INST627

Data Analytics for Information Professionals



***Project Report:***

Cyberbullying among Adolescents

***Submitted By:***

Akhil Gupta

Deon Crasto

Srikanth Jaikumar

**Table of Contents**

Page No

1. INTRODUCTION 3
2. RESEARCH QUESTIONS 5
3. METHODOLOGY 6
   1. VARIABLES USED 6
   2. WORKING WITH DATA 9
4. ANALYSIS 9
5. INTERPRETATION OF FINDINGS 13
6. CONCLUSION 15
7. REFERENCES 15

**INTRODUCTION**

The advent of the internet, meteoric rise of social media, increase in use of smartphones and rapid development of technologies used for immediate communication have all played a role in what is termed as “Cyberbullying”. Bullying is a phenomena generally experienced in school by children in the form of verbal, physical or psychological abuse. One of several characteristic of a bully is to prey on what they consider as inferior to them in order to feel better about themselves. They try and target individuals who have an introverted personality, are shy, feel insecure and anxious. Research conducted at two high schools in USA has shown that 71% of students are traditional victims (Raskauskas & Stoltz, 2007). Such traditional form of bullying that starts in school has progressed on with the help of increasing communication mediums into the technology spectrum and is now widely known as “Cyberbullying”. It is any form of bullying that involves the use of electronic media to propagate messages that have a negative connotation used primarily to belittle an individual.

In cyberbullying, the victim is subjected to continuous harassment around the clock and is unable to find a medium of escape from the same. Moreover, the platform on which it takes place caters to a wider audience. The person who engages in such behavior may choose to act anonymously or in a telling manner but in the end is not able to or chooses not to comprehend the impact it has on the individual being cyberbullied. Further research has also suggested the impact that gender difference has on cyberbullying (Li, 2006). There is a difference in the way males and females themselves are cyberbullied and in the way both respond to the cyberbullies. Qing Li in her paper noted that “there was no difference in the number of men and women being cyberbullied”. However, it was noted from statistical results obtained that “males were more likely to be cyberbullied than females” (Li, 2006).

The unimaginable impact that social media has had on our lives is also a prime factor responsible for the state of cyberbullying as it is now. Kowalski and Limber in their research noted that “instant messaging and email” were the most likely reported methods used for cyberbullying (Agatston, Kowalski, & Limber, 2007). Davidson in his research article entitled “The Dangers of Cyberbullying” notes that social media sites such as Facebook and twitter will become the new leading platform for cyberbullying and serve as a springboard for cyberbullies (Davison & Stein, 2007).

Ask.fm and Formspring, both are question and answer based social networking sites with number of users in millions. Individuals including small children can sign up on this website and ask questions of other users. The main advantage with this system is in its concept of anonymity which is exploited by cyberbullies. Both of these websites have fallen under controversy owing to the sinister nature of questions and answers that are posted on the websites. There have been cases that have led to suicide among teens due to cyberbullying faced on these websites.

A study was conducted to understand the “perception of college students on cyberbullying” (Paullet & Pinchot, 2014). The findings of the same revealed that “21% of 168 undergraduate students surveyed” had been a victim of cyberbullying. Out of those “97% stated they had experienced cyberbullying in high school”. “9% of the sample stated they were currently being cyberbullied” which goes to show it is prevalent at a university level as well.

Everyone would agree that cyberbullying is most certainly wrong and needs to be stopped. However, how many of us actually come forward to help others who are cyberbullied. If this question was rephrased in another manner as-how many of us would like to help others who are being cyberbullied but are not able to do so would garner more positive responses. As cyberbullies make use of the cloak of invisibility to target their victims, is it so that even people who come forward to help victims need anonymity. This is a very interesting research question that is trending among data scientists. Prevention of cyberbullying is an important aspect that must be addressed at a very early stage considering the dominance that the world of technology today has on us from an early age. Some suggestions provided by college students to prevent cyberbullying were inclusion of training programs at high school, awareness of legal laws associated with cyberbullying, students should be taught how to report incidents and call for more support from the online community who witness cyberbullying but still choose to remain quiet (Paullet & Pinchot, 2014).

Weight-based cyberbullying is another form of cyberbullying that has taken a foothold especially in social media circles. Weight-based bullying is experienced by adolescents in high school wherein they are targeted for being obese. In order to propagate such form of bullying, cyberbullying is made use of with statistics showing “61% of adolescents have received mean or embarrassing posts online and 59% have received mean texts, e-mails or instant messages” (Puhl, Peterson, & Luedicke, 2013)

The research we have conducted focuses primarily on understanding cyberbullying behavior among young adults by making use of research questions that touch upon aspects as highlighted above.

**RESEARCH QUESTIONS**

There are four research questions that we have focused on stated as below:

**RQ1:** Does increased interaction with strangers lead to an increase in Cyberbullying?

**RQ2:** Are Cyberbullying victims more likely to support other cyberbullying victims by posting supportive comments on their page?

**RQ3:** Does Gender play a role in being cyberbullied?

**RQ4:** Does difference in Nationality play a role in activities that could be termed as Cyberbullying?

The questions listed above try and predict how user activity online on social media platforms such as Ask.fm, Formspring, Facebook, Twitter, LinkedIn etc. are linked to Cyberbullying. In order for us to predict to a certain extent the answers to these questions, we make use of data that is collected through a survey conducted. This includes responses to questions which when analyzed were used as independent and dependent variables. It helped us conduct statistical analyses to test hypothesis and assumptions related to each question.

Cyberbullying increases with increased usage of Internet (Feinberg, Ted, & Nicole Robey, 2008) and interaction with strangers (Lenhart, 2008), but does interacting with a stranger often lead to increased Cyberbullying? In RQ1 we try to answer this question by taking self-reported frequency of interaction with strangers as an Independent variable. This question is interesting from a perspective of how anonymity plays a role when it comes to interaction with people in social media.

A general consensus is that cyberbullying victims will support other victims as they themselves have been traumatized by the experience and can understand the torment better than anyone else (Patchin, Justin, & Hinduja, 2010). RQ2 helps us analyze if this consensus holds true or if the victims prefer to not engage in any online activity in the future which could lead to them having to relive their experience all over again.

RQ3 helps us understand if an individual’s sex has any role to play with him/her being cyberbullied. We analyze two groups of data – males versus females and their frequency of being cyberbullied online. This helps us draw a comparison between the two samples to determine if there is a significant deviation of means from one another. It would be interesting to note here the direction in which cyberbullying leans on more in terms of gender.

Where a person comes from, the culture, the environment in which he/she has been brought up has a great role to play in the way in which their lives will shape out (Berger, 2003). Keeping this in mind, RQ4 focuses our attention on American and Non-American groups of people interlinked with their engagement in activities that can be termed as cyberbullying.

**METHODOLOGY**

The dataset that we have worked with is a survey that was distributed among users of Ask.fm and Formspring comprising of young adults aged between 18-21 years old created by Dr. Vitak (Vitak, 2014). The dataset comprises of variables that were coded from responses to questions posed to understand human behavior for dealing with cyberbullying. The data has a sample size of 244 responses out of which 158 are males and 86 females. Majority of participants were white North Americans followed by Indians and people of other nationality. It is interesting to note here that more than 70% of the participants used social media websites such as Facebook and Twitter as compared to other services such as Tinder, Instagram, WhatsApp, Google Plus etc.

**VARIABLES USED**

**Research Question 1**

**Interaction with strangers:** This is a dichotomous variable which is taken as an Independent variables. Response to question ‘How often do you interact with strangers’ with responses ‘Never, Rarely, Sometimes, Often, Very Often’ has been recoded into new variable with two levels ‘Low Interaction’ and ‘High Interaction’.

**Cyberbullying\_Freq**: Dependent variable which indicates the frequency of being cyberbullied. The values it can take are Never, Rarely, Sometimes, Often and Very Often.

**Research Question 2**

**Cyberbullying\_YesorNo:** This is a dichotomous Independent variable created by recoding variable “Cyberbullying\_Freq”. The only values this variable can take are Yes or No. It is used to signify whether a person has been cyberbullied or not. Participants who responded as having “Never” experienced cyberbullying were recoded as “No” and all others who responded as “Rarely, sometimes, often and very often” were coded as “Yes”.

**Cyberbullying\_Response\_Scale:** This is a reliable scale constructed out of the below mentioned items which are rated on the following scale – Strongly Disagree, Disagree, Neither Agree nor Disagree, Agree and Strongly Agree.

SeeCyberbullying\_2: When I see someone being bullied, I will defend them

SeeCyberbullying\_3: I’m more likely to defend someone being bullied if I can post anonymously

SeeCyberbullying\_5: I think it’s important to defend people who are being bullied for no reason

SeeCyberbullying\_6: I feel good about myself when I defend someone else who is being bullied

The scale was constructed by averaging the above mentioned variables SeeCyberbullying\_2**,** SeeCyberbullying\_3, SeeCyberbullying\_5, and SeeCyberbullying\_6 to create our dependent variable **“Cyberbullying\_Response\_Scale**”. Reliability analysis was also performed on the above mentioned items and scale variable generated is reliable.

**Research Question 3**

**Sex:** Identifies the sex of the respondent. It is an independent variable at two levels- Male and Female.

**Cyberbullying\_Freq**: Indicates the frequency of being cyberbullied. The values it can take are Never, Rarely, Sometimes, Often and Very Often.

**Research Question 4**

**Nationality:** This is dichotomous independent variable created by recoding variable “Nationality” into **“Nationality\_response**”. The only values this variable can take are “American or non-American”. This has been recoded in this way considering that majority of the participants in the survey are white Americans and a significantly less number were either Indian or of other nationality.

**Cyberbullying\_activities\_scale:** This is a reliable scale constructed out of the below mentioned items which are rated on the following scale – Never, Rarely, Sometimes, Often and Very Often. Bullying has been associated with cyberbullying as it has been found out in many research that a bully in school is much more likely to be a cyberbully (Li, 2007).

BullyingActivities\_Tease: Teasing another student in front of others

BullyingActivities\_NameCalling: Calling another student names or using negative language to describe them

BullyingActivities\_Laughing: Laughing at jokes other students made about someone else

BullyingActivities\_Threats: Repeatedly making threats

BullyingActivities\_Rumors: Spreading rumors

BullyingActivities\_Attacking: Attacking someone physically or verbally

BullyingActivities\_Excluding: Excluding someone from a group on purpose.

The scale was constructed by averaging the above mentioned variables BullyingActivities\_Tease**,** BullyingActivities\_NameCalling, BullyingActivities\_Laughing, BullyingActivities\_Threats, BullyingActivities\_Rumors, BullyingActivities\_Attacking, and BullyingActivities\_Excluding to create our dependent variable **“Cyberbullying\_activities\_scale**”. Reliability analysis was also performed on the above mentioned items with Cronbach’s alpha computed as .911 which is highly reliable.

**WORKING WITH DATA**

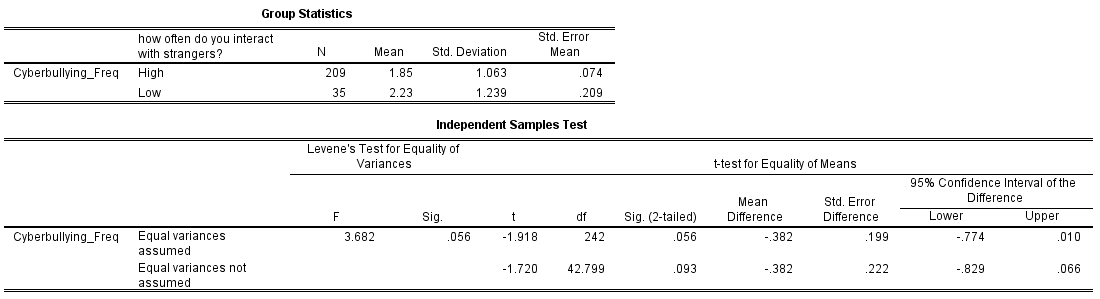
This process involved checking normality of all variables that we used for our analyses. We ran descriptive statistics on all variables to check levels of Skewness, Kurtosis, and Frequency of variables and plotted histograms with normal curve for the same. We concluded that all variables approximate a normal distribution. We did not come across variables that had a significant amount of missing data such that the statistical tests could not be run. We also ensured that any assumptions for the tests conducted were not violated with regards to use of categorical or continuous data.

**ANALYSIS**

**Research Question 1: Test Used: Two Sample Independent t-test**

From the output of the tests run in SPSS (Figure 1), the Levene’s test for equality of variances obtained for this sample is 0.56 which is non-significant and we report with the assumption that equal variances have been assumed. With a significance level of 0.05, we obtained a P value of 0.056 and hence fail to reject the null hypothesis that increased interaction with strangers does not lead to an increase in being cyberbullied. The number of participants where high interaction with strangers was found was 209 and low interaction with strangers was 35. We conclude that we found participants who interacted less with strangers (2.23±1.239) had a statistically significant higher exposure to being cyberbullied than those with higher interaction (1.85±1.063), t(242) = -1.918, P = 0.056.

The effect size (Cohen’s d) calculated works out to -0.329 which indicates a small effect understandably due to a P value of 0.056. The Power calculated corresponding to this effect size is 0.560. In order to obtain a power of 0.8 and thus avoid a type-2 error, we require a total sample size comprising of 230 participants with 115 participants belonging to each group.



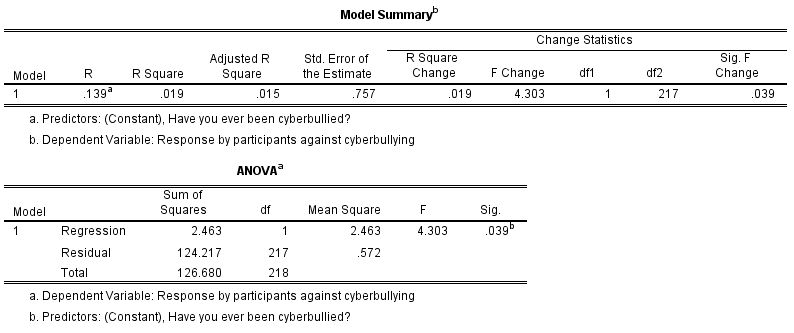
**Figure 1.** SPSS output for RQ1 analysis

**Research Question 2: Test Used: Linear Regression**

From the output of tests run in SPSS (Figure 2), the R value is calculated as 0.139 which shows a small correlation. The R Square value indicates that the independent variable **“Cyberbullying\_YesorNo”** explains 1.9% of the variability of the dependent variable **“Cyberbullying\_Response\_Scale”**. Adjusted R square provides us with an estimate of the effect size which is 0.015 which qualifies as a very small effect.

The Regression model is statistically Significant as **F(1,217) =4.303., P = 0.039**. This indicates that the model is statistically significant enough to predict the dependent variable **“Cyberbullying\_Response\_Scale”**.

The standardized co-efficient Beta was calculated as **0.212, P=0.039**. Based on this t-statistic, we reject the Null hypothesis that the independent variable “Cyberbullying\_YesorNo” does not predict the dependent variable “Cyberbullying\_Response\_Scale”.

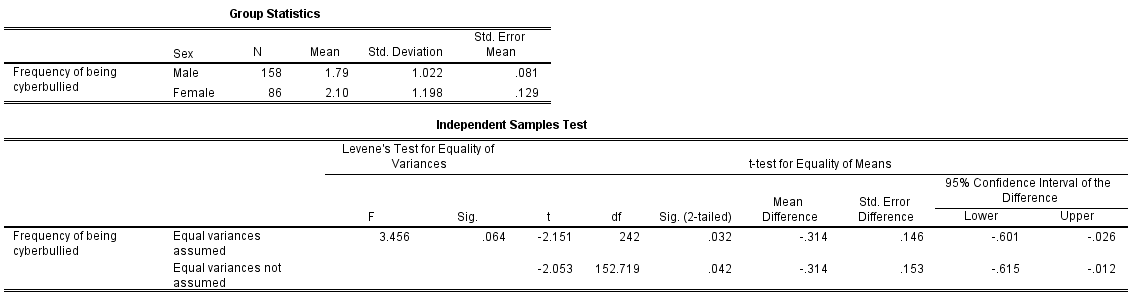


**Figure 2.** SPSS output for RQ2 analysis

**Research Question 3: Test Used: Independent Sample T Test**

From the tests run in SPSS (Figure 3), the Levene’s test for equality of variances obtained for this sample is 0.06 which is non-significant and we report with the assumption that equal variances have been assumed. With a significance level of 0.05, we obtained a P value of 0.032 and hence reject the null hypothesis that there is no gender difference when it comes to individuals being cyberbullied. The number of participants who were males were 158 and females 26. We conclude that we found female participants had a statistically significant higher exposure to cyberbullying **(2.10±1.198)** than male participants **(1.79±1.022), t(242) = -2.151, P=0.032.**

The effect size (Cohen’s d) calculated works out to **-0.285** which indicates a medium effect. The Power calculated corresponding to this effect size is **0.682**. In order to obtain a power of 0.8, we require a total sample size comprising of 306 participants.

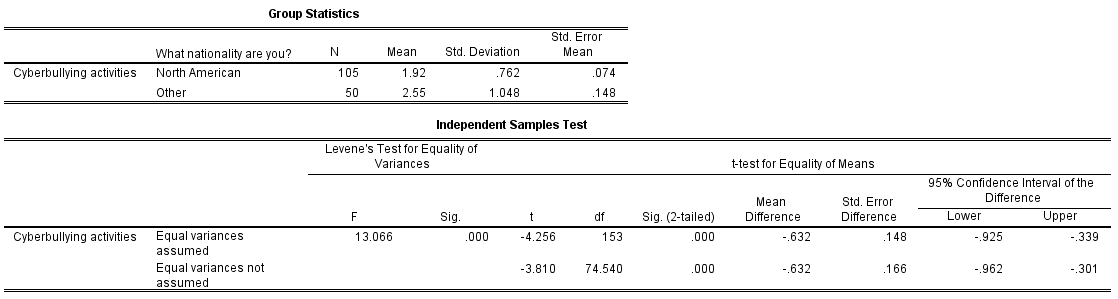


**Figure 3.** SPSS output for RQ3 analysis

**Research Question 4: Test Used: Independent Sample T Test**

The Levene’s test for equality of variances obtained for this sample is significant and we report with the assumption that equal variances have not been assumed (Figure 4). With a significance level of 0.05, we obtained a P value less than 0.001 and hence reject the null hypothesis that there is no difference in the nationality of an individual when it comes to engaging in activities that could be termed as cyberbullying. The number of participants who were North American were 105 and those of other nationality were 50. We conclude that we found participants belonging to other nationality group (2.55±1.048)have a higher tendency to engage in cyberbullying activities than North Americans (1.92±0.762), **t(153) = -4.256, P < 0.001**

The effect size (Cohen’s d) calculated works out to **-0.687** which indicates a large effect. The Power calculated corresponding to this effect size is a high value of **0.99**.



**Figure 4.** SPSS output for RQ4 analysis

**INTERPRETATION OF FINDINGS**

The first research question to an extent contributes to the discussion of whether interacting with strangers serves as an invitation to being cyberbullied. Both Ask.fm and Formspring work on a question and answer based forum with the ability to record answers anonymously as strangers. What we found is that there increased interaction with strangers does not necessarily lead to an increase in cyberbullying. However, this was based only on a marginal difference in the P value obtained which was 0.56. We can also say that the test suffered from limitations of having an uneven sample size with 209 participants calculated as having high interaction and 35 participants with low interaction. This analysis did not definitively provide evidence of effect of interaction of strangers with cyberbullying, arguably because of the P value bordering on the significance level of 0.05. Further research questions along the same lines can be more focused towards obtaining a sample that is truly representative of the population which includes individuals who make use of platforms where anonymity is allowed in the collaboration that takes place online.

The second research question helped us fortify the notion that we already hold which is victims of cyberbullying are more likely to help other cyberbullying victims. The linear regression test we ran provided us with a statistically significant result to conclude that victims do help other victims by being supportive towards them. This finding will especially be helpful for individuals who take a stand on cyberbullying online. To know that you are not alone in a fight that has plagued the internet is in itself reassuring to take a stance on fighting back. Individuals who are afraid of being more active online for fear of being cyberbullied can take assurance from the fact that there are those out there willing to support and defend them when in need. Social media websites could also modify their framework of working such that victims of cyberbullying should be able to contact other victims and a black list be created for known cyberbullies.

The third research question is one of the most important questions to consider and we obtained a statistically significant result for the same. Knowing that there is a gender difference in the frequency of being cyberbullied and by finding that there is directionality involved with females being more susceptible to cyberbullies helps us understand if factors such as male dominance or teenage aggression in males have any role to play in cyberbullying. It would be interesting to further study if the cyberbullying experienced by females is conducted primarily by males or equally by both genders. In a research conducted on gender differences found “25.8% of females reported that they were cyberbullied as compared to only 16.8% of boys” (Hinduja & Patchin, 2010). A general notion that stems from traditional bullying is that boys are likely to bully girls in high school (Sampson, 2009) and the same trend progresses on in the cyber world. However, it is important to conduct further research to fully understand the source of bullying for female victims and the technology that is used to carry out the same. Research along these lines have already been conducted but to do it from use of different kinds of technology perspective may bring about interesting results.

Our findings for the fourth research question helped us cement the theory that difference in one’s culture, country where an individual belongs from and their upbringing will have an impact on the way they engage in cyberbullying activities. We obtained a statistically significant result for the same. However, the directionality obtained stating individuals who were “Indians or Other” were more likely to commit cyberbullying activities than “North Americans” was significant. This can help in constituting further research on the same with a sample size that is truly representative of the population in terms of demographics. If future findings lean more towards a particular nationality that in itself can lead to further study targeting a specific sample to better understand the nature of cyberbullying. In order to take action like setting policies for cyberbullying prevention based on what the findings reveal further will prove to be difficult considering the sensitivity attached to an individual by putting his/her nationality in question to prevent cyberbullying.

**CONCLUSION**

Through our research we tried finding conclusive results to the different questions arising in the cyberbullying domain. From the analysis and results we can conclude that interacting with strangers does lead to Cyberbullying, but frequency of interaction with strangers does not increase the frequency of being cyberbullied. But, there is gender difference in the frequency of being cyberbullied. According to the findings, Females are more vulnerable to cyberbullying as compared to males. Another important finding from our research is that people of nationality other than North Americans have a higher probability of being a cyberbully. Also, we can conclude that Cyberbullying victims are more likely to support other cyberbullying victims by posting supportive comments on their page. We hope our research can parallel initiatives to explore the human and online elements that lead to cyberbullying. We hope to build on this in future iterations to get an even deeper insight into the social science behind this problem.

**REFERENCES**

Agatston, P. W., Kowalski, R., & Limber, S. (2007). Students’ perspectives on cyber bullying. *Journal of Adolescent Health*, *41*(6), S59-S60.

Berger, K. S. (2003). *The developing person through childhood and adolescence*. Macmillan.

Davison, C. B., & Stein, C. H. (2014). The Dangers of Cyberbullying. *North American Journal Of Psychology*, *16*(3), 595-606.

Feinberg, Ted, and Nicole Robey. "Cyberbullying." *Principal Leadership* 9.1 (2008): 10-14.

Hinduja, S., & Patchin, J. W. (2010). Resources. Cyberbullying Research Center - findings, stories, cases, downloads, fact sheets, tips and strategies, news headlines, a blog, and other helpful resources

Lenhart, A. (2008). *Teens, Online Stranger Contact & Cyberbullying: What the Research is Telling Us--*. Pew Internet & American Life Project.

Li, Q. (2006). Cyberbullying in schools a research of gender differences. *School psychology international*, *27*(2), 157-170.

Li, Q. (2007). New bottle but old wine: A research of cyberbullying in schools. *Computers in human behavior*, *23*(4), 1777-1791.

Patchin, Justin W., and Sameer Hinduja. "Cyberbullying and self‐esteem\*."*Journal of School Health* 80.12 (2010): 614-621.

Paullet, K., & Pinchot, J. (2014). Behind the screen where today's bully plays: perceptions of college students on cyberbullying. *Journal of information systems education*, *25*(1), 63.

Puhl, R. M., Peterson, J. L., & Luedicke, J. (2013). Weight-based victimization: Bullying experiences of weight loss treatment–seeking youth. *Pediatrics*, *131*(1), e1-e9.

Raskauskas, J., & Stoltz, A. D. (2007). Involvement in traditional and electronic bullying among adolescents. *Developmental psychology*, *43*(3), 564.

Sampson, R. (2009). Bullying in schools

Vitak, Jessica. (2014). Cyberbullying [Data Set].