Pakistan Journal of Engineering and Technology, PakJET
Multidisciplinary & Peer Reviewed
Volume: 04, Number: 01, Pages: 145- 150, Year: 2021

Analysis of Twitter Related to Eating and Mental Disorders

Misbah Iqbal, Akmal Khan, Noor-e-Aqsa, Muhammad Omar
Department of Computer Science & IT, The Islamia University Bahawalpur, Pakistan
Corresponding Author: Akmal Khan (e-mail: akmal.shahbaz@iub.edu.pk)

Abstract- **The study aims to examine the Eating disorder (ED) and Mental disorder (MD) related tweets that increase the volume and distribution of Twitter conversations and explore what content produces high attention on Twitter. We explore 35,440 ED and 14,800 MD-related Tweets from 18 December 2019- 14 January 2020. Topic modeling (LDA) is applied to identify different topics discussed in the tweets. Our analysis identifies eleven topics, which are group into five main themes (Education and Treatment, Social Media, Personnel Experience, Mentally Suffer, MD types) for MDs and six main themes (Body Image, ED types, Education and Treatment, Social Media, Personnel Experience, Stigma attitude) for ED. The top topic category is Treatment and Support in MD and Education and Treatment in ED.**

***Index Terms*— Eating disorder, Mental disorder, Social media, Twitter, Topic model**

I. INTRODUCTION

On social media, Twitter and Facebook where users share their thoughts, ideas, and health-related experiences which promote public health behavior [1], [2]. Previous research investigated the impact of infectious diseases such as Zika, and influenza on various local people. Using Twitter sheds light on public opinion regarding Zika with a focus on identifying emerging topics in the negative hearing category [3]. However, previous studies have explored their ideas and thoughts related to eating disorders (EDs) [4]–[6] and mental disorders (MDs) [7], [8] with a Twitter content analysis.

Public health authorities and information providers used Topic Models on social media to efficiently extract themes and clustering of related tweets [9]. Research on the inclusion of topics on Twitter has been done by some researchers such as finding current articles and events on Twitter [10], [11]. Sasaki et al. [12] proposed a new way to extend the Twitter-LDA approach to optimizing well-modeled tweets. Hidayatullah et al. [13], [14] detected topics on road traffic and football news in Indonesia. Mamidi et al. [3] focused on identifying emerging negative sentiments on Zika related topics, Recently, Abd-Alrazaq et al. [15] identify key topics related to the COVID-19 epidemic. Our contribution is applying a text-mining algorithm to analyze ED and MD Twitter data.

- Topic modeling (LDA) to identify the latent topics in tweets.
- Word clouds to describe the major themes of EDs and MDs-related tweets.

This study aims to explore Eating and Mental disorder-related tweets that increase the volume and distribution of Twitter conversations and explore what content produces high attention on Twitter. We use the text mining analysis method to observe (1) what latent topics related to ED and MD can we detect from the Tweets? (2) What are the major themes of EDs and MDs related-tweets?

II. METHODS

A. RESEARCH DESIGN

Our research consists of four main steps. Data collection, data preparation, implementation of text mining approaches (Topic Modeling and Word Cloud), and evaluation.

B. DATA COLLECTION

We used the Twitter advance search tool to collect EDs and MDs Tweets with NCapture [16]. The study period was from 18 December 2019 to 14 January 2020. The following information was collected from every single tweet (1) Tweet related information (a)date and time; (b) the number of retweets; (c) hashtags; and (d) quotes. (2) User details (a) username; (b) brief bio; and (c) location.

Recovered tweets were found in CSV files. We reviewed the date and time of each tweet to ensure that all existing tweets are timely and targeted.

C. PRE-PROCESSING THE RAW DATASET

Before preprocessing Retweets and non-location tweets are removed. In pre-processing, the corpus is converted into lower case. Punctuations, numbers, stopwords, URLs, and emojis are removed. Lemmatization is applied to tweets. Lastly, we tokenize the extracted raw text.

D. DATA ANALYSIS

1) TEXT MINING ANALYSIS

For the mining of data, we have chosen to use Topic Modeling.

a) LATENT DIRICHLET ALLOCATION

Latent Dirichlet allocation (LDA) is a reproductive probabilistic model of a corpus. LDA is an unsupervised soft clustering algorithm that views documents as bags of words. The pseudocode is as follows:

For each Document $d \in \{1,\ldots,N\}$:
 For each word w in document d :
 Generate $z \sim p$ (z | d)
 Generate $w \sim p$ (w | z)

The processed tweets were analyzed using the topic modeling technique [17] identifying the most common topics in tweets We used the Latent Dirichlet Analysis algorithm (LDA) from the Python sklearn package. We have selected eleven to be the topics for using the LDA. We have taken the top names representing each of the eleven topics produced by the LDA [18].

b) WORD CLOUD

In the content of tweet conversation, the processed tweets are analyzed using single-word combinations (unigram) and double-word combinations (bigrams) and are displayed in word clouds to identify major themes. [19].

III. RESULTS

A. TWITTER DATA

We collected tweets on EDs and MDs as shown in FIGURE 1. A total of 35,440 Eating Disorders (EDs) and 14,800 Mental Disorders (MDs) tweets were retrieved from Twitter over the 4 weeks (18 December 2019 to 14 January 2020). After removing Non- location tweets of EDs (4,863) and MDs (2,125) we get EDs tweets (30,557) and MDs (12, 675) location tweets. After removing retweets on EDs (22,851) and MDs (9,120), we have unique 7,726 EDs tweets (22%) and 3,555 MDs tweets (24%) were included in the analyses. We perform a text-mining analysis of unique tweets.

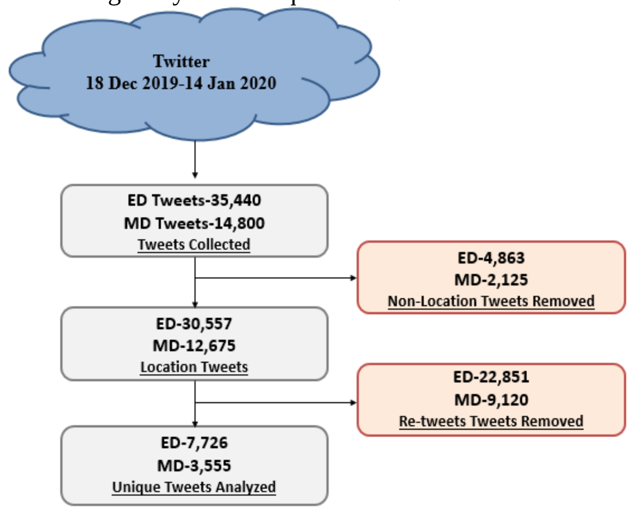


FIGURE 1. Flowchart of the selection of tweets.

B. TEXT MINING OF EATING DISORDERS AND MENTAL DISORDERS

1) TOPIC MODELING (LDA)

After pre-processing of the data set, we obtain 7,726 tweets on EDs and 3,555 tweets on MDs to feeding into the topic model. EDs tweets contain 1,041 unique topics with their Coherence Score: 0.44. MDs tweets contained 5,75 unique tokens with their Coherence Score: 0.50. We implement the topic models on both EDs and MDs with topic numbers set to eleven. The Coherence score is shown in FIGURE 2.

a) DISTRIBUTION AND EVALUATION OF EATING DISORDER TOPICS

The eleven topics further analyzed are produced by LDA. We summarize the eleven different topics and grouped these topics (similar themes) into six categories: 'ED types', 'Education & Treatment', 'Body-Image', 'Social Media', 'Personnel Experience, 'Stigma attitude'. FIGURE 3 shows the topics of each category and TABLE I list all the topics, representative keywords, and topic categories.

In TABLE I, the percentage shows the accuracy of topic-related ED tweets. We noticed that the maximum number of tweets belonging to the 1st topic and the focus on Education and Treatment. We compare our study with another ED content analysis [20], in which 18,288 tweets were manually reviewed from June 1, 2012, to January 31, 2018. We found high consistency on 'Education & Treatment', 'Body-Image', and 'Social Media topics'. ED Types, Personnel Experience, Stigma attitude are not found in the previous study [20].

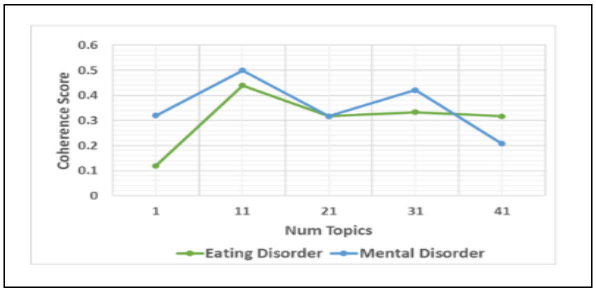


FIGURE 2. Coherence score for a various number of topics on EDs and MDs.

b) DISTRIBUTION AND EVALUATION OF MENTAL DISORDER TOPICS

The eleven topics further analyzed are produced by the LDA. We summarize the eleven different topics and grouped these topics (similar themes) into five categories: ‘Education and Treatment’, ‘Social Media’, ‘Personnel Experience, ‘Mentally Suffer’, ‘MD types’. FIGURE 4 shows the topics of each category and in TABLE II, a list of all the topics, representative keywords, and topic categories.

In TABLE II, the percentage shows the accuracy of topic-related MD tweets. We noticed that the maximum number of tweets belonging to the 2nd topic and the focus on Treatment and support.

We compare our study with another MD content analysis [21], in which 132 tweets were manually reviewed from September 2015 to November 2015. We found high consistency on Education & Treatment, Personnel Experience topics. Social Media, Mentally suffers, MD types are not previous study [21].

2) WORD CLOUD-BASED ANALYSIS OF EDs AND MDs TWEETS

The majority of high-frequency uni-grams and bi-grams reflect the most important words in eating disorders tweets. In, FIGURE 5 the word cloud of “eating disorders” shows that most people are suffering from anorexia and bulimia and they are also struggling to lose weight to not feel fat and also not see themselves as overweight. They need help and support to recover from ED. FIGURE 6 the word cloud of “mental disorders” shows that people describe their depression and anxiety. Many people are diagnosed with serious mental illness and they get treatment through medication and therapy to recover from MD.

TABLE I.
EATING DISORDER TOPICS ACCORDING TO THEIR CATEGORY.

Topic Categories	Topics	Representative Keywords	%
Education & Treatment	Education& Treatment	Recovery, Treatment, Supportive	16.5%
Body-Image	Body image	Body dysmorphia, promoting, shaming	14.6%
	Fitness	Body, body image, diet	13%
ED Types	Types	Eating disorder, anorexia, bulimia, binge	8.2%
Social Media	Media disappointment	Social medium, depression anxiety, body shamming	6.9%
	Media support	Social media, awareness, promote	4.8%
Personnel Experience	Suffering	Self-harm, young women, suffer,	6.1%
	health	Mental health, Weight	9.4%
	struggling	Struggling with, control, tackling fast	5.5%
Stigma attitude	Warning	Mental illness, trigger warning	2.6%
	Negative	Joke about, feel like, behavior	1.5%

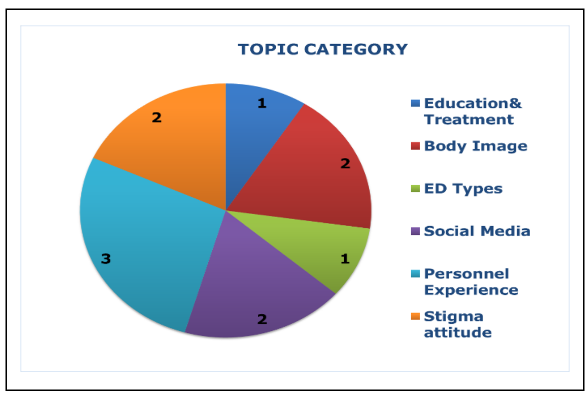


FIGURE 3. Topics distribution with different categories of EDs.

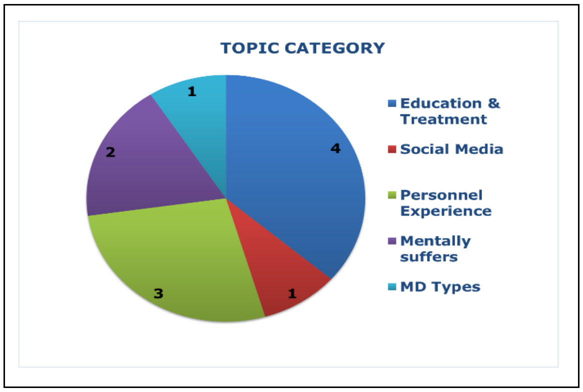


FIGURE 4. Topics distribution with different categories of MDs.

IV. DISCUSSION AND CONCLUSION

Social media is a place where people discussed their experiences, interactions, and get information. The results show that people use Twitter to communicate about their experiences related to EDs and MDs and share information and advice on the issue. While most tweets were about EDs (n = 7,726), fewer related to MDs (n=3,555). Our study analyzed data from Twitter to evaluate the EDs and MDs-related content of web platforms. By using topic modeling (LDA), generates eleven different topics for both EDs and MDs, spanning five content areas for MDs, including "Education and Treatment", "Social Media", "Personnel Experience", "Mentally Suffer", "MD types" and spanning six content areas for EDs, including "Body Image", "ED types", "Education, and Treatment", "Social Media", "Personnel Experience", "Stigma attitude".

Similarities with previous research in this area [4], [5] and found a large portion of web content replicate psychological understanding of EDs, as well as "eating and weight concerns", "bulimic symptoms" "thinness", "weight-loss behaviors", and "body checking" [22] and mental disorders and the desire for a "sense of community", "raising awareness and combatting stigma", "safe space for the expression", and "coping and empowerment"[21].

This suggests that MDs behave similarly online in the same way in the real world. These results described prevention and interventions aimed at eating disorders orientation and the content of mental disorders on communication platforms.

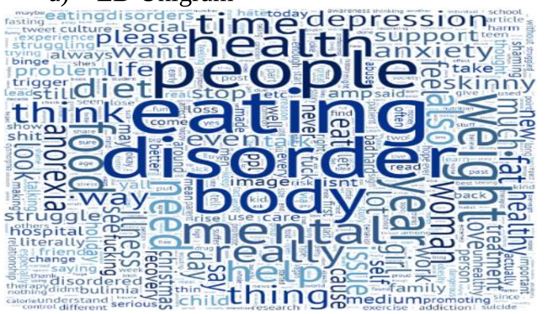
Our analysis identified topics that were not previously described, providing the behavioral dissemination of EDs and MDs. For example, topics related to EDs and MDs reflect on suffering, struggling, and self-health as a means of personnel experience that relates to earlier research related to self-report [23]. Also, some previous articles seem to draw attention to the negative behaviors of EDs and MDs. By being given that communication that reflects bad moral values always before a change of behavior [24]. Our studies show MDs and EDs treatment and support in tweets. Most posts promoted recovery, sooner than illness [25].

TABLE II.

MENTAL DISORDER TOPICS ACCORDING TO THEIR CATEGORY.

Topic Categories	Topics	Representative Keywords	%
Education & Treatment	MD Research	medical, diagnosed, treatment, research, support	7.4%
	Treatment & support	cause, suggest, symptom, treatment, positive	11%
	MD Diagnostics	Mental disorder, Diagnostic, study, Addiction	5.7%
	Medication	Severe, medication, therapy, diagnose	4.7%
Social Media	Media disappointment	Social, understand, fucking, change	4.7%
Personnel Experience	Health	Person, illness, behavior, Depression	3.7%
	Struggling	People, trying, mentally, continue	4.3%
	Suffering	Suffering, experience, emotional, Syndrome	2%
Mentally suffers	Mental distressing	Suffers, problem, trauma, addiction	3%
	Mental condition	Health, condition, serious, disability	4%
MD types	Types	Depression, anxiety, disease, schizophrenia, bipolar	5.7%

a) ED Unigram



b) ED Bigram

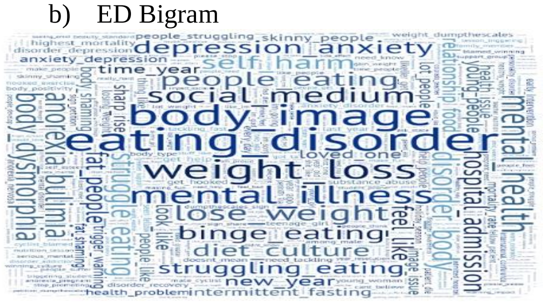
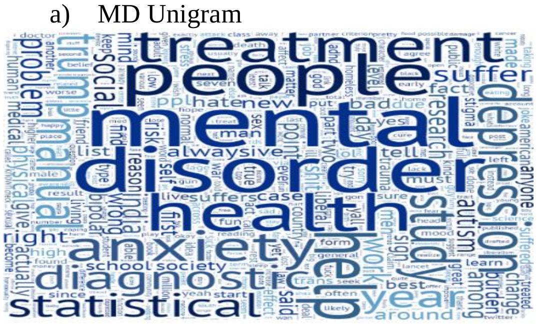


FIGURE 5. Unigram cloud (a) and Bigram cloud (b) of EDs-related tweets.

a) MD Unigram



b) MD Bigram

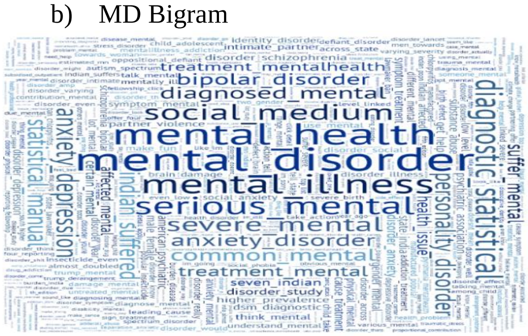


FIGURE 6. Unigram cloud (a) and Bigram cloud (b) of MDs-related tweets.

REFERENCES

[1] M. J. Paul and M. Dredze, "You Are What You Tweet : Analyzing Twitter for Public Health," *Artificial Intelligence.*, vol. 38, pp. 265–272, 2011.

[2] J. Luo, "Exploring temporal suicidal behavior patterns on social media : Insight from Twitter analytics," *Health Informatics*, vol. 26, no. 2, pp. 738–752, 2019, DOI: https://doi.org/10.1177/1460458219832043.

[3] R. Mamidi, M. Miller, T. Banerjee, W. Romine, and A. Sheth, "Identifying key topics bearing negative sentiment on Twitter: Insights concerning the 2015-2016 Zika epidemic," *Journal of Medical Internet Research*, vol. 21, no. 6, pp. 1–18, 2019, DOI: 10.2196/11036.

[4] A. Arseniev-Koehler, H. Lee, T. McCormick, and M. A. Moreno, "#Proana: Pro-Eating Disorder Socialization on Twitter," *Journal of Adolescent Health*, vol. 58, no. 6, pp. 659–664, 2016, DOI: 10.1016/j.jadohealth.2016.02.012.

[5] P. A. Cavazos-Rehg *et al.*, "'I just want to be skinny.': A content analysis of tweets expressing eating disorder symptoms," *PLoS ONE*, vol. 14, no. 1, pp. 1–11, 2019, DOI: 10.1371/journal.pone.0207506.

[6] D. B. Branley and J. Covey, "Pro-ana versus pro-recovery: A content analytic comparison of social media users' communication about eating disorders on Twitter and Tumblr," *Frontiers in Psychology*, vol. 8, no. AUG, pp. 1–8, 2017, doi: 10.3389/fpsyg.2017.01356.

[7] Y. Wang, "Detecting Signals of Associations between Dietary Supplement Use and Mental Disorders from Twitter," *2018 IEEE International Conference on Healthcare Informatics Workshop (ICHI-W)*, no. 2018, pp. 1–5, 2019, doi: 10.1109/ICHI-W.2018.00016.

[8] A. Budenz, A. Klassen, J. Purtle, E. Yom Tov, M. Yudell, and P. Massey, "Mental illness and bipolar disorder on Twitter: implications for stigma and social support," *Journal of Mental Health*, vol. 29, no. 2, pp. 191–199, Mar. 2020, doi: 10.1080/09638237.2019.1677878.

[9] M. Gomide, Janaína & Veloso, Adriano & Meira Jr, Wagner & Almeida, Virgilio & Benevenuto, Fabrício & Ferraz, Fernanda & Teixeira, "Dengue surveillance based on a computational model of spatio-temporal locality of Twitter," *Proceedings of the 3rd International Web Science Conference*, pp. 1-8., 2011, doi: 10.1145/2527031.2527049.

[10] A. Ottesen, S. Jonas, F. Therkelsen, and B. Gambäck, "Twitter Topic Modeling by Tweet Aggregation," *Proceedings of the 21st Nordic Conference of Computational Linguistics*, no. May, pp. 77–86, 2017.

[11] M. S. C. Sapul, T. H. Aung, and R. Jiamthapthaksin, "Trending topic discovery of Twitter Tweets using clustering and topic modeling algorithms," *Proceedings of the 2017 14th International Joint Conference on Computer Science and Software Engineering, JCSSE 2017*, 2017, doi: 10.1109/JCSSE.2017.8025911.

[12] K. Sasaki, T. Yoshikawa, and T. Furuhashi, "Online topic model for twitter considering dynamics of user interests and topic trends," *EMNLP 2014 - 2014 Conference on Empirical Methods in Natural Language Processing, Proceedings of the Conference*, pp. 1977–1985, 2014, doi: 10.3115/v1/d14-1212.

[13] A. F. Hidayatullah, E. C. Pembrani, W. Kurniawan, G. Akbar, and R. Pranata, "Twitter Topic Modeling on Football News," *2018 3rd International Conference on Computer and Communication Systems, ICCCS 2018*,

pp. 94–98, 2018, doi: 10.1109/CCOMS.2018.8463231.

[14] A. F. Hidayatullah and M. R. Ma'Arif, "Road traffic topic modeling on Twitter using latent dirichlet allocation," *Proceedings - 2017 International Conference on Sustainable Information Engineering and Technology, SIET 2017*, vol. 2018-Janua, pp. 47–52, 2018, doi: 10.1109/SIET.2017.8304107.

[15] A. Abd-Alrazaq, D. Alhuwail, M. Househ, M. Hamdi, and Z. Shah, "Top Concerns of Tweeters During the COVID-19 Pandemic: Infoveillance Study," *Journal of Medical Internet Research*, vol. 22, no. 4, p. e19016, 2020, doi: 10.2196/19016.

[16] G. L. Passerello, J. E. Hazelwood, and S. Lawrie, "Using Twitter to assess attitudes to schizophrenia and psychosis," *BJPsych Bulletin*, vol. 43, no. 4, pp. 158–166, 2019, doi: 10.1192/bjb.2018.115.

[17] D. M. Blei, A. Y. Ng, and M. I. Jordan, "Latent Dirichlet Allocation," *Journal of Machine Learning Research 3*, vol. 3, pp. 993–1022, 2003, doi: https://ai.stanford.edu/~ang/papers/jair03-lda.pdf.

[18] S.-W. Kao and P. Luarn, "Topic Modeling Analysis of Social Enterprises: Twitter Evidence," *Sustainability*, vol. 12, no. 8, p. 3419, 2020, doi: 10.3390/su12083419.

[19] M. Tiggemann, O. Churches, L. Mitchell, and Z. Brown, "Tweeting weight loss: A comparison of #thinspiration and #fitspiration communities on Twitter," *Body Image*, vol. 25, pp. 133–138, 2018, doi: 10.1016/j.bodyim.2018.03.002.

[20] S. Zhou, Y. Zhao, R. Rizvi, J. Bian, A. F. Haynos, and R. Zhang, "Analysis of Twitter to Identify Topics Related to Eating Disorder Symptoms," *2019 IEEE International Conference on Healthcare Informatics, ICHI 2019*, pp. 1–4, 2019, doi: 10.1109/ICHI.2019.8904863.

[21] N. Berry, F. Lobban, M. Belousov, R. Emsley, G. Nenadic, and S. Bucci, "#WhyWeTweetMH: Understanding why people use Twitter to discuss mental health problems," *Journal of Medical Internet Research*, vol. 19, no. 4, 2017, doi: 10.2196/jmir.6173.

[22] O. Whooley, "Diagnostic and Statistical Manual of Mental Disorders (DSM)," 2014.

[23] L. Behavior, "Self-Criticism as a Transdiagnostic Process in Nonsuicidal Self-Injury and Disordered Eating : Systematic Review and Meta-Analysis," *Suicide and Life-Threatening Behavior*, pp. 1–18, 2018, doi: https://doi.org/10.1111/sltb.12436.

[24] K. Beckman *et al.*, "Mental illness and suicide after self-harm among young adults: Long-term follow-up of self-harm patients, admitted to hospital care, in a national cohort," *Psychological Medicine*, vol. 46, no. 16, pp. 3397–3405, Dec. 2016, doi: 10.1017/S0033291716002282.

[25] K. Saffran, E. E. Fitzsimmons-, A. E. Kass, D. E. Wilfley, C. B. Taylor, and M. Trockel, "Facebook Usage Among Those Who Have Received Treatment for an Eating Disorder in a Group Setting," *International Journel of Eating disorders*, pp. 1–14, 2016, doi: https://doi.org/10.1002/eat.22567.