depression and elevated depressive symptoms among adolescents: A systematic review and meta-analysis. *British Journal of Clinical Psychology*, 61(2), 287–305.
<https://doi.org/10.1111/bjc.12333>
- Statista. (2023, March 6). Number of Instagram users worldwide from 2018 to 2027 (in millions) [Graph]. In Statista. Retrieved May 20, 2023, from <https://www-statista-com.ezproxy2.utwente.nl/forecasts/1138856/instagram-users-in-the-world>
- Thapar, A., Eyre, O., Patel, V. & Brent, D. A. (2022). Depression in young people. *The Lancet*, 400(10352), 617–631. [https://doi.org/10.1016/s0140-6736\(22\)01012-1](https://doi.org/10.1016/s0140-6736(22)01012-1)

- Vannucci, A., Flannery, K. M. & Ohannessian, C. M. (2017). Social media use and anxiety in emerging adults. *Journal of Affective Disorders*, 207, 163–166.
- Van De Mortel, T. F. (2008). Faking it: Social desirability response bias in self-report research. *Australian Journal of Advanced Nursing*, 25(4), 40–48. http://epubs.scu.edu.au/cgi/viewcontent.cgi?article=1001&context=hahs_pubs
- Vázquez-Herrero, J., Direito-Rebollal, S. & García, X. L. (2019). Ephemeral Journalism: News Distribution Through Instagram Stories. *Social media and society*, 5(4), 205630511988865. <https://doi.org/10.1177/2056305119888657>
- Wang, P., Wang, X., Wu, Y., Xie, X., Zhao, F. & Ouyang, M. (2018). Social networking sites addiction and adolescent depression: A moderated mediation model of rumination and self-esteem. *Personality and Individual Differences*, 127, 162–167. <https://doi.org/10.1016/j.paid.2018.02.008>
- We Are Social & DataReportal & Meltwater. (January 26, 2023). Distribution of Facebook users worldwide as of January 2023, by age and gender [Graph]. In Statista. Retrieved May 20, 2023, from <https://www-statista-com.ezproxy2.utwente.nl/statistics/376128/facebook-global-user-age-distribution/>
- Woods, H. C. & Scott, H. (2016). #Sleepyteens: Social media use in adolescence is associated with poor sleep quality, anxiety, depression and low self-esteem. *Journal of Adolescence*, 51(1), 41–49. <https://doi.org/10.1016/j.adolescence.2016.05.008>
- World Health Organization: WHO & World Health Organization: WHO. (2023). Depressive disorder (depression). *www.who.int*. <https://www.who.int/news-room/fact-sheets/detail/depression>
- Zhang, T., Yang, K., Alhuzali, H., Liu, B. & Ananiadou, S. (2023). PHQ-aware depressive symptoms identification with similarity contrastive learning on social media. *Information Processing and Management*, 60(5), 103417. <https://doi.org/10.1016/j.ipm.2023.103417>

Appendix

Appendix A: Informed consent

Dear participant,

You are invited to participate in an online survey “The Relationship between Social Media Use and Mental Health”. From this, you have the right to withdraw at any time without any negative consequences or providing any reasons. For that, please note that there are no right or wrong answers so it is about indicating your subjective view. You will be asked to provide demographic information. The research procedure consists of a questionnaire which will take approximately 15 minutes. The results will be saved for later investigation purposes. The data we collect will be handled confidentially. The analysis of the data will be anonymous and only for purposes of the research. After completing the project, the data will be deleted. Thus there are no expected risks within the participation in our study. You need to be at least 18 years old to participate.

If you have any further questions, feel free to contact us:

Malin Cybaier (m.k.cybaier@student.utwente.nl)

Daniel Barth (d.d.barth@student.utwente.nl)

Supervisor: Martha Kreuzberg (m.s.kreuzberg@utwente.nl)

Appendix B: R-Studio script

```
## Set working directory ##
```

```
setwd("C:/Users/danie/OneDrive/Desktop/University/University of Twente/Bachelor/R-  
script/R-script Bachelor")
```

```
## Load packages ##
```

```
library(readxl)
```

```
library(foreign)
```

```
library(lavaan)
```

```
library(lavaanPlot)
```

```
library(dplyr)
```

```
library(haven)
```

```
library(ggpubr)
```

```
library(knitr)
```

```
library(semPlot)
```

```
library(MVN)
```

```
library(tidyr)
```

```
library(tidyverse)
```

```
library(WriteXLS)
```

```
#library(ltm)
```

```
#library(outliers)
```

```
#library(EnvStats)

#library(broom)

#library(Matrix)

#library(lme4)

#library(knitr)

#library(officer)

library(car)

#### Import data ####

Control_Bachelor <- read_xlsx("Bachelor_survey_02.05.2023.xlsx")

View(Control_Bachelor)

summary(Control_Bachelor)

#### Clean data set ####

### Remove Completed = False ###

## Delete rows where the value in "my_column" is "False" ##

Bachelor_survey <- Control_Bachelor[Control_Bachelor$Finished != "False",]

### Remove columns ###

## Remove multiple columns from the data set ##

Bachelor_survey <- Bachelor_survey[, !(names(Bachelor_survey) %in% c("StartDate",
  "EndDate", "Status", "IPAddress", "Progress", "Duration (in seconds)",
  "RecordedDate", "Responseid", "RecipientLastName", "RecipientFirstName",
  "RecipientEmail", "ExternalReference", "LocationLatitude", "LocationLongitude",
  "DistributionChannel", "UserLanguage", "Q14_3_TEXT", "PHQ-9_1", "PHQ-9_2",
```

```

"PHQ-9_3", "PHQ-9_4", "PHQ-9_5", "PHQ-9_6", "PHQ-9_7", "PHQ-9_8", "PHQ-
9_9"))]

### Remove first row ###

Bachelor_survey <- slice(Bachelor_survey, -1)

### Remove other scales ###

# Self compassion #

columns_to_remove_SC <- grep("^Self-compassion_", names(Bachelor_survey), value =
  TRUE)

Bachelor_survey <- Bachelor_survey[, !(names(Bachelor_survey) %in%
  columns_to_remove_SC)]

# Fear of missing out #

columns_to_remove_FOMO <- grep("^Fear of Missing Out", names(Bachelor_survey), value
  = TRUE)

Bachelor_survey <- Bachelor_survey[, !(names(Bachelor_survey) %in%
  columns_to_remove_FOMO)]

# SMUIS #

columns_to_remove_SMUIS <- grep("^SMUIS", names(Bachelor_survey), value = TRUE)

Bachelor_survey <- Bachelor_survey[, !(names(Bachelor_survey) %in%
  columns_to_remove_SMUIS)]

### create participant numbers ###

Bachelor_survey %>%

mutate(Subjects = row_number())

```



```
#### Turn values numeric ####

## SONTUS scale ##

Bachelor_survey <- Bachelor_survey %>%

  mutate(across(starts_with("SONTUS"), as.numeric))

## Q-14_ ##

Bachelor_survey <- Bachelor_survey %>%

  mutate(across(starts_with("Q14_"), as.numeric))

## Numeric Check ##

# Check if values in my_column are numeric

numeric_check <- is.numeric(Bachelor_survey$Sontus_14)

# Print logical vector

print(numeric_check)

## remove participants over the age of 24 ##

Bachelor_survey <- Bachelor_survey[Bachelor_survey$Age <= 24, ]

##### Calculate scores #####

### Descriptives ###

## Sontus ##

mean_sontus <- mean(Bachelor_survey$Sontus_global_score)

sd_sontus <- sd(Bachelor_survey$Sontus_global_score)

## PHQ-9 ##

mean_phq <- mean(Bachelor_survey$PHQ_9_score)
```

```
sd_phq <- sd(Bachelor_survey$PHQ_9_score)

### SONTUS ###

## Code the individual scales ##

assign_value_general <- function(x) {

  if (x >= 1 & x <= 3) {

    return(1)

  } else if (x >= 4 & x <= 6) {

    return(2)

  } else if (x >= 7 & x <= 9) {

    return(3)

  } else if (x >= 10 & x <= 11) {

    return(4)

  }

}

Bachelor_survey$Sontus_1 <- sapply(Bachelor_survey$`SONTUS scale_1`,
  assign_value_general)

Bachelor_survey$Sontus_2 <- sapply(Bachelor_survey$`SONTUS scale_2`,
  assign_value_general)

Bachelor_survey$Sontus_3 <- sapply(Bachelor_survey$`SONTUS scale_3`,
  assign_value_general)

Bachelor_survey$Sontus_4 <- sapply(Bachelor_survey$`SONTUS scale_4`,
  assign_value_general)
```

```
Bachelor_survey$Sontus_5 <- sapply(Bachelor_survey$`SONTUS scale_5`,  
  assign_value_general)
```

```
Bachelor_survey$Sontus_6 <- sapply(Bachelor_survey$`SONTUS scale_6`,  
  assign_value_general)
```

```
Bachelor_survey$Sontus_7 <- sapply(Bachelor_survey$`SONTUS scale_7`,  
  assign_value_general)
```

```
Bachelor_survey$Sontus_8 <- sapply(Bachelor_survey$`SONTUS scale_8`,  
  assign_value_general)
```

```
Bachelor_survey$Sontus_9 <- sapply(Bachelor_survey$`SONTUS scale_9`,  
  assign_value_general)
```

```
Bachelor_survey$Sontus_10 <- sapply(Bachelor_survey$`SONTUS scale_10`,  
  assign_value_general)
```

```
Bachelor_survey$Sontus_11 <- sapply(Bachelor_survey$`SONTUS scale_11`,  
  assign_value_general)
```

```
Bachelor_survey$Sontus_12 <- sapply(Bachelor_survey$`SONTUS scale_12`,  
  assign_value_general)
```

```
Bachelor_survey$Sontus_13 <- sapply(Bachelor_survey$`SONTUS scale_13`,  
  assign_value_general)
```

```
Bachelor_survey$Sontus_14 <- sapply(Bachelor_survey$`SONTUS scale_14`,  
  assign_value_general)
```

```
Bachelor_survey$Sontus_15 <- sapply(Bachelor_survey$`SONTUS scale_15`,  
  assign_value_general)
```

```
Bachelor_survey$Sontus_16 <- sapply(Bachelor_survey$`SONTUS scale_16`,  
  assign_value_general)
```

```
Bachelor_survey$Sontus_17 <- sapply(Bachelor_survey$`SONTUS scale_17`,  
  assign_value_general)
```

```
Bachelor_survey$Sontus_18 <- sapply(Bachelor_survey$`SONTUS scale_18`,  
  assign_value_general)
```

```
Bachelor_survey$Sontus_19 <- sapply(Bachelor_survey$`SONTUS scale_19`,  
  assign_value_general)
```

```
Bachelor_survey$Sontus_20 <- sapply(Bachelor_survey$`SONTUS scale_20`,  
  assign_value_general)
```

```
Bachelor_survey$Sontus_21 <- sapply(Bachelor_survey$`SONTUS scale_21`,  
  assign_value_general)
```

```
Bachelor_survey$Sontus_22 <- sapply(Bachelor_survey$`SONTUS scale_22`,  
  assign_value_general)
```

```
Bachelor_survey$Sontus_23 <- sapply(Bachelor_survey$`SONTUS scale_23`,  
  assign_value_general)
```

```
Bachelor_survey$Sontus_24 <- sapply(Bachelor_survey$`SONTUS scale_24`,  
  assign_value_general)
```

```
Bachelor_survey$Sontus_25 <- sapply(Bachelor_survey$`SONTUS scale_25`,  
  assign_value_general)
```

```
Bachelor_survey$Sontus_26 <- sapply(Bachelor_survey$`SONTUS scale_26`,  
  assign_value_general)
```

```
Bachelor_survey$Sontus_27 <- sapply(Bachelor_survey$`SONTUS scale_27`,  
  assign_value_general)
```

```
Bachelor_survey$Sontus_28 <- sapply(Bachelor_survey$`SONTUS scale_28`,  
  assign_value_general)
```

```
Bachelor_survey$Sontus_29 <- sapply(Bachelor_survey$`SONTUS scale_29`,  
  assign_value_general)
```

```
## Group Items ##
```

```
# Component 1 #
```

```
Bachelor_survey <- Bachelor_survey %>%

  mutate(Comp_1 = `Sontus_2` + `Sontus_6` + `Sontus_7` + `Sontus_12` + `Sontus_14` +
    `Sontus_21` + `Sontus_22` + `Sontus_24` + `Sontus_26`)

# Component 2 #

Bachelor_survey <- Bachelor_survey %>%

  mutate(Comp_2 = `Sontus_1` + `Sontus_5` + `Sontus_10` + `Sontus_13` + `Sontus_28` +
    `Sontus_29`)

# Component 3 #

Bachelor_survey <- Bachelor_survey %>%

  mutate(Comp_3 = `Sontus_4` + `Sontus_9` + `Sontus_17` + `Sontus_19` + `Sontus_23`)

# Component 4 #

Bachelor_survey <- Bachelor_survey %>%

  mutate(Comp_4 = `Sontus_3` + `Sontus_8` + `Sontus_15` + `Sontus_16` + `Sontus_27`)

# Component 5 #

Bachelor_survey <- Bachelor_survey %>%

  mutate(Comp_5 = `Sontus_11` + `Sontus_18` + `Sontus_20` + `Sontus_25`)

# Component 1 scoring #

assign_value1 <- function(x) {

  if (x >= 9 & x <= 12) {

    return(1)

  } else if (x >= 13 & x <= 16) {
```

```
    return(2)

  } else if (x >= 17 & x <= 20) {

    return(3)

  } else if (x >= 21 & x <= 24) {

    return(4)

  } else if (x >= 25 & x <= 28) {

    return(5)

  } else if (x >= 29 & x <= 32) {

    return(6)

  } else {

    return(7)

  }

}

Bachelor_survey$Comp_1_score <- sapply(Bachelor_survey$Comp_1, assign_value1)

# Component 2 scoring #

assign_value2 <- function(x) {

  if (x >= 6 & x <= 9) {

    return(1)

  } else if (x >= 10 & x <= 13) {

    return(2)

  } else if (x >= 14 & x <= 17) {
```

```
    return(3)

  } else if (x >= 18 & x <= 21) {

    return(4)

  } else {

    return(5)

  }

}

Bachelor_survey$Comp_2_score <- sapply(Bachelor_survey$Comp_2, assign_value2)

# Component 3 scoring #

assign_value3 <- function(x) {

  if (x >= 5 & x <= 8) {

    return(1)

  } else if (x >= 9 & x <= 12) {

    return(2)

  } else if (x >= 13 & x <= 16) {

    return(3)

  } else if (x >= 17 & x <= 20) {

    return(4)

  }

}

Bachelor_survey$Comp_3_score <- sapply(Bachelor_survey$Comp_3, assign_value3)
```

```
# Component 4 scoring #
```

```
assign_value4 <- function(x) {
```

```
  if (x >= 5 & x <= 8) {
```

```
    return(1)
```

```
  } else if (x >= 9 & x <= 12) {
```

```
    return(2)
```

```
  } else if (x >= 13 & x <= 16) {
```

```
    return(3)
```

```
  } else if (x >= 17 & x <= 20) {
```

```
    return(4)
```

```
  }
```

```
}
```

```
Bachelor_survey$Comp_4_score <- sapply(Bachelor_survey$Comp_4, assign_value4)
```

```
# Component 5 scoring #
```

```
assign_value5 <- function(x) {
```

```
  if (x >= 4 & x <= 7) {
```

```
    return(1)
```

```
  } else if (x >= 8 & x <= 11) {
```

```
    return(2)
```

```
  } else {
```

```
    return(3)
```



```
}  
  
}  
  
Bachelor_survey$Comp_5_score <- sapply(Bachelor_survey$Comp_5, assign_value5)  
  
## Sontus Global scoring ##  
  
Bachelor_survey <- Bachelor_survey %>%  
  
  mutate(Sontus_global_score = Comp_1_score + Comp_2_score + Comp_3_score +  
         Comp_4_score + Comp_5_score)  
  
#### PHQ-9 ####  
  
## Combine values ##  
  
Bachelor_survey <- Bachelor_survey %>%  
  
  mutate(PHQ_9_score = Q14_1 + Q14_2 + Q14_3 + Q14_4 + Q14_5 + Q14_6 + Q14_7 +  
         Q14_8 + Q14_9)  
  
#### Check whether the data is normally distributed ####  
  
shapiro.test(Bachelor_survey$Sontus_global_score)  
  
shapiro.test(Bachelor_survey$PHQ_9_score)  
  
# Use grep() to select the columns based on their names  
  
column_names <- grep("^Sontus_", names(Bachelor_survey), value = TRUE)  
  
# Apply the Shapiro-Wilk test to the selected columns  
  
shapiro_results <- lapply(Bachelor_survey[, column_names], shapiro.test)  
  
# Print the results  
  
print(shapiro_results)
```

```
# Use grep() to select the columns based on their names

column_names2 <- grep("^Q14_", names(Bachelor_survey), value = TRUE)

# Apply the Shapiro-Wilk test to the selected columns

shapiro_results2 <- lapply(Bachelor_survey[, column_names2], shapiro.test)

# Print the results

print(shapiro_results2)

## visualize data ##

hist(Bachelor_survey$Sontus_global_score)

hist(Bachelor_survey$PHQ_9_score)

Bachelor_survey %>%

  summary(Sontus_global_score)

### Correlation analysis ###

cor.test(Bachelor_survey$Sontus_global_score,
         Bachelor_survey$PHQ_9_score,use="pairwise.complete.obs", method = "kendall")

# p-value: .0005487

# Tau: 0.2959884

## regression analysis ##

model_orignal <-lm(PHQ_9_score ~ Sontus_global_score, data = Bachelor_survey)

summary(model_orignal)

# Create a new variable called "male_dummy" with all male participants scoring 1

Bachelor_survey$male_dummy <- ifelse(Bachelor_survey$Sex == "Male", 1, 0)
```

```
# Print the updated data frame
```

```
print(Bachelor_survey)
```

```
## Moderated regression model ##
```

```
model <- lm(PHQ_9_score ~ Sontus_global_score*Sex, data=Bachelor_survey)
```

```
summary(model)
```

```
# Moderation with sex: p-value: .848
```

```
## Check linearity ##
```

```
ncvTest(model)
```

```
plot(model, 1)
```