

AN ABSTRACT OF THE DISSERTATION OF

Mandy M. Greaves for the degree of Doctor of Philosophy in Counseling presented on December 3, 2018.

Title: A Corpus Linguistic Analysis of Public Reddit and Tumblr Blog Posts on Non-Suicidal Self-Injury

Abstract approved:

Cass Dykeman

Self-harm is a frightening phenomenon that affects adolescents and young adults all over the world. While non-suicidal self-injury (NSSI) has a high prevalence rate, it is not a new behavior. One of the earliest documented cases of NSSI was in ancient Greece and it has gained popularity amongst adolescents. NSSI content appears consistently on pro-NSSI social media and blog websites. Social media's ubiquity has profound impacts on people's behavior; young people in particular spend hours consuming social media every day. The combination of NSSI behavior and social media creates debate and controversy. While some have argued that social media creates a community of people struggling with similar issues, others argue that social media is a space to perpetuate negativity and rumination. Even though research has attempted to understand and explain NSSI behavior, there has been little research on NSSI behavior and related language on social media. This dissertation examines selected pro-NSSI social media to explore the linguistic processes and psychological processes active in pro-NSSI social media posts.

This study, which used a synchronic corpus linguistic design, has two arms; the first arm (A) focuses on pro-NSSI posts on Tumblr and the second arm (B) examined pro-NSSI posts on Reddit. Both the study on Tumblr and the study on Reddit asked the following questions: (1) What is the NSSI specific content word use of the individuals making posts? (2) What is the pattern of use of linguistic processes of the individuals making posts compared to the pattern that appears in blogs overall? (3) What is the pattern of use of psychological processes of the individuals making posts compared to the pattern that appears in blogs overall? The study used an API to collect public, pro-NSSI posts from both Reddit and Twitter that appeared between January 1, 2017 and December 31, 2017; these posts were then anonymized and analyzed. The API was instructed to scrape posts that: (a) monolingual English and (b) contain at least one term on a NSSI terms list compiled by the researchers. The API was instructed to eliminate user names, URLs, hashtags, location of the posts, posts from outside of the US, photographs, and foreign languages. When added together, the posts collected by the API for Research Arm A (Tumblr) had a total 330,868 total word count. For Research Arm B (Reddit), the corpus' total word count was 228,172. The data was analyzed using the Linguistic Inquiry and Word Count (LIWC) software.

For the first research question, the percentage of all words for each NSSI-specific category was reported. In addition, the raw count for each word in each category was listed in rank order. In terms of the second and third research questions, the following were reported for all categories except for emotional tone: (a) percentage of all words, (b) log likelihood statistic of the comparison of the NSSI

corpus to general blog norms, and (c) log ratio of the comparison of the NSSI corpus to general blog norms. For the emotional tone category, the following were noted: (a) percentage of all words, (b) one sample z test of proportion with study results as \hat{p} and general blog norms as p_0 . When the Tumblr posts from Research Arm A were analyzed, the findings revealed that people posting their NSSI behaviors on public blogs are people who have a high desire to connect with others while sharing their pain and experience of numbness, and use high levels of negative emotional words. Overall, these results demonstrated that people posting on Tumblr about NSSI are hurting individuals reaching out for help. The findings for Research Arm B (Reddit) suggest that the posts were made in the mental state that was characterized by high levels of depression, anxiety, and neuroticism. When both research arms are considered as a whole, the results illustrated a profile of a young person who is struggling with managing intense negative emotions. The results from this study have implications for the clinical treatment by informing the counseling theories used to treat this population, transforming how counselors are trained in NSSI behaviors and investigating how research can further explore the nuances of language to have a better understanding of a difficult problem the plagues so many young people.

©Copyright by Mandy M. Greaves

December 3, 2018

All Rights Reserved

A Corpus Linguistic Analysis of Public Reddit and Tumblr Blog Posts on Non-Suicidal Self-Injury

by

Mandy M. Greaves

A DISSERTATION

submitted to

Oregon State University

in partial fulfillment of
the requirements for the
degree of

Doctor of Philosophy

Presented December 3, 2018
Commencement June 2019

Doctor of Philosophy dissertation of Mandy M. Greaves presented on December 3, 2018.

APPROVED:

Major Professor, representing Counseling

Dean of the College of Education

Dean of the Graduate School

I understand that my dissertation will become part of the permanent collection of Oregon State University libraries. My signature below authorizes release of my dissertation to any reader upon request.

Mandy M. Greaves, Author

ACKNOWLEDGEMENTS

The author expresses sincere appreciation to Dr. Cass Dykeman, Dissertation Chair and Doctoral Advisor. You have been a guiding light in this dissertation process. I could not have done this without your constant positive encouragement, support and help. It is because of your constant encouragement that I was able to complete my dissertation and PhD. In moments when it was difficult to continue, and I doubted myself, you helped to inspire me to be persistent. Thank you for believing in me, even when I didn't believe in myself.

CONTRIBUTION OF AUTHORS

Cass Dykeman assisted with the methodology and research design in addition to refining the narrative document. He also assisted with statistical computation of the data in this dissertation.

TABLE OF CONTENTS

Chapter 1: Introduction	1
Study Rationale	1
Organizational Plan for Chapter 1	2
Description of Manuscript #1	5
Manuscript rationale	5
Target journal for publication	5
Statement of research questions	5
Description of methodology	6
Description of Manuscript #2	6
Manuscript rationale	6
Target journal for publications	7
Statement of research questions	7
Description of methodology	7
Specialized Glossary	7
Thematic Links Between Phases of the Research	10
Dissertation Organization	11
References	12
Chapter 2: A Research Manuscript	16
Abstract	18
Overview	19
Method	26

TABLE OF CONTENTS (Continued)

Design	26
Corpus	26
Measures	28
Data Analysis	30
Results.....	31
Discussion.....	31
Limitations	35
Implications.....	36
References.....	40
Chapter 3: A Research Manuscript	48
Abstract.....	50
Overview.....	49
Method	59
Design	59
Corpus	59
Measures	61
Data Analysis	64
Results.....	64
Discussion.....	65
Limitations	69

TABLE OF CONTENTS (Continued)

Implications.....	70
References.....	73
Chapter 4: A General Conclusion.....	81
Overview.....	82
Summary of Manuscript #1	83
Summary of Manuscript #2	86
Thematic Linkage of the Two Manuscripts	88
Contribution to the Knowledge Base	90
Implications of Dissertation for Research Agenda	91
References.....	93
Bibliography	95
Appendix A.....	107
Appendix B:	109

LIST OF TABLES

<u>Table</u>	<u>Page</u>
Table 2.1. Item word counts.....	44
Table 2.2. Results for Linguistic and Psychological Processes.....	45
Table 3.1. Item word counts.....	77
Table 3.2. Results for Linguistic and Psychological Processes	78

Chapter 1: Introduction

#Selfharm is a growing trend on social media and blog sites. This trend is troubling because it reflects the real-life and hidden phenomena occurring among adolescents and young adults in large numbers. Non-suicidal self-injury (NSSI) is frequent among adolescence and young adults. Recently, since the social media boom, NSSI content has become increasingly common on the internet. This study will use data from pro-NSSI Tumblr blogs and Reddit posts to identify and analyze the language that people are using to describe their NSSI behavior.

Study Rationale

Adolescents live their life through social media, which has become a popular focus for many young people. Adolescents average nine hours a day on social media and watching online videos (Tsukayama, 2015). Information is being conveyed faster and more conveniently through social media micro-blogs like Twitter and through image sharing platforms like Instagram. Adolescents (and others) use social media to communicate with their friends and family, meet new people and even find job opportunities. Since social media is the realm where young people spend the majority of their time, it is important for mental health clinicians to not only be aware of this realm, but also to actively discuss it in counseling sessions.

While there have been many studies on mental health and social media and researchers have used linguistic analysis to study the language people use on social media, little attention has been paid to social media and NSSI. Most of the studies that have applied a corpus linguistic analysis approach to social media posts have studied depression, anxiety, postpartum depression and eating disorders. NSSI is still

a relatively underappreciated area of research; previous studies focused on NSSI have attempted to understand why adolescents engage in NSSI and investigated why adolescents post their NSSI pictures on social media (Brown et al., 2018). This study aims to explore the specific language of NSSI and what psychological markers are involved in the language of the pro-NSSI social media post. Because of this study's unique approach, it provides a different perspective on the needs of individuals who engage in NSSI and can be used to help mental health professionals improve treatment protocols.

Organizational Plan for Chapter 1

This chapter contains six sections to provide a background for this study. In order of presentation, these sections are as follows: (1) definition of NSSI, (2) overview of how NSSI usually begins in adolescence, (3) pro-NSSI blog posts, (4) the influence of social media and blogs, (5) mental health and the use of the internet and (6) exploring the use of LIWC as a new method of understanding language online. These sections will be followed by a section that describes the organizational plan for the remaining dissertation.

There are varying definitions of non-suicidal self-injury (NSSI). In one of the most widely cited studies on the subject, Favazza (1998) defined NSSI as “the deliberate, non-suicidal destruction of one’s own body tissue” (p. 260). Favazza’s definition includes an important phrase that would change the how the mental health field viewed NSSI: “non-suicidal”. Favazza argued that people who engage in NSSI are using it to cope with personal distressing feelings. Trichotillomania (hair pulling), excoriation (skin picking) and overdoing were later added to the definition of NSSI.

Studies show that NSSI usually begins in early adolescents (Healthyplace.com, 2016). It is difficult to obtain exact statistics because hospitals do not have separate classification systems to distinguish between NSSI and suicidal behavior. According to the statistics that researchers are able to verify, one in 12 adolescents utilize NSSI (Allen, 2011). Other studies show that 90% of people who do engage in NSSI begin this behavior in their pre-adolescence (Healthyplace.com). There are many reasons why someone might engage in NSSI, but one of the most common reasons is so that they can cope with distressing feelings; adolescents are at greater risk for using NSSI because they lack emotional regulation, which makes it more difficult for individuals to manage their strong emotions (Van Vliet & Kalnins, 2011).

Many young people no longer watch the news and instead learn about world and local events through blogs and microblogs, particularly Twitter and Instagram. This is inherently dangerous, since micro-blogs are filled with opinions (Li & Du, 2014). Furthermore, Chancellor (2016) argues that visual images, like those shared on social media, influence and perpetuate distorted thinking. The use of images and the prevalence of opinions regarding NSSI behavior in pro-NSSI posts can potentially influence others to continue and increase their NSSI behaviors. Pro-NSSI blogs and social media pages can encourage “better” methods of self-injury and instruct affected individuals on how to cover their scars from others.

Social media and blogs’ influence on individuals, particularly adolescents, can be seen in news media coverage. Adolescents engage in social media challenges that can be dangerous and lethal, such as the most recent destructive challenge, the Tide

Pod Challenge. When a dangerous challenge becomes a trend, people are left questioning how this can happen and how it can spread to so many people. One possible explanation of the influence of social media is social contagion, or the influence of behaviors across a community (Aalal, 2018). A classic, well-documented example is youth suicide clusters, in which exposure to suicidal behavior is positively associated with the emergence clusters of suicidal behavior and cluster suicides (Jarvi, Jackson, Swenson & Crawford, 2013). This is particularly dangerous for adolescents, who are especially susceptible to imitation due to their low self-regulation and low impulse control (Jarvi et al., 2013). While this study does not examine suicidal behavior, it is a well-understood demonstration of the influence that social contagion has on the behaviors of young people. In the case of individuals engaged in NSSI, social media has the potential to influence young people to either begin engaging in NSSI or continue and escalate their NSSI behaviors.

As participation in social media increases, mental health experts started to conduct research on mental health markers in language expression. The study by Choudhury, Counts and Horvitz (2013) was one of the first to find that people with mental health issues have psychological markers in their language. Their research suggested that it is possible to identify people who are struggling with mental health issues based on how they post on social media platforms like Instagram. In order to examine whether social media positively or negatively impacts on mental health, O'Reilly et al. (2018) studied self-reports from adolescents on the topic of social media. The majority of adolescents and young adults reported that social media gives them a community of people who are struggling with similar issues. However, there

is much more evidence that demonstrates that social media's effects are more likely to be damaging than positive (Mars et al., 2015; O'Reilly et al., 2018).

One relatively new methodological tool for conducting this kind of research is corpus linguistics. In this method, a corpus is constructed using large amounts of data. For the purpose of this study, the data was collected from two public social media websites with pro-NSSI content, Tumblr and Reddit. The created corpus is then inputted into a computer software to analyze the data in ways that human beings are not able to. This study used the Linguistic Inquiry and Word Count (LIWC 2017). For the purpose of this study LIWC was used to explore three main domains: (1) specific NSSI word content, (2) linguistic process, and (3) psychological process.

Description of Manuscript #1

Manuscript rationale. Data from public pro-NSSI Tumblr posts will be used to analyze the NSSI specific words used, the linguistic properties of the texts and the psychological properties found in the language of the posts.

Target journal for publication. *Journal of Adolescent Mental Health* aims to contribute to an inclusive knowledge base. Specifically, they seek to address child and adolescent mental health studies.

Statement of research questions. The following research questions will be addressed regarding the language of NSSI on public social media pages.

1. What is the NSSI specific content word use of the individuals making posts?
2. What is the pattern of use of linguistics processes of the individuals making posts?

3. What is the pattern of use of psychological processes of the individuals making posts?

Description of methodology. This study employed a synchronic corpus linguistic design for exploring public Tumblr posts on NSSI (Weisser, 2017). The Linguistic Inquiry and Word Count (LIWC) was used to analyze the Tumblr posts containing pro-NSSI content. Tumblr is considered a micro-blog website in which individuals can post on a blog whatever topic they would like. In an attempt to limit pro-NSSI, pro-eating disorder and pro-suicide posts, Tumblr has banned all posts related to these topics. However, adolescents have discovered ways around the ban and continue to post on these topics using hidden hashtags and slag terms.

For the first research question regarding the NSSI specific word content the following categories were analyzed: Methods of NSSI, Cutting-Specific Terms, NSSI Terms, Instruments Used, Reasons for NSSI, and Hidden Hashtag Terms. The second research question regarding linguistic processes the following categories were analyzed: 1st person singular, 1st person plural, 3rd person singular and 3rd person plural. The third research question regarding the psychological processes were analyzed using the following categories: emotional tone, negative emotions, anxiety, anger, sadness and risk focus.

Description of Manuscript #2

Manuscript rationale. Data from public pro-NSSI Reddit posts will be used to analyze the NSSI specific words used, the linguistic properties and the psychological properties found in the language of the posts.

Target journal for publication. *Journal of Clinical Psychology* aims to improve research, especially in the area of psychotherapy effectiveness. This journal is a multidisciplinary journal that covers a broad array of topics. The journal also strives to publish articles that focus on interventions, training, current issues and psychological research.

Statement of research questions. The following research questions will be addressed regarding the language of NSSI on public social media pages.

1. What is the NSSI specific content word use of the individuals making posts?
2. What is the pattern of use of linguistics processes of the individuals making posts?
3. What is the pattern of use of psychological processes of the individuals making posts?

Description of methodology. This study employed a synchronic corpus linguistic design for exploring public Reddit posts on NSSI (Weisser, 2017). The Linguistic Inquiry and Word Count (LIWC) was used to analyze the Reddit posts containing pro-NSSI content. The analyzed variables were the following linguistic categories: 1st person singular pronouns, 3rd person singular pronouns, 1st personal plural pronouns, 3rd person plural pronouns, emotional tone, negative emotion, anxiety, anger, sadness and risk. The following categories are specific to NSSI: methods of NSSI, cutting-specific terms, NSSI terms, instruments used, reasons for NSSI, and hidden hashtag terms.

Specialized Glossary

This study will include specialized use of terms that might be unfamiliar to the reader. These terms are defined here for the purposes of clarity:

Corpus/Corpora

“A collection of text, now usually in machine-readable form and compiled to be representative of a particular kind of language and provided with some kind of annotation” (essex.ac.uk, 1998).

Corpus Linguistics

“Depending on the author, may mean either (a) any approach to language that uses corpus data and methods, or (b) an approach to linguistics that uses corpus methods but does not subscribe to principles of the so-called *corpus-driven* approach” (McEnery & Hardie, 2012, p. 241).

Greaves NSSI Dictionary

Custom list of compiled of NSSI specific terms.

Hashtag

“Originally designed for categorizing posts, the hashtag can now be a tool for a supplementary coy or witty comment. The word tag can mean ‘a word or phrase used for a description or identification.’ Hash is short for hash mark, a term for what we more commonly call a pound sign.” (Merriam-webster.com, n.d.)

Lexical Variants (LV)

“Lexical variants are different word forms for the same expression. These forms may derive from spelling or grammatical variation or from abbreviated formats” (Zeng, 2003).

LIWC

“LIWC is a transparent text analysis program that counts words in psychologically meaningful categories. Empirical results using LIWC demonstrate its ability to detect meaning in a wide variety of experimental settings, including to show attentional focus, emotionality, social relationships, thinking styles, and individual differences” (Tausczik & Pennebaker, 2010, p. 24).

Log Likelihood Test

“A significance test similar to the chi-square test, but generally considered more reliable, especially when working with small values” (McEneaney & Hardie, 2012, p. 246).

Non-suicidal self-injury (NSSI)

“The deliberate, non-suicidal destruction of one’s own body tissue” (Favazza, 1998, p. 260). People who engage in NSSI harm themselves without suicidal intent. Behavior usually includes cutting, burning and picking one’s skin also known as excoriation disorder in the DSM-5.

Out-of-vocabulary (OOV) Words

“It includes neologisms and proper nouns like *hopeable* or *WikiLeaks* which have not made their way into the dictionary” (Han & Baldwin, 2011, p. 370).

Pro-NSSI

Positivity toward or encouragement of non-suicidal self-injury.

Reddit

Reddit calls itself the “front page of the internet”. It is a website that is a “collection of forums, where people can share news and content or comment on other

people's posts. Reddit is broken up into over a million communities known as 'subreddits,' each covering a different topic" (Nichol, 2018, para. 1).

Synchronic

"Relating to the study of a language or languages as they exist at a particular moment in time, without reference to how they might change over time. A *synchronic corpus* contains texts drawn from a single period- typically the present or the very recent past" (McEnery & Hardie, 2012, p. 251).

Textese

"A form of written language as used in text messages and other digital communications, characterized by many abbreviations and typically not following standard grammar, spelling, punctuation, and style" (dictionary.com, n.d.)

Tumblr

"It is a cross between a social networking site (like Facebook and Twitter) and a blog. It is often described as 'microblog' as people usually post short snippets of text and quick snaps as opposed to longer diary style entries found in more traditional blogs. Tumblr appears to have a younger user demographic in comparison with other social networks" (Webwise.ie, n.d.).

Thematic Links Between Phases of the Research

These studies both seek to better the understanding of the linguistic process and the NSSI terms that are used for pro-NSSI public blog posts. This research will be conducted in two manuscripts. The first manuscript, Chapter 2, will utilize data from public pro-NSSI Tumblr posts. The data will be analyzed to better understand specific language markers for pro-NSSI posts. The second manuscript, Chapter 3, will

utilize data from public pro-NSSI Reddit posts. The data will be analyzed to understand NSSI better with a focus on studying NSSI specific language, linguistic process and psychological process language markers. This research will fill a gap in counseling literature by providing data and insight into NSSI behavior, the thought processes of those who engage in NSSI and will provide more information for the treatment of individuals who use NSSI to cope with distressing feelings.

Dissertation Organization

This dissertation is divided into four chapters. Chapter 1 provides a general introduction that includes an overview of the major research on the topics of NSSI and explains major themes that have emerged in earlier research. Chapter 2 is an original manuscript that uses data from public Tumblr posts to identify NSSI specific terms, linguistic processes and psychological processes in the language of the posts. Chapter 3 is a separate original manuscript that uses data from public Reddit posts to identify NSSI specific terms, linguistic processes and psychological processes in the language of the posts.

References

- Aalai, A. (2018, June 8). *The social contagion of suicide*. Retrieved from <https://www.psychologytoday.com/us/blog/the-first-impression/201806/the-social-contagion-suicide>
- Allen, J. E. (2011, November 17). 1 in 12 Teens cut, deliberately hurt themselves. *ABC News*, ABC News Medical Unit. Retrieved from <https://abcnews.go.com/Health/MindMoodNews/12-teens-cut-harm-themselves/story?id=14969232>
- Brown, R. C., Fischer, T., Goldwich, A. D., Keller, F., Young, R., & Plener, P. L. (2018). #cutting: Non-Suicidal Self-Injury (NSSI) on Instagram. *Psychological Medicine*, 48, 337-346. <https://dx.doi.org/10.1017/S0033291717001751>
- Chancellor, S., Lin, Z., Goodman, E. L., Zerwas, S., & Choudhury, M. D. (2016). Quantifying and predicting mental illness severity in online pro-eating disorder communities. *Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work & Social Computing - CSCW 16*, 1171-1184. <https://dx.doi.org/10.1145/2818048.2819973>
- Choudhury, M. D., Counts, S., & Horvitz, E. (2013). Major life changes and behavioral markers in social media. *Proceedings of the 2013 Conference on Computer Supported Cooperative Work - CSCW 13*. <https://dx.doi.org/10.1145/2441776.2441937>

Corpus Linguistics. (n.d.). Retrieved from

https://www1.essex.ac.uk/linguistics/external/clmt/w3c/corpus_ling/content/glossary.html

Explainer: What is Tumblr? (2018, August 08). Retrieved from

<https://www.webwise.ie/parents/explainer-what-is-tumblr-2/>

Favazza, A. R. (1998). The coming of age of self-mutilation. *The Journal of Nervous & Mental Disease, 186*, 259-268. <https://dx.doi.org/10.1097/00005053-199805000-00001>

Han, B., & Baldwin, T. (2011). Lexical normalisation of short text messages: Makn sens a #twitter. In *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies-Volume 1* (pp. 368-378). Stroudsburg PA: Association for Computational Linguistics. Retrieved from: <https://dl.acm.org/citation.cfm?id=2002520>

Hashtag. (n.d.). Retrieved from <https://www.merriam-webster.com/dictionary/hashtag>

Jarvi, S., Jackson, B., Swenson, L., & Crawford, H. (2013). The impact of social contagion on non-suicidal self-injury: A review of the literature. *Archives of Suicide Research, 17*, 1-19. <https://dx.doi.org/10.1080/13811118.2013.748404>

Li, F., & Du, T. C. (2014). Listen to me — Evaluating the influence of micro-blogs. *Decision Support Systems, 62*, 119-130. <https://dx.doi.org/10.1016/j.dss.2014.03.008>

Mars, B., Heron, J., Biddle, L., Donovan, J. L., Holley, R., Piper, M., . . . Gunnell, D. (2015). Exposure to, and searching for, information about suicide and self-harm on the Internet: Prevalence and predictors in a population based cohort

of young adults. *Journal of Affective Disorders*, 185, 239-245.

<https://dx.doi.org/10.1016/j.jad.2015.06.001>

McEnery, T., & Hardie, A. (2012). *CORPUS LINGUISTICS: Method, theory and practice*. Cambridge, UK: Cambridge University Press.

Nicol, W. (2018, July 20). What is Reddit? A Beginner's Guide to the Front Page of the Internet. Retrieved from <https://www.digitaltrends.com/social-media/what-is-reddit/>

O'Reilly, M., Dogra, N., Whiteman, N., Hughes, J., Eruyar, S., & Reilly, P. (2018). Is social media bad for mental health and wellbeing? Exploring the perspectives of adolescents. *Clinical Child Psychology and Psychiatry*, 23, 601-613.

<https://dx.doi.org/10.1177/1359104518775154>

Self-injurious Behavior, Self-injury Treatment. (2010, February 22). Retrieved from <https://www.healthypace.com/abuse/articles/self-injurious-behavior-self-injury-treatment>

Tausczik, Y. R., & Pennebaker, J. W. (2009). The psychological meaning of words: LIWC and computerized text analysis methods. *Journal of Language and Social Psychology*, 29, 24-54. <https://dx.doi.org/10.1177/0261927x09351676>

Textese. (n.d.). Retrieved from <https://www.dictionary.com/browse/textese>

Van Vliet, K., & Kalnins, G. (2011). A compassion-focused approach to nonsuicidal self-injury. *Journal of Mental Health Counseling*, 33, 295-311.

<https://dx.doi.org/10.17744/mehc.33.4.j7540338q223t417>

Weisser, M. (2016). *Practical corpus linguistics: An introduction to corpus-based language analysis*. NY: John Wiley & Sons.

Zeng, M. L. (2005). *4. Semantic relationships used in controlled vocabularies.*

Retrieved from <http://marciazeng.slis.kent.edu/Z3919/42equivalency.htm>

Chapter 2: A Research Manuscript

A Corpus Linguistic Analysis of Public Tumblr Blog Posts on Non-Suicidal Self-
Injury

Mandy M. Greaves and Cass Dykeman

Oregon State University

Author Note

Mandy Greaves, Counseling Academic Unit, Oregon State University; Cass Dykeman, Counseling Academic Unit, Oregon State University.

The research contained in this manuscript was conducted under the approval of the Oregon State University Institutional Review Board (Study ID No. 8618) and was part of the first author's dissertation research project.

Correspondence concerning this article should be addressed to Cass Dykeman, Counseling Academic Unit, Oregon State University, 104 Furman Hall, Corvallis, OR 97331-3502. E-mail: dykemanc@onid.oregonstate.edu

Abstract

Rates of non-suicidal self-injury (NSSI) have steadily risen among adolescents and young adults. This study collected pro-NSSI public blog posts from Tumblr and analyzed the content linguistically using LIWC software, examining the NSSI-specific words, the linguistics properties and the psychological linguistic properties. The results align with similar studies conducted on the language markers of mental health. For NSSI specific word categories, the methods of engaging in NSSI was the most frequently used in the Tumblr blogs. This aligns with literature about the need for people to express their emotional pain to others who may understand. The linguistic properties demonstrated some unique results that can be best explained by the numbing feeling that is found in individuals who utilize NSSI and their tendency to avoid expressing painful experiences in a manner that would increase their vulnerability. The psychological properties of these public Tumblr posts were revealed through the dominantly negative emotional tone of the writing, which is also indicative of anyone struggling with severe mental illness. These findings suggest that treatment which specializes in shame and increasing self-compassion, such as Compassion Focused Therapy, would be more beneficial to these individuals who struggle with pain, shame and emotional distress than current treatment programs.

Keywords: NSSI, LIWC, linguistic analysis, mental health, self-injury

A Corpus Linguistic Analysis of Public Tumblr Blog Posts on Non-Suicidal Self-Injury

Adolescents and emerging adults live in a social media world and as mental health professionals we need to meet them where they are at. The rise of social media has dramatically transformed how information is conveyed and received; now, unlike a decade ago, information can be delivered within seconds of an event. Additionally, non-suicidal self-injury (NSSI) and other mental illnesses have been on the rise, especially among adolescents. The combination of NSSI and social media has created confusion about whether social media is more beneficial or more harmful for adolescents. Since Tumblr, a microblogging site, has attempted to eliminate pro-NSSI posts, by banning all pro-NSSI and other forms of self-injury such as pro-eating disorder blog pages, studies have been conducted to learn whether or not censorship can help reduce the spread of pro-NSSI and pro-eating disorder information via social media. Unsurprisingly, for those of us who have teenagers or work with them professionally, these studies have demonstrated that adolescents find a way around the censorship and continue to post pro-NSSI and pro-eating disorder information (Eveleth, 2013).

When the blogging phenomenon began, they were typically journals for other people to read and connect to. Over the years, blogs have been transformed; while they are still a form of self-expression and way to share one's interests and passions, they have become platforms for sharing knowledge on a particular topic, practicing writing skills, earning exposure and building a network (Dekmezian, 2015). New, website-based social media platforms now cater to those who want to have blogs on

various topics. The most widely known and used website is Twitter, a micro-blog site that provides information in bite-sized bits called “tweets.” Other sites, like Tumblr, provide a greater variety of blogging options, including a wide range of topics for blogging and commenting on. While Tumblr has tried to ban pro-NSSI blogs and pro-eating disorder blogs, little has made an impact on the actual posting of these blogs.

Social media and blogs have achieved their highest use rate ever (Smith & Anderson, 2018), and they have become a place where NSSI is frequently discussed. There are many reasons why people use self-injury and why they post about it on personal social media pages and blogs. One study conducted by Harris and Roberts (2013) explored the reasons why adolescents and young adults go to social media and blog sites to read and post about their self-injury. The researchers discovered a number of reasons for individuals that engage in NSSI to engage with these social media websites; they categorized these into five broad themes: (1) “help and support”, (2) “isolation reduction and community engagement”, (3) “distraction and expression”, (4) “help others” and (5) “triggering material/tips.” Harris and Roberts (2013) discovered that self-injury is a very socially isolating behavior since it is performed by one’s self. Other researchers, like Gilbert (2010), add that there is a level of shame that is attached to engaging in self-destructive behaviors. Participants in the Harris and Roberts (2013) study emphasized that these pro-NSSI websites and blogs provide a community of people who understand their experiences; it seems as though these websites and blogposts have both positive and negative aspects. While the studies of Harris and Roberts and others may provide a better understanding about *why* people post on NSSI blogs and social media communities, there is still little

information about the language in which they use and how their use of language describing their experiences can clue mental health therapists into the person's inner workings and thoughts.

Despite how young people are clearly heavily influenced by social media, there is little knowledge about the persuasive and psychological processes occurring within the pro-NSSI space on the internet. There is also very little knowledge about how these individuals talk about their NSSI with others. Without such knowledge, the nature of what mental health interventions can be most useful to this population remains obscure.

After reviewing the literature on both non-suicidal self-injury and blog use, it was clear that there are six important issues to consider in the study: (1) providing a definition of non-lethal self-injury, (2) understanding that NSSI usually begins in adolescence, (3) examining pro-NSSI posts in blogs, especially on Tumblr, (4) the influence of social media and blogs on people's decisions, (5) mental health and the use of the internet, and (6) exploring the use of LIWC as a new method of understanding language online. Finally, the research questions will be detailed.

An examination of relevant scientific literature reveals that there are multiple definitions of and explanations for NSSI. Favazza (1998) is one of the most widely cited resources for self-harm/self-mutilation. According to Google Scholar (2015), this reference has been cited 688 times in the professional literature. Favazza defines "self-mutilation (SM), [as] the deliberate, non-suicidal destruction of one's own body tissue" (p. 260). He concludes by stating, "SM includes compulsive acts such as trichotillomania and skin picking and such episodic acts as skin-cutting and burning"

(p. 259). Skegg's definition of self-harm first appeared in the literature in 2005 and has also been cited widely (448 times; Google Scholar, 2015); Skegg's definition expanded Favazza's definition of self-mutilation/self-harm to include other behaviors, namely overdosing. Furthermore, Skegg argues that self-harm takes place on a continuum of behaviors; all of which are "deliberate" and "self-wounding" behaviors (p. 1471). The difference between NSSI and suicidal behavior is the individual's intent; individuals engaged in NSSI typically aim to cope (non-fatally) with their disturbing feelings, while suicidal behavior aims to end all feelings through the loss of life (Jarvi, Jackson, Swenson & Crawford, 2013).

There are few precise numbers and statistics describing NSSI rates, in part because it is difficult to collect accurate data. of the, the classification system in hospitals do not distinguish between NSSI behaviors and suicidal behavior. In addition to the flawed classification system, not everybody will report their NSSI behaviors to someone, especially a professional. Researchers have estimated that by the age of 14 or 15, approximately one in 12 adolescents engage in NSSI behaviors; this is estimated to be about six to ten percent of all teenagers (Allen, 2011). Newer statistics reveal that about 90% of people who engage in NSSI began during their teen or pre-adolescent years (Healthyplace.com, 2016). These statistics reveal that more research needs to be done to better help adolescents and pre-adolescents before the NSSI behavior becomes worse.

As micro-blogs like Twitter and Tumblr have increased in popularity and frequency of use, they have become the new way that people receive information (Li and Du 2014). Li and Du noted the difference between a micro-blog and blog in their

study, which is that micro-blogs are mostly used as a channel to express one's opinion, which often skews the information being received by so many. Individuals who engage in NSSI who use micro-blogs often, problematically, influence others either to begin or to continue with the viewer's NSSI. A study by Chancellor (2016) noted that the use of visual images in social media, especially Tumblr and Instagram, encourages the continuation of distorted thoughts and behaviors. As social media becomes more popular and the number of pro-NSSI posts increases, this tendency to influence others with opinions and images that continue distorted thoughts and behaviors could potentially cause other vulnerable individuals to either begin using NSSI behaviors to cope with negative feelings or to continue and escalate their current NSSI behaviors.

Social media can influence people's thoughts and behaviors, such as consumer decisions or personal feelings about themselves. This is known as social contagion, which can be loosely defined as "behaviors [that are] spread throughout a community" (Aalal, 2018, para. 3). A prominent example of social contagion is a positive association with exposure to suicidal behavior and peer suicide attempts (Jarvi et al., 2013). In many communities, we see a cluster of adolescent suicides that occur within weeks of each other (Mueller, 2017). This phenomenon of social contagion is supported the theory by Alfred Bandura's social learning theory (1977), which posits that individuals learn through imitating a model. For example, in terms of delinquent behaviors, adolescents appear to be more susceptible to peer influence and social contagion. Studies have found that adolescents are especially susceptible to

social contagion because they have difficulties with self-regulation and thus impulse control is minimal (Jarvi et al., 2013).

As social media has increasingly impacted peoples' lives, it has become an important topic for study in mental health circles. Research has been divided on the topic of whether social media is helpful or harmful to individual's mental health. While adolescents and young adults have reported that social media gives them a community of people struggling with similar issues and as a source of support (O'Reilly et al., 2018), there is much more evidence to support the idea that social media is more damaging to adolescents' and young adult's mental health than helpful. Studies have demonstrated that social media potentially leads to depression and suicidal behavior (Mars et al., 2015; O'Reilly et al., 2018). For individuals who engage in NSSI, there are many positive resources and websites to help someone cope with negative feelings and on how to eliminate NSSI. The dark side, however, is that there are also websites and blogs that promote and encourage NSSI by providing tips on how to self-harm and how to hide self-harm. While this controversy continues, there is little known about the impact of social media on NSSI behavior and still less information about the language that is used to describe NSSI.

As social media emerged and grew in popularity, researchers became interested in how people use language in social media posts and what mental health professionals can learn from their language. As early as 2013, Choudhury, Counts, and Horvitz examined the language in social media posts to identify people who may be suffering with post-partum depression. They noted that language can provide useful "psychological markers" that give useful information into an individual's inner

workings. More recently, researchers have started to make an effort to understand the language of self-harm.

The most frequently used method for studying language is corpus linguistics. First, the researcher constructs a corpus using large amounts of data drawn from online social media, which is subsequently inputted into software for analysis. The software then completes a deeper and more nuanced in-depth analysis than what people are capable of doing (McGlashan, 2018). The Linguistic Inquiry and Word Count (LIWC; Pennebaker et al., 2007) software is the most frequently used tool in corpus linguistics related scholarship; the LIWC analyzes and calculates aspects of language including the use of emotional words, social word use, self-referencing, drives and pronoun use (Adrian et al., 2011). Carefully examining these aspects of language has the potential to provide mental health professionals with further insight into the individual's experience and environment.

Given the promise and demonstrated usefulness of the corpus linguistics method of studying textual content, three research questions were designed to guide this study. The first research question was: What is the NSSI specific content word use of the individuals making posts? The NSSI specific content word use categories were: Methods of NSSI, Cutting-Specific Terms, NSSI Terms, Instruments Used, Reasons for NSSI, and Hidden Hashtag Terms. The second research question was: What is the pattern of use of linguistics processes of the individuals making posts compared to the pattern that appears in blogs overall? The linguistics process categories were: 1st person singular, 1st person plural, 3rd person singular, and 3rd person plural. The third research question was: What is the pattern of use of

psychological processes of the individuals making posts compared to the pattern that appears in blogs overall? The psychological process categories were: emotional tone, negative emotions, anxiety, anger, sadness and risk focus.

Method

Design

This study employed a synchronic corpus linguistic design for exploring public Tumblr posts on NSSI (Weisser, 2017). The Linguistic Inquiry and Word Count (LIWC) was used to analyze the Tumblr posts containing pro-NSSI. The variables were the following linguistic categories: 1st person singular pronouns, 3rd person singular pronouns, 1st person plural pronouns, 3rd person plural pronouns, emotional tone, negative emotion, anxiety, anger, sadness and risk. It also was used to analyze another set of categories that were NSSI-specific: methods of NSSI, cutting-specific terms, NSSI terms, instruments use, reasons for NSSI, and hidden hashtag terms.

The sample was compiled from Tumblr during a year's worth of data ranging from January 1, 2017 to December 31, 2017. The sample was compared to individuals who make Tumblr posts from national norms for social media posts.

Corpus

The researchers used Tumblr's API system to collect public posts that occurred between January 1, 2017 and December 31, 2017 to construct the corpus. The API was instructed to scrape posts that: (a) used monolingual English and (b) were comprised of the list of NSSI terms compiled by the researchers (list available from first author). The API was instructed to eliminate the following while collecting

the text: user names, URLs, hashtags, location of the posts, posts from outside of the United States, photographs and foreign languages.

As a result of these API instructions, over two thousand Tumblr posts were collected and then cleaned. This cleaning included the following:

1. All abbreviations were spelled out (e.g. “max” to “maximum”)
2. UK English word usage and spelling were converted to US English word usage.
3. Lexical Variants (LV)/Textese/Out-of-vocabulary (OOV) words were transformed by lexical normalization
 - a. LV/Textese/OOV Category “Letter” was normalized. Letter refers to instances where letters are missing or there are extraneous letters, but the lexical correspondence to the target word form is trivially accessible (e.g. shuld “should”) (Han & Baldwin, 2011).
 - b. LV/Textese/OOV Category “Number Substitution” was normalized. Number Substitution refers to instances of letter-number substitution, where numbers have been substituted for phonetically-similar sequences of letters (e.g. 4 “for”) (Han & Baldwin, 2011).
 - c. LV/Textese/OOV Category “Letter&Number” was normalized. Letter&Number refers to instances which have both extra/missing letters and number substitution (e.g. b4 “before”) (Han & Baldwin, 2011).

- d. LV/Textese/OOV Category “SpaceDeleted” was normalized. SpaceDeleted refers to where spaces having being deleted between words (e.g. sucha “such a”) (Han & Baldwin, 2011).
- e. LV/Textese/OOV Category “NSSI Slang” was normalized. NSSI Slang refers to words used to hide NSSI related Tumblr posts (Moreno, Ton, Selkie, & Evans, 2016).

After cleaning, the average word count per post was 200. The total word count for the entire corpus was 330,868 words, which provided sufficient material for analysis.

Measures

GNLS- Methods of NSSI. Methods of NSSI are ways in which people engage in non-suicidal self-injury. This is a broad category that includes all methods for self-harming, including biting, burning, erasing, hitting and picking.

GNLS- Cutting-Specific Terms. Cutting is a specific category under methods of NSSI. Since cutting is one of the most common methods of non-suicidal self-injury, it has its own sub-category. Some words in this category are cut, the seed word cut*, and cutting.

GNLS- NSSI Terms. NSSI terms are any terms that describe NSSI. Some examples include non-suicidal, NSSI, suicide and the seed word self-mutil*.

GNLS- Instruments Used. The category includes the tools that people use to engage in NSSI behaviors. Some of these instruments include blade, bleach, eraser, fingernail and laxatives.

GNLS- Reasons for NSSI. There are many reasons why people choose to engage in NSSI behaviors. Some examples of reasons for NSSI are addiction, anger, anxiety, attention, control and feeling empty.

GNLS- Hidden Hashtag Terms. Due to new social media restrictions, people who engage in NSSI behaviors and post about it on social media have become creative in the hashtags they use to communicate with the community of people who engage in NSSI. Some hidden hashtags include selfharmmmmm, mysecretfamily, selfinjuriyyyy, secretsociety_123, ehtilb and mauled.

LIWC- 1st Person Singular Pronoun. According to Pennebaker (2013), people who use 1st person singular pronouns track their focus of attention. People who are depressed tend to “pay more attention to themselves” (Pennebaker, p. 40). Pennebaker has discovered in his research that the more depressed a person is the more they use I-words (Pennebaker, p. 108). Examples of 1st person singular pronoun is “I”, “me”, and “my”. Pennebaker, Boyd, Jordan, and Blackburn (2015) report adequate reliability and validity for this LIWC scale.

LIWC- 3rd Person Singular Pronoun. Third person singular pronouns allow the writer to talk about other people. Common examples are he, she, it, him, or her. Pennebaker, Boyd, Jordan, and Blackburn (2015) report adequate reliability and validity for this LIWC scale.

LIWC- 1st Person Plural Pronoun. According to Pennebaker (2013), first person plural pronouns have multiple meanings. He uses the example of “we”. On the one hand, the use of the word “we” can provide a sense of warmth and community, but on the other hand, there is a more distant meaning of we, which

creates distance. For the second, Pennebaker uses the example of “we really need to take out the trash” (p. 41). This type of “we” is a way to nicely order someone to do something. Common examples are “we”, “us” and “our”. Pennebaker, Boyd, Jordan, and Blackburn (2015) report adequate reliability and validity for this LIWC scale.

LIWC- 3rd Person Plural Pronoun. Examples of third person plural pronouns are “they”, “them” and “theirs”. Pennebaker, Boyd, Jordan, and Blackburn (2015) report adequate reliability and validity for this LIWC scale.

Data Analysis

For the first research question, the raw count for each word in each category will be calculated and the percentage of all words for each NSSI-specific category will be reported. In terms of the second and third research questions, the following will be reported for all categories except for emotional tone: (a) percentage of all words, (b) log likelihood statistic of the comparison of the NSSI corpus to blog general norms, and (c) log ratio of the comparison of the NSSI corpus to blog general norms. The log likelihood analysis requires raw scores; raw scores were calculated by multiplying the percentage of all words and the size of the corpus. The log likelihood formula used was from Cressie and Read (1989) as cited in Rayson and Garside (2010). The log ratio is a measure of effect size (Hardie, 2014). As Hardie notes, a log ratio of 1 means a word is twice as common in a corpus than a comparison corpus. The score for emotional tone is based upon a standard score rather than a raw score. As such, a one sample z test of proportion was used to analyze this category rather than the log likelihood test. All analyses were conducted using Excel. Given the small number of analyses conducted, the standard .05 alpha level was retained.

Results

In the analysis of the first research question, the most frequently used NSSI-specific category was Reasons for NSSI. The percentage of all words were as follows: Methods of NSSI (25.77%), Cutting-Specific Terms (18.67%), NSSI Terms (15.23%), Instruments Used (3.16%), Reasons for NSSI (0%), and Hidden Hashtag Terms (0%). A list of the raw counts for the words or word stems in all categories can be found in Table 1. In terms of the second research question, the linguistic process category that differed most from the general blog norms was first person plural ($LL = 18.59$). The effect size for this category was a log ratio of -1.34. Except for the emotional tone, a list of the results of all the categories can be found in Table 2. In terms of emotional tone, the one sample z test of proportion was significant ($z = -6.727, p < .05$). In terms of the third research question, the psychological process category that differed most from the general blog norms was negative emotion ($LL = 8.80$). The effect size for this category was a log ratio of 0.50.

Discussion

In this study, which examined the language used in pro-NSSI blogs on Tumblr, researchers examined whether there were differences in the use of language in pro-NSSI blogs when they were compared to blogs overall. Specifically, it focused on addressing the following questions about the presence and use of NSSI-specific language in the corpus: What is the NSSI specific content word use of the individuals making posts? What is the pattern of use of linguistics processes of the individuals making posts? What is the pattern of use of psychological processes of the individuals

making posts? Next, results are presented and discussed. Toward the end of the section, limitations and implications of the findings are presented.

In answering the first research question, the researchers sought to understand the specific NSSI words that were used in these blog posts and found that the most frequently used words in pro-NSSI blog posts were the methods that individuals use to engage in NSSI. In this category, the most frequently used words were cut, erase and break. One explanation for the high frequency of these words is the most common method used to engage in NSSI is cutting.

While cutting one's skin is not the only method of self-injury, it is likely one of the most common ones. It is evident in the corpus that these individuals cut themselves in private and when they are in the public domains. It appears that cutting is one of the easiest behaviors to self-injure without people taking notice. It is becoming evident in research that there is a difference in the occurrence of cutting and other methods of NSSI. Studies have found that cutting is the most common method of NSSI and least common among hospital presentations (Arensman et al., 2013). There also appears to be a difference in the gender distribution among methods to engage in NSSI. Cutting is one of the methods that demonstrates a more even gender distribution when compared to other methods such as intentional overdose (Arensman et al., 2013).

Words like cut, erase and break evokes strong feelings of distress and pain. These words may appear with a high frequency because individuals engaged in NSSI may be seeking coping strategies for the intense emotional pain they feel. Skegg (2005) noted that most people who engage in NSSI are attempting to relieve their

distress, to seek help, to cope with their feelings and to gain a sense of control. Williams (2000) theorized that NSSI is more of a cry of pain than a cry for help. From what we understand in the literature, individuals who engage in NSSI feel intense emotional and psychological distress and pain. It makes sense that given the intense feelings and pain that these individuals experience daily, that those would be the most commonly used words while they are describing their experiences. While both reasons are plausible in understanding the specific NSSI word use, a more likely explanation is the common use of the method of cutting to engage in NSSI.

The second research question concerns the linguistic patterns in pro-NSSI blogs. The study data demonstrates that the use of the first-person plural was the most frequent and statistically significant pattern. Having worked with adolescents who engage in NSSI behaviors for over six years this significant pattern of using first person plural was surprising, yet not unusual. One explanation for the frequent use of the first-person plural is the community's involvement in the social media realm. On Tumblr, many of the users appeared to express a sense of support from one another, whether that be emotional support and encouragement or support in continuing and improving their NSSI behaviors; in short, their discourse typically demonstrates that they experience a sense of belonging to something outside of themselves (Pennebaker, 2013). Another explanation for this pattern is the sense of numbness from their trauma that usually accompanies NSSI; Pennebaker (2013) in his book explains that usually after a trauma people often experience numbness and pain. His theory is that NSSI engaged individuals use NSSI to reduce and cope with the pain they experience in the moment and as a way to self-regulate; they intend to divert and

avoid the pain. He states that it is common for people in this kind of pain to avoid using I words because the word I expresses vulnerability. Following Pennebaker's logic of avoiding emotional pain and vulnerability, it is unlikely that this population would choose to express its feelings and desires using first person singular pronouns. In addition to avoiding a sense of vulnerability, the use of we pronouns may also be the result of Tumblr's website protocol. Tumblr is not an anonymous posting website; because Tumblr accounts are linked to one's name and personal email, an individual is more likely to feel exposed. This relative ease of identification and exposure may be an additional reason for individuals engaged in NSSI to avoid using language that is too vulnerable. Of these three possible explanations, the most likely one is the sense of numbness from trauma. NSSI is typically associated with trauma symptoms which causes dissociation and distance from first person singular pronouns. The distancing nature of the first-person plural pronouns makes sense given the symptoms these individuals experience.

The third research question pertains to the psychological word use patterns in pro-NSSI blogs. The most common and statistically significant psychological pattern was the use of negative emotional language. Given the aforementioned literature, this use of negative emotion language would be expected in pro-NSSI blog posts. People who engage in NSSI are in extreme emotional pain; logically, it makes sense that the language they use is going to have a negative emotional tone. Several studies by Pennebaker (2013) suggest that people who use negative emotional language do not benefit from writing about their experiences (i.e. blogs), but instead remain stuck in the negativity, as though ruminating. This explanation suggests that people who post

about their behavior in pro-NSSI blogs and social media continue to ruminate in their negative emotions and has the potential to also ruminate on other's behaviors as well, thus serving to trigger and continue the cycle of NSSI. An alternate explanation for the results obtained for the third research question is that these individuals are experiencing low self-esteem. Pennebaker, Mehl, and Niederhoffer (2003) found that there was a correlation between an individual with low scores in self-esteem assessments and the high use of negative emotion words. Of these two explanations, the former is most likely because while there is a correlation between negative emotion words and low self-esteem, it is a relatively small correlation. These individuals using NSSI and posting on blogs about their behavior are more likely to be ruminating on their negative circumstances and feel that need to connect those negative experiences with others.

The findings from this study reflects the work of many researchers who have explored how the language used by individuals provides insight into their mental health. This study discovered that the language used in pro-NSSI social media posts and blogs demonstrates a level of mental illness; it includes a high frequency of negative emotion language which signals rumination, the construction of a sense of community that may continue the cycle of NSSI behaviors and indicates that those who engage in NSSI may desperately be seeking coping skills and help.

Limitations

There are four major limitations to this study that should be noted. Firstly, linguistic analysis is not perfect. A single word can have many meanings, so it is possible for a word to be categorized as an indicator of self-harm or negative

emotional language when the author is using it in a different context; this could slightly skew the data. In order to control for this known limitation, the data was reviewed before it was inputted into the LIWC software in order to clarify any confusing word meanings. The second limitation to this study was that the study was limited to only English words due to the fact that the researcher only speaks English. This substantially limits the corpus because according to a 2013 study, approximately 21% of people in the United States speak a language other than English in their home (Camarota & Zeigler, 2014). It is possible that some people who post on Tumblr blogs feel more comfortable writing in their native language. The third limitation is that the researcher was limited to using public Tumblr posts to create a corpus, because they chose not to use a private Tumblr account. Tumblr's settings allow their users to control their audience; an individual can post a blog and have the privacy settings set to be either fully public (so that it is visible to users with or without Tumblr accounts), public to members with Tumblr members only, visible to followers only, and private. This study limited itself to Tumblr blogs that were fully public, which did not require a log in, in an effort to respect the inherent privacy of a Tumblr blog post. The fourth limitation of this study is that its results might not be generalizable to the entire population of people who post on pro-NSSI social media and blog sites. Even though the effect size is relatively large for this study, inferences into the data should be done with care.

Implications

Two implications for practice and one implication for research emerged from this study. Many counselors who treat individuals who use NSSI are unaware of

client online behaviors. This was evident in the Tumblr posts, where many individuals commented that their counselors are unaware that they are engaging in self-injury and that they are encouraged to self-injure by others online. Because counselors have the ethical responsibility to assess their clients' safety—including suicidal/ homicidal thoughts and self-injury behaviors, if an adolescent client is self-harming and the counselor has never conducted an assessment, the counselor is liable for not fulfilling their minimum ethical mandate if the client harms themselves seriously enough to be hospitalized.

The first implication for practice is that it can inform the counselor about NSSI behaviors to ask patients about. It is necessary for counselors to know to ask their about NSSI behaviors; clients may not be comfortable bringing up NSSI behavior and may think that it is not appropriate to share in counseling—this is particularly true if the patient is experiencing any level of transference and wants to protect the counselor from their destructive patterns. In order to help counselors better fulfill their ethical mandate and to help them feel more comfortable bringing up NSSI, Cornell created a fact sheet for counselors. Included in the fact sheet were techniques for approaching NSSI, which discussed the importance of discussing online behavior. This is the first resource that encourages specific questions to clients regarding their online NSSI behavior (Ernhout, & Whitlock, 2014).

The second practical implication revolves around treatment choices. Individuals who use NSSI to cope with extreme emotional distress are frequently isolated and are desperate to belong to community where they are not shamed. The

language of shame can be clearly visible in the narratives of the blogs on Tumblr and in the results of this study, when analyzing the frequency of negative emotional tone.

NSSI behaviors, like cutting, are disturbing to any counselor and counselors, rightly, will choose an approach known to be effective in changing destructive behaviors. (i.e., Motivational Interviewing [MI] and Cognitive Behavioral Therapy [CBT]) (Washburn et al., 2012). Washburn et al., (2012) suggest that there is no effective treatment theory that was designed specifically for evidenced based treatment for NSSI. Therefore, they suggest that many counselors choose CBT as an easy to use standardized and evidenced based approach. This study, however, reveals a client profile that could help counselors understand individuals who practice NSSI behaviors; this profile suggests that individuals who practice NSSI have a harsh internal self-critic that is fueled by the shame that surrounds NSSI behaviors. Therefore, it may be necessary to first help the clients “warm-up” to themselves before starting a more action oriented approach.

One theoretical approach that is suggested by this new knowledge is Compassion Focused Therapy (CFT). While CFT shares many basic features with CBT, like exploring and examining triggers and how the brain reacts to said triggers, CFT takes the treatment a step beyond CBT by focusing on shame and self-compassion. In CFT, shame is recognized as a harmful emotion that keeps people in a state of negative emotions and rumination, and the practice of self-compassion can help a patient with severe mental illness to learn better ways to cope with negative emotional reactions (Gilbert, 2010). Neff (2009) adds that increasing self-compassion

reduces anxiety, a leading reason for NSSI behavior, and promotes better emotion regulation.

Another important implication for the research is a better understanding of the language of NSSI. Very few research studies have examined the specific language used by those who engage in NSSI behaviors. There is only a modicum of research using language analyzing software to understand the language characteristics of NSSI, which could have implications for improving treatment. This study adds to the existing research exploring language markers for mental health disorders by expanding the lens and giving voice to a misunderstood population of people.

References

- Aalai, A. (2018, June 8). *The social contagion of suicide*. Retrieved from <https://www.psychologytoday.com/us/blog/the-first-impression/201806/the-social-contagion-suicide>
- Adrian, M., Zeman, J., Erdley, C., Lisa, L., & Sim, L. (2010). Emotional dysregulation and interpersonal difficulties as risk factors for Nonsuicidal Self-Injury in adolescent girls. *Journal of Abnormal Child Psychology*, 39, 389-400. <http://dx.doi.org/10.1007/s10802-010-9465-3>
- Allen, J. E. (2011, November 17). 1 in 12 Teens cut, deliberately hurt themselves. *ABC News*, ABC News Medical Unit. Retrieved from <https://abcnews.go.com/Health/MindMoodNews/12-teens-cut-harm-themselves/story?id=14969232>
- Arensman, E., Larkin, C., Corcoran, P., Reulbach, U., & Perry, I. J. (2013). Factors associated with self-cutting as a method of self-harm: Findings from the Irish National Registry of Deliberate Self-Harm. *The European Journal of Public Health*, 24(2), 292-297. <https://dx.doi.org/10.1093/eurpub/ckt087>
- Camarota, S. A., & Zeigler, K. (2014, October 3). One in five U.S. residents speaks foreign language at home, record 61.8 million. Retrieved from <https://cis.org/One-Five-US-Residents-Speaks-Foreign-Language-Home-Record-618-million>
- Chancellor, S., Lin, Z., Goodman, E. L., Zerwas, S., & Choudhury, M. D. (2016). Quantifying and predicting mental illness severity in online pro-eating disorder communities. *Proceedings of the 19th ACM Conference on*

- Computer-Supported Cooperative Work & Social Computing - CSCW 16*, 1171-1184. <http://dx.doi.org/10.1145/2818048.2819973>
- Choudhury, M. D., Counts, S., & Horvitz, E. (2013). Major life changes and behavioral markers in social media. *Proceedings of the 2013 Conference on Computer Supported Cooperative Work - CSCW 13*.
<https://dx.doi.org/10.1145/2441776.2441937>
- Cressie, N., & Read, T. R. (1989). Pearson's χ^2 and the loglikelihood ratio statistic G2: a comparative review. *International Statistical Review/Revue Internationale de Statistique*, 57, 19-43. <http://dx.doi.org/10.2307/1403582>
- Dekmezian, G. (2016, September 23). Why do people blog? The benefits of blogging. Retrieved from https://www.huffingtonpost.com/gary-dekmezian/why-do-people-blog-the-be_b_8178624.html
- Ernhout, C., & Whitlock, J. (2014). *Bringing up self-injury with your clients*. The Fact Sheet Series, Cornell Research Program on Self-Injury and Recovery. Cornell University, Ithaca, NY.
- Eveleth, R. (2013, July 22). *What happened to 'self-harm blogs' after Tumblr banned them?* Retrieved from <https://www.smithsonianmag.com/smart-news/what-happened-to-self-harm-blogs-after-tumblr-banned-them-15883320/>
- Favazza, A. R. (1998). The coming of age of self-mutilation. *The Journal of Nervous & Mental Disease*, 186, 259-268. <https://dx.doi.org/10.1097/00005053-199805000-00001>
- Gilbert, P. (2010). *Compassion focused therapy*. NY: Routledge.

- Han, B., & Baldwin, T. (2011). Lexical normalisation of short text messages: Makn
sena a #twitter. In *Proceedings of the 49th Annual Meeting of the Association
for Computational Linguistics: Human Language Technologies-Volume 1* (pp.
368-378). Stroudsburg PA: Association for Computational Linguistics.
Retrieved from: <https://dl.acm.org/citation.cfm?id=2002520>
- Hardie, A. (2014). Log Ratio—an informal introduction. Lancaster, UK: Centre for
Corpus Approaches to Social Science. Retrieved from
<http://cass.lancs.ac.uk/log-ratio-an-informal-introduction/>
- Harris, I. M., & Roberts, L. M. (2013). Exploring the use and effects of deliberate
self-harm Websites: An Internet-Based Study. *Journal of Medical Internet
Research, 15*. <http://dx.doi.org/10.2196/jmir.2802>
- Jarvi, S., Jackson, B., Swenson, L., & Crawford, H. (2013). The impact of social
contagion on non-suicidal self-injury: A review of the literature. *Archives of
Suicide Research, 17*, 1-19. <http://dx.doi.org/10.1080/13811118.2013.748404>
- Lewis, S. P., & Seko, Y. (2015). A double-edged sword: A review of benefits and
risks of online non-suicidal self-injury Activities. *Journal of Clinical
Psychology, 72*, 249-262. <http://dx.doi.org/10.1002/jclp.22242>
- Li, F., & Du, T. C. (2014). Listen to me — Evaluating the influence of micro-blogs.
Decision Support Systems, 62, 119-130.
<http://dx.doi.org/10.1016/j.dss.2014.03.008>
- Mars, B., Heron, J., Biddle, L., Donovan, J. L., Holley, R., Piper, M., . . . Gunnell, D.
(2015). Exposure to, and searching for, information about suicide and self-
harm on the Internet: Prevalence and predictors in a population based cohort

of young adults. *Journal of Affective Disorders*, 185, 239-245.

<http://dx.doi.org/10.1016/j.jad.2015.06.001>

McGlashan, M. (Ed.). (2018). Corpus: Some key terms. *CASS: Briefings*. Retrieved from <http://cass.lancs.ac.uk/wp-content/uploads/2013/12/CASS-Gloss-final1.pdf>

Moreno, M. A., Ton, A., Selkie, E., & Evans, Y. (2016). Secret Society 123: understanding the language of self-harm on Instagram. *Journal of Adolescent Health*, 58, 78-84. <https://doi.org/10.1016/j.jadohealth.2015.09.015>

Mueller, A. S. (2017). Does the media matter to suicide?: Examining the social dynamics surrounding media reporting on suicide in a suicide-prone community. *Social Science & Medicine*, 180, 152-159.

<http://dx.doi.org/10.1016/j.socscimed.2017.03.019>

Neff, K. D. (2009). The role of self-compassion in development: A healthier way to relate to oneself. *Human Development*, 52, 211-214.

<http://dx.doi.org/10.1159/000215071>

O'Reilly, M., Dogra, N., Whiteman, N., Hughes, J., Eruyar, S., & Reilly, P. (2018). Is social media bad for mental health and wellbeing? Exploring the perspectives of adolescents. *Clinical Child Psychology and Psychiatry*, 23, 601-613.

<http://dx.doi.org/10.1177/1359104518775154>

Pennebaker, J. W. (2013). *The secret life of pronouns: What our words say about us*. NY: Bloomsbury Press.

- Pennebaker, J.W., Boyd, R.L., Jordan, K., & Blackburn, K. (2015). *The development and psychometric properties of LIWC2015*. Austin, TX: University of Texas at Austin. <http://dx.doi.org/10.15781/T29G6Z>
- Pennebaker, J.W., Chung, C.K., Ireland, M., Gonzales, A., & Booth, R.J., (2007). *The development and psychometric properties of LIWC2007*. Austin, TX: University of Texas at Austin.
- Rayson, P., & Garside, R. (2000). Comparing corpora using frequency profiling. In *Proceedings of the workshop on comparing corpora-Volume 9* (pp. 1-6). Stroudsburg, PA: Association for Computational Linguistics. <http://dx.doi.org/10.3115/1117729.1117730>
- Healthyplace.com. (2016). *Self-injurious behavior, self-injury treatment*. San Antonio, TX: Author. Retrieved from <https://www.healthyplace.com/abuse/articles/self-injurious-behavior-self-injury-treatment>
- Skegg, K. (2005). Self-harm. *Lancet*, 366, 1471-1483. [https://doi.org/10.1016/S0140-6736\(05\)67600-3](https://doi.org/10.1016/S0140-6736(05)67600-3)
- Smith, A., & Anderson, M. (2018, September 19). *Social Media Use 2018: Demographics and Statistics*. Washington, DC: Pew Research Center. Retrieved from <http://www.pewinternet.org/2018/03/01/social-media-use-in-2018/>
- Tausczik, Y. R., & Pennebaker, J. W. (2009). The psychological meaning of words: LIWC and computerized text analysis methods. *Journal of Language and Social Psychology*, 29, 24-54. <http://dx.doi.org/10.1177/0261927x09351676>

- Washburn, J. J., Richardt, S. L., Styer, D. M., Gebhardt, M., Juzwin, K. R., Yourek, A., & Aldridge, D. (2012). Psychotherapeutic approaches to non-suicidal self-injury in adolescents. *Child and Adolescents Psychiatry and Mental Health*, 6(14). <https://doi.org/10.1186/1753-2000-6-14>
- Weisser, M. (2016). *Practical corpus linguistics: An introduction to corpus-based language analysis*. NY: John Wiley & Sons.
- Williams, J. M. G., & Pollock, L. R. (2000). The psychology of suicidal behaviour. In K. Hawton and K. van Heeringen K (Eds.), *The international handbook of suicide and attempted suicide* (pp. 79-93). Chichester, UK: John Wiley & Sons. <https://doi.org/10.1002/9780470698976.ch5>

Table 1

Item Word Counts

Word/Stem	Methods	Cutting Terms	NSSI Terms	Instru- ments Used	Reasons for NSSI	Hidden Hashtag Terms	Word Ct.
Cut		X					585
Cutt*		X					543
Cutting		X					146
Anxiety					X		62
Control					X		98
Blade				X			119
Burning	X		X				96
Anger					X		110
Picking	X		X				93
Attention					X		69
Hitting	X		X				85
Addiction					X		30
Bite	X						27
Fingernail				X			8
Biting			X				15
Empty					X		68
Eraser				X			362
Mauled						X	0
Bleach				X			4
Erasing	X		X				106
ehilb						X	0
Laxatives				X			0
mysecretfamily						X	0
secretsociety_123						X	0
selfharmmmm						X	0
selfinjuryyy						X	0

Table 2

Results for Linguistic and Psychological Processes

Category	Process Type	Tumblr Word Count	Blog Norm Word Count	Over/Under Use	Log Likelihood	Log Ratio
1st Person Singular	Linguistic	21772.47	200.72	+	0.29	0.05
Negative Emotion	Psychological	9747.37	66.05	+	8.80**	0.50
1st Person Plural	Linguistic	1205.86	29.18	-	18.59**	-1.34
Anxiety	Psychological	1373.34	8.66	+	1.73	0.60
Sad	Psychological	2244.24	14.11	+	2.86	0.61
Risk	Psychological	2009.77	14.75	+	1.13	0.38
Anger	Psychological	2847.17	21.80	+	1.16	0.32
3rd Person Plural	Linguistic	2344.73	21.80	+	0.02	0.04
3rd Person Singular	Linguistic	5359.38	48.10	+	0.20	0.09

* $p < .05$, ** $p < .01$

Chapter 3: A Research Manuscript

A Corpus Linguistic Analysis of Public Reddit Blog Posts on Non-Suicidal Self-
Injury

Mandy M. Greaves and Cass Dykeman

Oregon State University

Author Note

Mandy Greaves, Counseling Academic Unit, Oregon State University; Cass Dykeman, Counseling Academic Unit, Oregon State University.

The research contained in this manuscript was conducted under the approval of the Oregon State University Institutional Review Board (Study ID No. 8617) and was part of the first author's dissertation research project.

Correspondence concerning this article should be addressed to Cass Dykeman, Counseling Academic Unit, Oregon State University, 104 Furman Hall, Corvallis, OR 97331-3502. E-mail: dykemanc@onid.oregonstate.edu

Abstract

While non-suicidal self-injury (NSSI) is not a new phenomenon, there is still a limited understanding of the behavior, the intent behind the behavior and what the individuals themselves say about their behavior. This study collected pro-NSSI public blog posts from Reddit and analyzed the content linguistically using LIWC software, in order to examine the use of NSSI specific words, linguistic properties and the psychological linguistic properties. The results inform current counseling practices by dispelling myths and providing insight into the inner world of people who engage in NSSI. The most frequently appearing category of NSSI specific words in the Reddit blogs was the reasons for engaging in NSSI. The linguistic properties found in the analysis reflected the predicted results; authors of pro-NSSI posts used first-person singular pronouns extensively, indicating high levels of mental health distress and isolation. The psychological linguistic properties that could be observed in these public Reddit posts were dominantly in a negative emotional tone which demonstrates youth and impulsivity. The linguistic properties found when these posts were analyzed supports the work of earlier studies that dispelled common myths about NSSI that were circulating in the mental health community.

Keywords: NSSI, LIWC, linguistic analysis, mental health, self-injury

Linguistic Analysis of Words Used in Public Reddit Blog Posts on Non-Suicidal Self-Injury

Non-suicidal self-injury (NSSI) is not a new phenomenon. In his *Lectures on Ethics and Metaphysics of Morals*, Immanuel Kant (1804) concludes that “We must reverence humanity in our own person” and argues that those “who harm themselves harm all of humanity” (as cited in Portmann, 2004, p. 19). One of the first documented cases of self-inflicted harm was in ancient Greece, when a Greek historian Herodotus illustrated a Spartan leader who asked for a knife and began to harm himself starting with his shins (Spiegel, 2005). During the 19th and early 20th century, documented cases of self-harm included extreme behaviors such as eye enucleation and genital mutilation. By the end of the 1930’s, NSSI behavior was associated with suicide due to the extreme nature of the acts. It was not until the 1960’s that mental health professionals began to focus on wrist-cutting and classified this behavior as a specific disorder that could be distinguished from both cutting behavior and suicidal behavior. In their study of wrist-cutting syndrome Rosenthal, Rinzler, Walsh and Klausner (1972), patients described feeling a sense of numbness, dissociation and emptiness leading up to the cutting behavior. Favazza (1998) argued that self-harm is distinguished from suicidal acts by the individual’s intentions; Favazza argues that a person who is suicidal intends to end all feelings, but a person who self-harms is attempting to cope with negative feelings. While NSSI as a separate, distinctive phenomenon is now well-documented, little else is known about this behavior in adolescents or emerging adults.

Presently, scholars believe that adolescents self-harm because they are trying to avoid feeling intense emotions (Allen, 2011). American adolescents are turning to non-lethal self-harming behaviors at an increasing rate; approximately 8% adolescents between of 14- and 19- years of age experience self-harm nationwide (Allen, 2011). Moran, Coffey, Romaniuk, Olsson, Borschmann, Carlin and Patton (2011) noticed a trend that middle-to-late adolescents are the most vulnerable to experiencing self-harm because of their stage of brain development, which leads to a lack of emotional control and promotes high risk-taking behaviors. At the conclusion of their study, Moran et al., (2011) suggest that self-harming behavior is linked to anxiety and depression, which may contribute to continued self-harming behavior later into adulthood. The authors, therefore, recommend specifically treating for the self-harming behaviors.

Social media use amongst adolescents and emerging adults is at an all-time high (Smith & Anderson, 2018). The Pew Research Center survey found that the majority of Americans use Facebook and YouTube as social media outlets. The study found that younger Americans (18–24 years old) use other platforms like Instagram, Snapchat, and Twitter more frequently than Facebook. In their most recent (2018) study, The Pew Research Center found “roughly three-quarters of Facebook users and around six-in-ten Snapchat and Instagram users visit each site daily” (Smith & Anderson, 2018). Given the fact that the majority of Americans live in a reality informed by social media, it is particularly important to address the lack of research into how people who engage in NSSI use social media and use that information to develop new treatment plans.

Previous research on social media use and NSSI behavior seems to have been limited to exploring the benefits and the risks of the usage of social media for NSSI. Much of the prior research has been very similar, in that it has focused on investigating the benefits and risks of posting about NSSI behaviors and pro-NSSI Instagram pages. Some studies have taken a balanced approach; on the one hand, they list the benefits of reducing social isolation and finding social support, while on other hand they also find that posting NSSI activity can invite reinforcement of NSSI or triggering NSSI urges (Lewis & Seko, 2015).

A review of the literature on both non-suicidal self-injury and social media use revealed five major points to consider. These points are: (1) definition of non-lethal self-injury, (2) NSSI behavior is not an uncommon phenomenon in adolescents, (3) prevalence of pro-NSSI posts on social media, (4) pervasiveness of social media impacts people's behavior and choices, (5) mental health and the impact of the internet and (6) exploring use of LIWC as new method of understanding language online. Finally, the research questions will be detailed.

There are varying definitions of self-harm in the literature. One of the most widely cited resources for self-harm/self-mutilation is by Favazza (1998), which according to Google Scholar (2015) has been cited 688 times in the professional literature. Favazza defines self-mutilation as “the deliberate, non-suicidal destruction of one's own body tissue” (p. 260). He concludes his overview by writing, “Superficial/moderate SM includes compulsive acts such as trichotillomania and skin picking and such episodic acts as skin-cutting and burning” (p. 259). Skegg's definition, which has built on Favazza's definition, has also been cited widely (448

times; Google Scholar, 2015). Skegg's definition adds depth to Favazza's definition of self-mutilation/self-harm; in particular, Skegg (2005) adds overdosing to the definition of self-harm and further states that self-harm takes place on a continuum of "deliberate" and "self-wounding" behaviors (p. 1471). Gilbert (2009) states that people who experience self-harm are often experiencing criticism and shame which contributes to "self-attacking" (p. 199).

NSSI behavior is not an uncommon phenomenon in adolescents. It is estimated that two million Americans use some method of self-harm (Spiegel, 2005). By the age of 14 or 15, approximately one in 12 adolescents—between six and ten percent of all teenagers—engage in NSSI (Allen, 2011). According to several studies, adolescents who self-injure experience difficulty with affect regulation, have a desire to punish themselves and want to communicate with others (Briere & Gil, 1998; Chapman & Dixon-Gordon, 2007; Klonsky & Muehlenkamp, 2007 as cited in Van Vliet & Kalnins, 2011). Building on previous research, Van Vliet and Kalnins (2011) argue that there is a need in clinical treatment for increased "emotion regulation, self-acceptance, and positive ways of relating with others" (p. 296).

As social media rises in popularity and accessibility, there has been a corresponding increase in social media posts about NSSI. Adolescents and emerging adults have been posting pictures of their NSSI on social media and posting text that describes NSSI behavior, including tips on what to use to self-injure and how to hide their injuries. There have been studies conducted that explore the possible consequences users face for posting their personal NSSI on social media. In a meta-analysis exploring the risks and benefits of online posting of NSSI, Lewis and Seko

(2015) found major themes in NSSI social media involvement. Their study found that one possible reason that people who utilize NSSI post on social media is to connect with others who are going through similar experiences. Many people who engage in NSSI experience high levels of shame, fear and embarrassment. One argument from the Lewis and Seko study is that people who utilize NSSI behaviors can post anonymously without judgment and shame. The researchers also found that there was a “reduction of social isolation through online interactions” (Lewis & Seko, 2015, p. 254). Studies that explore online posting of NSSI, like Lewis and Seko’s, encourage therapists to use this information in therapy. Specifically, they encourage therapists to explore and discuss social media use and what benefits the clients feel they receive from posting their NSSI behaviors (Marchant et al., 2017). While prior studies have attempted to understand the reasons and consequences for discussing NSSI on social media, there has been little or no engagement about the language that they use to describe their NSSI and what that language can tell us

The pervasive nature of social media provides a normative influence, affecting people’s social behaviors, consumer habits, personal choices and even mood. In 1955, Deutsch and Gerard described the phenomenon known as “normative influence” as an intrinsic characteristic of human behavior, in which social influences lead to conformity. While much has changed in American society since 1955, the concept of conformity is very much alive in the technological age. Huh, Vosgerau and Morewedge (2014) have further developed the concept of normative influence by describing a phenomenon known as the “social default”, which includes how other people’s known decisions will impact a person’s personal choices. These concepts

are grounded in people's natural tendency to mimic other people's behaviors and describe how observing other's choices can create social default effects. Businesses in the age of widespread social media usage, seeking to harness this tendency to mimic others for their own bottom lines, have studied these human behaviors to understand the effects of social media on consumer choices; they have used the results of their studies to inform and advance their marketing strategies. Mental health, however, has lagged behind business, and appears to continue to lack in understanding just how social media influences mental health behaviors. It stands to reason that if a person is influenced by observing other's choices in person, then the same would occur online.

Online behaviors seem to both fascinate and astound social media consumers. We can see how observing other's choices being documented on social media affects individual's decision-making processes. While many are aware of social media challenges, little is known about why or how these challenges become so popular. Social media challenges range from harmless to helpful, like The Mannequin Challenge and the ALS Ice Bucket challenge (which raised money for ALS research), to dangerous and deadly, like the Salt and Ice Challenge, the Fire Challenge (in which the individual lights themselves on fire and jumps into a pool or shower), and the recent Tide Pod Challenge (Calmenson, 2018). In opinion pieces and elsewhere, people have speculated that peer pressure and pressure to increase one's social media popularity and gain followers have driven the widespread, dangerous popularity of these challenges (Setzekorn, 2018). While there is no definitive answer on why these deadly behaviors become glamorized on social media, it is clear that they reveal that

there is an online pressure to conform and that people can be influenced even by online evidence of other's decisions when making their own personal choices.

Social media can also allow people to become involved in online communities, like Reddit. Reddit is the fifth most visited site in the United States and is home to over 138 thousand active communities ("The Conversation Starts Here,"). Unlike blogging platforms like Tumblr, the website provides its users with anonymity, which may make them feel more protected; the only requirement for posting on Reddit is a screen name (Harriagian, 2018). Each community in Reddit, which is called a sub-reddit, allows people to post on a topic, and others can respond to these posts. Reddit's anonymity may encourage people to post more honestly about their experiences and feelings. Reddit's anonymity and its users' increased comfort level makes Reddit a good fit for a linguistic analysis study (Kamarudin, Rakesh, Beigi, Manikonda & Liu, 2018). For the purpose of this study, the sub-reddits explored and used in the corpus involved pro-NSSI behavior.

As social media usage rises, mental health researchers have become more interested in how people's language use can give indications of their mental health. In a study on social media posting by people struggling with post-partum depression, Choudhury, Counts, and Horvitz (2013) noted that the language that these participants used provided helpful "psychological markers" that gives information on the individual's inner workings. Similarly, there has been an increased effort among mental health experts to understand how the language of self-harm has developed and is expressed on social media.

This study uses the corpus linguistics method to evaluate the use of NSSI related language on Reddit. This approach uses a corpus constructed from large amounts of data from any text source, like social media posts. The data is then inputted into software for analysis; the analytical software can complete a more in-depth analysis than what people are capable of doing (McGlashan, 2018). The Linguistic Inquiry and Word Count (LIWC; Pennebaker et al., 2007) is commonly used in current literature because it is capable of analyzing a variety of categories and emotional words; it is also capable of analyzing and calculating aspects of language. LIWC is often used to explore mental health markers in which researchers specifically focus on emotion word use, social word use, self-referencing, drives and pronoun use (Adrian et al., 2011). Analyzing language has the potential to inform research into an individual's experience, environment and mental health.

Given the aforementioned, three research questions were designed to guide this study. The first research question was: What is the NSSI specific content word use of the individuals making posts? The NSSI specific content word use categories were: Methods of NSSI, Cutting-Specific Terms, NSSI Terms, Instruments Used, Reasons for NSSI, and Hidden Hashtag Terms. The second research question was: What is the pattern of use of linguistics processes of the individuals making posts compared to the pattern that appears in blogs overall? The linguistics process categories were: 1st person singular, 1st person plural, 3rd person singular, and 3rd person plural. The third research question was: What is the pattern of use of psychological processes of the individuals making posts compared to the pattern that

appears in blogs overall? The psychological process categories were: emotional tone, negative emotions, anxiety, anger, sadness and risk focus.

Method

Design

This study used synchronic corpus linguistic design to explore public Reddit posts on NSSI (Weisser, 2017). The Linguistic Inquiry and Word Count (LIWC) was used to analyze the Reddit posts containing pro-NSSI content. The variables were the following linguistic categories: 1st person singular pronouns, 3rd person singular pronouns, 1st person plural pronouns, 3rd person plural pronouns, emotional tone, negative emotion, anxiety, anger, sadness and risk. Secondly, the following categories are specific to NSSI: methods of NSSI, cutting-specific terms, NSSI terms, instruments used, reasons for NSSI, and hidden hashtag terms. The sample was compiled from Reddit during a year's worth of data ranging from January 1, 2017 to December 31, 2017. The sample was compared to individuals who make Reddit posts from national norms for social media posts.

Corpus

The corpus was constructed using Reddit's API system to collect public posts that published between January 1, 2017 and December 31, 2017. The API was instructed to scrape posts that: (a) were written in monolingual English and (b) included words or linguistic features from the list of NSSI terms compiled by the researchers (list available from first author). The API was instructed to eliminate user names, URLs, hashtags, location of the posts, posts from outside of the United States, photographs and foreign languages while collecting texts for the corpus.

Over one thousand Reddit posts were collected. These posts were then cleaned. This cleaning included the following:

1. All abbreviations were spelled out (e.g. “max” to “maximum”)
2. UK English word usage and spelling were converted to US English word usage.
3. Lexical Variants (LV)/Textese/Out-of-vocabulary (OOV) words were transformed by lexical normalization.
 - a. LV/Textese/OOV Category “Letter” was normalized. Letter refers to instances where letters are missing or there are extraneous letters, but the lexical correspondence to the target word form is trivially accessible (e.g. shuld “should”) (Han & Baldwin, 2011).
 - b. LV/Textese/OOV Category “Number Substitution” was normalized. Number Substitution refers to instances of letter-number substitution, where numbers have been substituted for phonetically-similar sequences of letters (e.g. 4 “for”) (Han & Baldwin, 2011).
 - c. LV/Textese/OOV Category “Letter&Number” was normalized. Letter&Number refers to instances which have both extra/missing letters and number substitution (e.g. b4 “before”) (Han & Baldwin, 2011).
 - d. LV/Textese/OOV Category “SpaceDeleted” was normalized. SpaceDeleted refers to where spaces having being deleted between words (e.g. sucha “such a”) (Han & Baldwin, 2011).

- e. LV/Textese/OOV Category “NSSI Slang” was normalized. NSSI Slang refers to words used to hide NSSI related Tumblr posts (Moreno, Ton, Selkie, & Evans, 2016).

After cleaning, the average word count per post was 187. The total word count for the entire corpus was 228,172 words, which provided sufficient material for analysis.

Measures

GNLS- Methods of NSSI. Methods of NSSI are ways in which people engage in non-suicidal self-injury. This is a broad category that includes all forms of self-harming behaviors, including biting, burning, erasing, hitting and picking.

GNLS- Cutting-Specific Terms. Cutting is a specific category under methods of NSSI. Since cutting is one of the most common methods of non-suicidal self-injury, it has its own sub-category. Examples of words for this category is cut, the seed word cut*, and cutting.

GNLS- NSSI Terms. NSSI terms are any terms that describe NSSI. Some examples include non-suicidal, NSSI, suicide and the seed word self-mutil*.

GNLS- Instruments Used. The category includes the tools that people use to engage in NSSI behaviors. Some of these instruments include blade, bleach, eraser, fingernail and laxatives.

GNLS- Reasons for NSSI. There are many reasons why people choose to engage in NSSI behaviors. Some examples of reasons for NSSI are addiction, anger, anxiety, attention, control and emptiness or numbness.

GNLS- Hidden Hashtag Terms. Due to new social media restrictions, people who engage in NSSI behaviors and post about it on social media have become

creative in the hashtags they use to communicate with the community of people who engage in NSSI. Some hidden hashtags that are hidden include selfharmmmmm, mysecretfamily, selfinjuryyyy, secretsociety_123, ehtilb and mauled.

LIWC- 1st Person Singular Pronoun. According to Pennebaker (2013), people who use 1st person singular pronouns track their focus of attention. People who are depressed tend to “pay more attention to themselves” (Pennebaker, p. 40). Pennebaker has discovered in his research that the more depressed a person is the more they use I-words (Pennebaker, p. 108). Examples of 1st person singular pronouns are “I”, “me” and “my”. Pennebaker, Boyd, Jordan, and Blackburn (2015) report adequate reliability and validity for this LIWC scale.

LIWC- 3rd Person Singular Pronoun. Third person singular pronouns allow the writer to talk about other people. Common examples are “he”, “she”, “it”, “him”, or “her”. Pennebaker, Boyd, Jordan, and Blackburn (2015) report adequate reliability and validity for this LIWC scale.

LIWC- 1st Person Plural Pronoun. According to Pennebaker (2013), first person plural pronouns have multiple meanings. He uses the example of “we”. We can give a sense of warmth and community meaning me and you. But then there is a more distant meaning of we, which creates distance. Pennebaker uses the example of “we really need to take out the trash” (p. 41). This type of “we” is a way to nicely order someone to do something. Common examples are “we”, “us” and “our”. Pennebaker, Boyd, Jordan, and Blackburn (2015) report adequate reliability and validity for this LIWC scale.

LIWC- 3rd Person Plural Pronoun. Examples of third personal plural pronouns are “they”, “them” and “theirs”. Pennebaker, Boyd, Jordan, and Blackburn (2015) report adequate reliability and validity for this LIWC scale.

LIWC- Emotional Tone. Emotional tone is a summary variable. Emotional tone is “the degree to which people express emotion, how they express emotion, and the valence of that emotion can tell us how people are experiencing the world” (Tausczik and Pennebaker, 2010, p. 32).

LIWC- Negative Emotion. Negative emotion words are typically used to describe a negative event. Examples of negative emotion words are hurt, ugly, nasty. Pennebaker, Boyd, Jordan, and Blackburn (2015) report adequate reliability and validity for this LIWC scale.

LIWC- Anxiety. Anxiety is a specific category under negative emotion words. An example of anxiety related words are worried and fearful. Pennebaker, Boyd, Jordan, and Blackburn (2015) report adequate reliability and validity for this LIWC scale.

LIWC- Anger. Anger is a specific category under negative emotion words. Example of anger related words are hate, kill and annoyed. Pennebaker, Boyd et al., (2015) report adequate reliability and validity for this LIWC scale.

LIWC- Sadness. Sadness is a specific category under negative emotion words. Example of sadness is crying, grief and sad. Pennebaker, Boyd et al., (2015) report adequate reliability and validity for this LIWC scale.

LIWC- Risk. Risk-taking behavior is common amongst youth who engage in NSSI. Risk related words include danger and doubt. Pennebaker, Boyd et al., (2015) report adequate reliability and validity for this LIWC scale.

Data Analysis

For the first research question, the percentage of all words for each NSSI-specific category will be reported. In addition, the raw count for each word in each category will be calculated. In terms of the second and third research questions, the following will be reported for all categories except for emotional tone: (a) percentage of all words, (b) log likelihood statistic of the comparison of the NSSI corpus to blog general norms, and (3) log ratio of the comparison of the NSSI corpus to blog general norms. The log likelihood analysis requires raw scores, which were obtained by multiplying the percentage of all words with the size of the corpus. The log likelihood formula used was from Cressie and Read (1989) as cited in Rayson and Garside (2010). The log ratio is a measure of effect size (Hardie, 2014). As Hardie notes, a log ratio of 1 means a word is twice as common in a corpus than it is in the comparison corpus. The score for emotional tone is based upon a standard score rather than a raw score. As such, a one sample z test of proportion was used to analyze this category rather than the log likelihood test. All analyses were conducted using Excel. Given the small number of analyses conducted, the standard .05 alpha level was retained.

Results

In reference to the first research question, the most frequently used NSSI-specific category was Reasons for NSSI. The percentage of all words were as follows: Methods of NSSI (0.92%), Cutting-Specific Terms (0.58%), NSSI Terms (0.36%),

Instruments Used (0.11%), Reasons for NSSI (2%), and Hidden Hashtag Terms (0.32%). A list of the raw counts for the words or word stems in all categories can be found in Table 1. In terms of the second research question, the linguistic process category that differed most from the general blog norms was first person singular ($LL = 66.37$). The effect size for this category was a log ratio of 0.76. Except for the emotional tone, a list of the results of all the categories can be found in Table 2. In terms of emotional tone, the one sample z test of proportion was significant ($z = -9.14, p < .05$). In terms of the third research question, the psychological process category that differed most from the general blog norms was negative emotion ($LL = 41.04$). The effect size for this category was a log ratio of 1.01. A list of the results of all the categories can also be found in Table 2.

Discussion

This study examined language used in pro-NSSI micro-blogs on Reddit and investigated whether there were differences between the use of language in pro-NSSI posts when compared to blogs overall. This study addressed the following research questions. What is the NSSI specific content word use of the individuals making posts? What is the pattern of use of linguistics processes of the individuals making posts? What is the pattern of use of psychological processes of the individuals making posts? Next, results are presented and discussed. Toward the end of the section, limitations and implications of the finds are presented.

The first research question aims to understand the specific NSSI words that were used in these posts. The findings indicated that the words most frequently used in pro-NSSI posts were reasons for their NSSI. In this category, the most frequently

used words in rank order are feel, help and scar. In the data drawn from posts, it seems as though the users share a strong desire to get help for their NSSI behaviors, but there is also a high focus on their scars. The raw data from Reddit reveals that individuals making pro-NSSI posts use language that is intensely focused on their scars, though some users talked about their scars in a prideful manner and others talked about their scars in a shame-filled, negative manner. One possible reason for this is the anonymity of Reddit; it is likely much easier for a person who uses NSSI to talk about their scars anonymously online rather than talking about them in person. Frequently, it is very difficult for people engaged in NSSI to talk about their struggles with their close friends and family; they often experience a lot of behaviors from others that add to the internal shame they already feel (Kane, 2017). Many of the blogs discuss how outside people react negatively to NSSI behaviors, which triggers an intense emotional reaction. It appears that Reddit provides a safe place to share about one's experience without the fear of judgement and shame.

Another possible reason for the high frequency of the words feel, help and scar may be the individual's intentions regarding self-injury. One of the long-standing myths about NSSI, which is still held by many counselors, is that NSSI is related to suicidal intent. While the word suicide is mentioned 86 times in the raw data, an analysis of the context reveals that suicide is mostly used when individuals are expressing that they do not want to attempt suicide, they want to end the pain they are experiencing now. The data suggests that many people who are engaged in NSSI feel misunderstood about their behavior and insist that they are not suicidal. This study supports the findings in previous research that adolescents who engage in NSSI

communicate more about their behavior with friends and mental health professionals than adolescents who engage in suicidal self-injury (Baetens, Claes, Muehlenkamp, Grietens & Onghena, 2011). These findings suggest that while the individuals are not seeking to end their lives, they are seeking help. While both the anonymity theory and the intention theory attempt to explain the results of specific NSSI word use, it is more likely that the language is can be attributed to the intention theory than that anonymity theory.

The second research question concerns the linguistic patterns in pro-NSSI blogs. The data demonstrates that the most frequent and statistically significant pattern was the use of the first-person singular—words such as I, me and my. According to Pennebaker (2013), people’s pronoun use tracks their focus of attention. Therefore, people who use first person singular pronouns are typically looking at themselves quite often and most likely experience anxiety, self-consciousness, pain and depression. People who engage in NSSI usually suffer from a form of mental illness like anxiety or depression and use NSSI to cope. It stands to reason that people who use NSSI to cope also struggle with mental health issues, and this is reflected in their pronoun use.

Another reason that Reddit users discussing NSSI use the first-person singular is the social isolation of NSSI. Many individuals who engage in NSSI self-harm in private and often feel a lot of shame about their behavior. The first-person singular pronouns demonstrate a lack of community or sense of belonging; instead they express a feeling of aloneness. Pennebaker (2013) also stated that “I” words denote vulnerability. Reddit is a unique platform that allows its users to post on their website

with a simple, fictional user name. It is one of only a few anonymous message board forums. This gives their users a sense of security because it would be difficult to identify who they are simply using a screen name. The two reasons support each other in the results, but it is likely that the most frequent use of first-person singular pronouns is due to the isolation these individuals experience.

The third research question pertains to the psychological word use patterns in pro-NSSI blogs. The most common and statistically significant psychological pattern was negative emotion language. Pennebaker (2013) suggests that through his findings in research younger people tend to use more negative emotion words when they write, unlike older people who write using more positive emotion words. He discovered that this difference in word usage between older and younger writers became apparent around the age of 40, and then became much more pronounced in the oldest age groups. Pennebaker theorized that older writers can better regulate their emotions and can look at the world from different perspectives. Given this research, it stands to reason that the people who post on pro-NSSI Reddit blogs are younger, more impulsive and view the world through a darker lens.

Another possible reason why those engaged in NSSI use a lot of negative emotion words is high levels of neuroticism. Pennebaker, Mehl and Niederhoffer (2003), discovered that neuroticism was positively correlated with the high use of negative emotion words. This finding was also consistent with the high use of first-person pronouns and the amount of text in which the individual is self-focused. Neuroticism is one of the big five personality traits (Costa & McCrae, 1985). Using Eysenck's influential Psychoticism, Extraversion and Neuroticism theory of

personality (Eysenck, & Eysenck, 1985), Shatz (2004) noted that neuroticism involves differences in emotional reactivity to negative environmental stimuli with high scorers exhibiting anxiety, depression, guilt, low self-esteem, and moodiness. The research from Pennebaker et al. supports this study's findings that demonstrate that people who are engaged in NSSI use more negative emotion words and first-person pronouns. While it is likely that both age and neuroticism cause this high usage rate of negative emotion words are very likely, but neuroticism is the more likely cause.

The study's results produce an image of the typical NSSI engaged person as a young individual who is affected by a form of mental illness, experiencing feelings of isolation and attempting to cope with their intense feelings. This study, like others, found that the posts written by individuals who engage in NSSI contain psychological markers, namely the use of negative emotion language, that demonstrate a level of mental illness. This negative emotion language also demonstrates that the user is experiencing a sense of longing for a safe place to talk about topics that may not be socially acceptable, a sense of aloneness and isolation and an indication of age and perspective. Therefore, this research implies that using what the client's strengths are which is written expression, and reframing in more positive terms can benefit the client's mental health and adaptive coping behaviors.

Limitations

There are several limitations to this study. Firstly, words often have multiple meanings and could be put into a category that does not reflect the true meaning of the word in its context. The second limitation the amount of data that was included in

the corpus could have been limited by the fact that only English language words were included. The third limitation is that this study may not be generalizable to the entire population of people who post on pro-NSSI social media and blog sites. Inferences into the data should be done with care.

Implications

Despite the study's limitations, the results have potentially significant implications for the greater understanding of NSSI and developing improved counseling for individuals who engage in NSSI. There are several persistent myths about NSSI in the mental health community. One of those myths is that people who use NSSI to cope are also suicidal. This study reveals that while NSSI engaged persons may experience suicidal thoughts, they are not trying to end their lives. These results support previous research that has also distinguished between NSSI and suicidal attempts, though this myth endures in the field of mental health. When posts on Reddit are evaluated linguistically, it seems clear that people who engage in NSSI are directly communicating with a broader community and saying that they are not suicidal and that they are not trying to end their life. Counselors may be better able to understand the true intentions of people who are engaged in NSSI, and thus improve their treatment plans, by examining the actual words used by these individuals.

Analyzing and understanding the language used by individuals to describe their mental health and NSSI behaviors can help counselors in the mental health field better understand their patients and provide insight into their inner world. This study's results from this study demonstrate how we as counselors can better communicate with these individuals and treat them. While research and lived

experience have already demonstrated that people who engage in NSSI are usually younger, we now can identify their language markers to be more negative in tone and see that they use more first-person singular pronouns which gives counselors insight into how alone they feel. This has the potential to help counselors to explore better and more efficiently the level of isolation, aloneness and level of mental illness, improving treatment.

This study found that people who engage in NSSI and post about it on blogs are seeking help by reaching out to a community of similar people. This study implies that these people are already using writing as one method for coping with their emotional pain. Pennebaker (2013) found in his study that people who used positive emotion words while writing about their personal trauma experienced improved physical and mental health during the weeks of the experiment. In this same study, researchers discovered that people who used negative emotion word frequently did not benefit from writing. This study demonstrates that while these adolescents are already writing to cope, their writing is not beneficial because of the high negative emotion use.

Counselors can use the information from this study to inform their practice by making use of theories that may better reach adolescents who use NSSI. As counselors we are told that we need to meet our clients where they are at. There is a high number of adolescents and young adults who post their NSSI behaviors on blogs. If we are to meet them where they are at, it is on the internet writing about their pain. However, their writing is filled with negative emotion words, which Pennebaker (2013) suggests may prevent this coping mechanism from being useful. One possible

way that counselors can use the insights from this study is by using a theory such as Narrative therapy to utilize what clients are already doing and altering their writing to be more beneficial to them. Hoffman and Kress (2008) have suggested that Narrative therapy is a beneficial treatment for those who use NSSI because this approach gives clients the power to name the problem and helps them learn to separate the problem from themselves, and this type of intervention can effectively be used with this population.

References

- Adrian, M., Zeman, J., Erdley, C., Lisa, L., & Sim, L. (2010). Emotional dysregulation and interpersonal difficulties as risk factors for non-suicidal self-injury in adolescent girls. *Journal of Abnormal Child Psychology*, *39*, 389-400. <https://dx.doi.org/10.1007/s10802-010-9465-3>
- Allen, J. E. (2011, November 17). 1 in 12 Teens cut, deliberately hurt themselves. *ABC News*, ABC News Medical Unit. Retrieved from <https://abcnews.go.com/Health/MindMoodNews/12-teens-cut-harm-themselves/story?id=14969232>
- Baetens, I., Claes, L., Muehlenkamp, J., Grietens, H., & Onghena, P. (2011). Non-suicidal and suicidal self-injurious behavior among Flemish adolescents: A Web-Survey. *Archives of Suicide Research*, *15*, 56-67. <https://dx.doi.org/10.1080/13811118.2011.540467>
- Calmenson, B., & Sullivan, P. (2018, February 06). *Why social media challenges are spiraling out of control*. Retrieved March 31, 2018, from <https://www.washingtonexaminer.com/why-social-media-challenges-are-spiraling-out-of-control>
- Choudhury, M. D., Counts, S., & Horvitz, E. (2013). Major life changes and behavioral markers in social media. *Proceedings of the 2013 Conference on Computer Supported Cooperative Work - CSCW 13*. <https://dx.doi.org/10.1145/2441776.2441937>

- Cressie, N., & Read, T. R. (1989). Pearson's χ^2 and the loglikelihood ratio statistic G2: a comparative review. *International Statistical Review/Revue Internationale de Statistique*, 57, 19-43. <http://dx.doi.org/10.2307/1403582>
- Deutsch, M., & Gerard, H. B. (1955). A study of normative and informational social influences upon individual judgment. *The Journal of Abnormal and Social Psychology*, 51, 629-636.
<http://dx.doi.org/10.1037/h0046408>
- Eysenck, H. J., & Eysenck, M. W. (1985). *Personality and individual differences: a natural science approach*. NY: Plenum Press.
- Favazza, A. R. (1998). The coming of age of self-mutilation. *The Journal of Nervous & Mental Disease*, 186, 259-268. <https://dx.doi.org/10.1097/00005053-199805000-00001>
- Han, B., & Baldwin, T. (2011). Lexical normalisation of short text messages: Mkn sens a #twitter. In *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies-Volume 1* (pp. 368-378). Stroudsburg PA: Association for Computational Linguistics.
Retrieved from: <https://dl.acm.org/citation.cfm?id=2002520>
- Hardie, A. (2014). *Log Ratio – an informal introduction*. Lancaster, UK: Centre for Corpus Approaches to Social Science. Retrieved from <http://cass.lancs.ac.uk/log-ratio-an-informal-introduction/>
- Harriagian, K. (2018, October 07). *Geocoding without geotags: A text-based approach for reddit*. Retrieved from <https://export.arxiv.org/abs/1810.03067>

- Hoffman, R. M., & Kress, V. E. (2008). Narrative therapy and non-suicidal-self-injurious behavior: Externalizing the problem and internalizing personal agency. *The Journal of Humanistic Counseling, Education and Development*, 47, 157-171. <https://dx.doi.org/10.1002/j.2161-1939.2008.tb00055.x>
- Huh, Y. E., Vosgerau, J., & Morewedge, C. K. (2014). Social defaults: Observed choices become choice defaults. *Journal of Consumer Research*, 41, 746-760. <https://dx.doi.org/10.1086/677315>
- Kamarudin, N. S., Rakesh, V., Beigi, G., Manikouda, L., & Liu, H. (2018). A study of Reddit-users response to rape. *2018 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM)*. <https://dx.doi.org/10.1109/asonam.2018.8508855>
- Kane, A. (2017, October 07). *What you can say if you see my self-harm scars*. Retrieved from <https://www.intrusivethoughts.org/blog/can-say-see-self-harm-scars/>
- Lewis, S. P., & Seko, Y. (2015). A double-edged sword: A review of benefits and risks of online non-suicidal self-injury activities. *Journal of Clinical Psychology*, 72, 249-262. <https://dx.doi.org/10.1002/jclp.22242>
- Li, F., & Du, T. (2014, April 03). *Listen to me - Evaluating the influence of micro-blogs*. Retrieved from <https://www.sciencedirect.com/science/article/pii/S0167923614001183>
- Marchant, A., Hawton, K., Stewart, A., Montgomery, P., Singaravelu, V., Lloyd, K., . . . John, A. (2018). A systematic review of the relationship between internet

- use, self-harm and suicidal behaviour in young people: The good, the bad and the unknown. *Plos One*, *13*. <https://dx.doi.org/10.1371/journal.pone.0193937>
- McGlashan, M. (Ed.). (2018). Corpus: Some key terms. *CASS: Briefings*. Retrieved from <http://cass.lancs.ac.uk/wp-content/uploads/2013/12/CASS-Gloss-final1.pdf>
- Moreno, M. A., Ton, A., Selkie, E., & Evans, Y. (2016). Secret Society 123: understanding the language of self-harm on Instagram. *Journal of Adolescent Health*, *58*, 78-84. <https://doi.org/10.1016/j.jadohealth.2015.09.015>
- Pennebaker, J. W. (2013). *The secret life of pronouns: What our words say about us*. NY: Bloomsbury Press.
- Pennebaker, J.W., Boyd, R.L., Jordan, K., & Blackburn, K. (2015). The development and psychometric properties of LIWC2015. Austin, TX: University of Texas at Austin. <https://dx.doi.org/10.15781/T29G6Z>
- Pennebaker, J.W., Chung, C.K., Ireland, M., Gonzales, A., & Booth, R.J., (2007). *The development and psychometric properties of LIWC2007*. Austin, TX: University of Texas at Austin.
- Pennebaker, J. W., Mehl, M. R., & Niederhoffer, K. G. (2003). Psychological Aspects of Natural Language Use: Our Words, Our Selves. *Annual Review of Psychology*, *54*, 547-577. <https://dx.doi.org/10.1146/annurev.psych.54.101601.145041>
- Portmann, J. (2004). *Bad for us: The lure of self-harm*. Beacon Press.
- Rayson, P., & Garside, R. (2000). Comparing corpora using frequency profiling. In *Proceedings of the workshop on comparing corpora-Volume 9* (pp. 1-6).

Stroudsburg, PA: Association for Computational Linguistics.

<http://dx.doi.org/10.3115/1117729.1117730>

- Rosenthal, R. J., Rinzler, C., Wallsh, R., & Klausner, E. (1972). Wrist-Cutting Syndrome: The Meaning of a Gesture. *American Journal of Psychiatry*, *128*(11), 1363-1368. <https://doi.org/10.1176/ajp.128.11.1363>
- Shatz, S. M. (2004). The relationship between Horney's three neurotic types and Eysenck's PEN model of personality. *Personality and Individual Differences*, *37*, 1255-1261. <https://doi.org/10.1016/j.paid.2003.12.009>
- Setzekorn, M. (2018, January 25). *Peer pressure factor in social media Tide Pod Challenge*. Retrieved March 31, 2018, from http://www.dailytoreador.com/news/peer-pressure-factor-in-social-media-tide-pod-challenge/article_8bbcd9ba-0188-11e8-9ec6-bb3b7faa067f.html
- Smith, A., & Anderson, M. (2018, March 01). *Social media use in 2018*. Retrieved from <http://www.pewinternet.org/2018/03/01/social-media-use-in-2018/>
- Spiegel, A. (2005, June 10). *The history and mentality of self-mutilation*. Retrieved from <http://www.npr.org/templates/story/story.php?storyId=4697319>
- Tausczik, Y. R., & Pennebaker, J. W. (2009). The psychological meaning of words: LIWC and computerized text analysis methods. *Journal of Language and Social Psychology*, *29*, 24-54. <https://dx.doi.org/10.1177/0261927x09351676>
- The Conversation Starts Here. (n.d.). Retrieved from <https://www.redditinc.com/>
- Vliet, K. V., & Kalnins, G. (2011). A Compassion-Focused approach to nonsuicidal self-injury. *Journal of Mental Health Counseling*, *33*, 295-311. <https://dx.doi.org/10.17744/mehc.33.4.j7540338q223t417>

Weisser, M. (2016). *Practical corpus linguistics: An introduction to corpus-based language analysis*. NY: John Wiley & Sons.

Table 1

Item Word Counts

Word/Stem	Methods	Cutting Terms	NSSI Terms	Instruments Used	Reasons for NSSI	Hidden Hashtag Terms	Word Ct.
Cut		X					1575
Cutt*		X					543
Cutting		X					520
Anxiety					X		109
Control					X		109
Blade				X			97
Burning	X		X				93
Anger					X		80
Picking	X		X				80
Attention					X		79
Hitting	X		X				302
Addiction					X		22
Bite	X						17
Fingernail				X			13
Biting			X				12
Empty					X		12
Eraser				X			2
Mauled						X	2
Bleach				X			1
Erasing	X		X				1
ehilb						X	0
Laxatives				X			0
mysecretfamily						X	0
secretsociety_123						X	0
selfharmmmm						X	0
selfinjuryyy						X	0

Table 2

Results for Linguistic and Psychological Processes

Category	Process Type	Reddit Word Count	Blog Norm Word Count	Over/Under Use	Log Likelihood	Log Ratio
1st Person Singular	Linguistic	24575.94	200.72	+	66.37**	0.76
Negative Emotion	Psychological	9612.65	66.05	+	41.04**	1.01
1st Person Plural	Linguistic	787.54	29.18	-	20.48**	-1.42
Anxiety	Psychological	1667.74	8.66	+	11.75**	1.42
Sad	Psychological	2131.00	14.11	+	9.86*	1.06
Risk	Psychological	2154.16	14.75	+	9.27*	1.02
Anger	Psychological	2015.18	21.80	+	1.42	0.36
3rd Person Plural	Linguistic	1436.11	21.80	-	0.18	-0.13
3rd Person Singular	Linguistic	3381.80	48.10	-	0.03	-0.04

* $p < .05$, ** $p < .01$

Chapter 4: A General Conclusion

Overview

This dissertation sought to explore and understand the language of NSSI in public pro-NSSI social media websites in order to better inform research and clinical practice for adolescents and young adults who engage in NSSI. This dissertation examined a specific subset of a population of people who use NSSI as a means of coping. The studies sought to understand the broad language processes of these individuals in an effort to better understand what they are attempting to communicate to others and to provide insight into their thought process.

This dissertation addressed two critical areas of research. The first area was exploring the language of pro-NSSI public posts on Tumblr, which is a unique format because it is a mix of micro-blogging and social media. In 2017, Tumblr had over 345 million register accounts and it was still growing (Gordon, n.d.). Tumblr's high number of users made it easy to find enough data to examine pro-NSSI posts. The specific NSSI terms, the linguistic properties and the psychological properties were all examined for the purpose of this study.

The second area was exploring the language of pro-NSSI public posts on Reddit, which is an important social media site in terms of users and the time spent by users on the site. Reddit is the third most popular website, after Google and YouTube; it has surpassed Facebook in users and its users spent more time browsing Reddit than Facebook around the time when Facebook was having its data sharing scandal (Nguyen, 2018). It is also a useful place for analysis because Reddit's beliefs about free speech mean that the site does not ban user's content unless there is a clear threat of violence or harm (Nguyen, 2018). Due to the anonymous and minimally censored

nature of Reddit, there was a plethora pro-NSSI public posts available for analysis. Factors that were analyzed for the purpose of this study was specific NSSI terms, the linguistic properties and psychological properties.

This chapter will review the findings of both studies and compare the linkages between the two manuscripts. This chapter covers the following five sections: (a) a summary of Manuscript #1; (b) a summary of Manuscript #2; (c) a presentation and discussion of linkage between the two manuscripts; (d) a discussion on the contribution of this dissertation to the knowledge base; and (e) future research directions.

Summary of Manuscript #1

Manuscript #1 explored the language used in pro-NSSI blogs on Tumblr. This study examined whether there were differences in the use of language in pro-NSSI blogs when they are compared to blogs overall. Three research questions were addressed in this study. The first research question explored the NSSI specific content word use. The second research question explored the linguistic processes of the posts. The third research question explored the use of psychological processes in the posts.

The findings from the first research question demonstrated that the most common NSSI specific words fell into the category of “reasons for the NSSI.” The most commonly occurring words in this category were feel, help and pain. These findings supported other studies that found that NSSI behaviors are a cry of pain for the distressing feelings they are experiencing (Williams, 2000). Research shows that

people who engage in NSSI are often experiencing distressing thoughts and feelings and use NSSI as a way to reduce and cope with the pain in the present.

The findings from the second research question demonstrated that first person plural was the most frequent and statistically significant pattern. This finding aligns with research that states that people who have experienced a trauma often write in first person plural pronouns due to their feelings of numbness and dissociation (Pennebaker, 2013). People who engage in NSSI often report feeling numbness. Pennebaker also found that first person plural pronouns are a way to avoid the use of I words which is a way to avoid pain and vulnerability.

The third research question pertaining to psychological word use patterns demonstrated that the most common pattern was negative emotion language. Negative emotion language is an expected outcome from this study. People who engage in NSSI are likely experiencing great emotional pain and it makes sense that this would be reflected in their writing (Pennebaker, 2013). There is also research supporting a correlation between negative emotion word use and low self-esteem (Pennebaker, Mehl, & Niederhoffer, 2003). People who engage in NSSI are likely to experience low self-esteem along with the emotional pain that they are in.

This study is the first to examine the language of NSSI on Tumblr. This is also one of the first studies to explore the language of NSSI on any public social media website. This study was able to find markers that demonstrates a level of mental illness and great emotion pain.

There are some limitations to this study that should be considered when analyzing the data. There are many meanings to a single word and thus can be

categorized in many different categories in the language analysis software. This has the potential to slightly skew the data. Another limitation is that only English was included in the corpus due to the native language of the researcher. The third limitation was due to the privacy settings on Tumblr. Only public posts on Tumblr were included in the corpus, which limits the posts that are public to those who are logged into a Tumblr account. This limited the amount of data that was able to be collected. The last limitation is that this study, while it had a rather large effect size may not be generalizable to the entire population of people who post on pro-NSSI social media and blog websites.

While this research has limitations, there are practical implications for this study. This research seeks to give practitioners a better understanding of the language of NSSI, which has been explored in very few studies. There is little information available on the understanding of NSSI language, behavior and what this could mean for improved treatment for this population. There are two practical implications and one implication to research that emerged from this study. The first implication is that often counselors are not aware that their clients are engaging in NSSI and post about their NSSI behaviors online. It is the counselor's ethical responsibility to assess for safety issues with their clients, including NSSI. If counselors are not upholding their ethical mandate, the counselor could be potentially doing more harm to the client.

The second implication is about the choice of treatment for NSSI individuals. It was evident in the language of the Tumblr posts that much of the language was filled with shame. Since NSSI behaviors like cutting are alarming and present some risk to the client's safety, many counselors use theories such as Motivational

Interviewing (MI) and Cognitive Behavioral Therapy (CBT) to treat patients (Washburn et al., 2012). However, this study's linguistic analysis, which attempts to learn more about the mental state of individuals who persist in NSSI, suggest that the client has high levels of shame and a harsh internal critic that is fueled by that sense of shame. Acknowledging these levels of shame and the presence of a harsh internal critic should inform a mental health professional's approach, and suggests the use of other theories, such as Compassion Focused Therapy (CFT). CFT has roots in CBT but addresses shame language and thoughts by focusing on self-compassion.

Summary of Manuscript #2

The second study explored the language of pro-NSSI posts on Reddit. This study compared if there were differences in the use of language of pro-NSSI blogs to blogs overall. Three research questions were addressed in this study. The first research question explored NSSI specific content word use. The second research question explored the linguistic processes of the post. And the third research question explored the use of psychological processes in the posts.

The findings from the first research question established that the most common NSSI specific words fell into the category of "reasons for the NSSI". The most common words were feel, help and scar. This data showed a strong desire to get help for their NSSI behaviors and a high focus on their scars. Some of the Reddit entries talked about their scars in both a prideful manner and yet shameful and negative. There is a myth that NSSI is connected to suicidal intent. The posts on Reddit did not show that these individuals engaged in NSSI as a means to end their life. These findings support previous studies that have found that individuals who

engage in suicidal self-injury do not discuss their behavior while individuals who engage in NSSI do discuss their behaviors with friends, family and mental health professionals (Baetens, Claes, Muehlenkamp, Grietens & Onghena, 2011).

The second research question concerning linguistic patterns in pro-NSSI blogs demonstrated statistically significant results. The results revealed that first person singular was the most frequent and significant pattern. According to Pennebaker (2013), people's pronoun use tracks their focus. Therefore, a high use of first-person singular pronouns demonstrates that the individual is most likely experiencing anxiety, self-consciousness, pain and depression. This aligns with the findings from the first research question and the third research question concerning psychological processes. It stands to reason that people who are struggling with mental illness such as depression and anxiety is indicated by their pronoun use.

The third research question pertaining to psychological word use patterns demonstrated negative emotion language. This finding supports Pennebaker, Mehl and Niederhoffer's (2003) that high levels of negative emotion language is positively correlated with neuroticism. People who have high levels of neuroticism also have high levels of anxiety, self-doubt, depression, shyness and other negative feelings (Neuroticism). This is in line with people who engage in NSSI because they too experience anxiety and depression.

One key limitations of this study is that words can have multiple meanings based on context and it is difficult for language software to pick up on these subtle nuances of the English language. Thus, there is the potential for some words to have

been miscategorized and, thus, for the data to be skewed. It is also important to be careful to make inferences into the data because this study may not be generalizable.

These results have major implications for dispelling the myth that NSSI is connected to being suicidal. These results are exciting for the field of mental health because this is a very common myth that still pervades counseling. It is the hope that these results can improve the treatment of clients who engage in NSSI and help counselors to better understand the behaviors of those who use NSSI. The results of this study can help counselors to better understand and communicate with the inner worlds of these clients.

This study also demonstrates that social media usage can often be a maladaptive way for these individuals to cope; while these individuals are using their writing to reach out to people in a community of others with similar challenges as a method for managing their NSSI behaviors and urges, they are not doing so effectively. As Pennebaker (2013) found, writing using negative emotion words and a negative tone is not helpful for relieving distressing feelings. Counselors should use this new information to guide their practice. For instance, narrative therapy could be one treatment modality in which the counselor can use the writings of their clients and help them to reframe it to use more positive emotion words to be more beneficial for their treatment outcome (Hoffman & Kress, 2008). By using a treatment modality that utilizes what the client is already doing and changing it to be more adaptive, the counselor may be better able to build a healthy rapport and to understand their client better.

Thematic Linkage of the Two Manuscripts

Both studies produced in this dissertation shared some common themes. The first manuscript explored the language of pro-NSSI language, linguistic processes and psychological process of posts made on Tumblr. The second manuscript explored the same properties but explored pro-NSSI posts made on Reddit. While Tumblr and Reddit may be similar, they serve very different in the populations. Many argue that Tumblr is a website that is populated with individuals who are more liberal and more closed-minded to other's opinions (Debate.org, n.d.) Those who frequent Reddit say that they have more open-minded discussions and experience less hate talk than on Tumblr (Debate.org, n.d.). While it appears that some people may participate on both Reddit and Tumblr, there seems to be a distinct difference between the two websites.

The results of these studies also demonstrated similar themes. Both studies resulted in similar findings for the first research question regarding NSSI specific language. Both studies found that the reasons for NSSI was the most common category of words used. Both studies also exhibited the same two most common used words which were feel and help. This is an interesting finding that suggests that both studies support each other in that people who engage in NSSI are seeking help and trying to cope with their high emotional distress.

The results were also similar for the third research question which pertained to the psychological processes. Both studies found that negative emotion words were the most common. These studies appear to support each other in that while these individuals are attempting to cope with distressing feelings, who are experiencing a level of mental illness, most likely to be anxiety or depression based on the content of their descriptions. These are important results because it highlights that these

individuals are struggling with mental illness and maladaptive behaviors. Counseling theories used to counter these behaviors need to take into account the level of negative emotion words so that they don't push their clients into action when they are not ready to do work to change behaviors.

Contribution to the Knowledge Base

Overall research has been focused more on NSSI due to the increased occurrence of these behaviors amongst young people. This dissertation is one of the first to examine the language use and language processes from people who engage in NSSI and post about it on social media and blogs. This is also one of the first studies to create a NSSI terms dictionary and compare the dictionary to known pro-NSSI blog posts. While the more recent studies involving pro-NSSI content on social media, the focus of these studies has been on the images that this population's posts and how these images reinforce NSSI behavior. This dissertation aims at understanding language markers and psychological markers that can give insight into the mental health and better understanding to the NSSI behaviors and thought processes of this population.

Professional counselors are often timid to ask directly about NSSI behaviors because very few know how to properly treat this population. Many professional counselors often treat NSSI clients using a behavior modification-based theory such as Cognitive Behavioral Therapy or Motivational Interviewing. While these are two empirically supported theories that demonstrate high effectiveness, the results from this dissertation show that shame, negative feelings and self-criticism is a major issue for this population. This suggests that a theory that first addresses the shame and self-

criticism is important in the treatment of this population. A theory such as Compassion Focused Therapy which addresses self-criticism and shame will be a good start with these clients struggling with NSSI.

This research suggests that counselors should be assessing for NSSI behaviors in their clients to make sure that their clients are safe and are not at risk for accidental suicide. An important aspect for counselors to assess for also, is the online behavior of their clients. Due to high levels of shame clients may not say outright that they are engaging in NSSI and posting about it on social media. Therefore, it is imperative for counselors to fulfill their ethical duty and assess for these behaviors because it is so prominent in many young people's lives.

Implications of Dissertation for Research Agenda

The research in this dissertation has opened many pathways for the future. The study of language and the application to mental health is still a new methodology in the field of counseling. It is interesting how as counselors we can examine language to have a better understanding of our clients. One study in the future that would be interesting, would be applying the data that was scraped from Tumblr and Reddit and analyze it for other aspects of language that can better help counselors. For example, the data could be analyzed for hedges, which are ways that people use words to soften what they are speaking or writing about. Such as, sort of, kind of or just are some common examples of hedges (Hyland, 2019). These phrases were often found in the data from Tumblr and Reddit but due to the nature of this study, this data was not analyzed for this dissertation. In addition to hedges I would also like to

analyze the data in terms of gendered language in an effort to understand which gender is posting more on social media their NSSI behaviors.

Another path of research from dissertation would be intervention-based studies. Since this research is exhibiting information that suggests that Compassion Focused Therapy would be a good fit for theory and client, it would be interesting to have a study in which progress is measured by the client's writing and assessing their language use, especially for negative emotion words. We would know if the client is improving if they are using an increased amount of positive emotion words rather than negative emotion words. This type of study has yet to be done on individuals who use NSSI to cope with distressing feelings.

References

- Debate.org. (n.d.). *Tumblr or Reddit; which is better?* Retrieved from <https://www.debate.org/opinions/tumblr-yes-or-reddit-no-which-is-better>
- Gordon, K. (n.d.). *Topic: Tumblr- Statistics and Facts*. Retrieved from <https://www.statista.com/topics/2463/tumblr/>
- Hoffman, R. M., & Kress, V. E. (2008). Narrative therapy and non-suicidal-self-injurious behavior: Externalizing the problem and internalizing personal agency. *The Journal of Humanistic Counseling, Education and Development*, 47, 157-171. <http://dx.doi.org/10.1002/j.2161-1939.2008.tb00055.x>
- Hyland, K. (2019). *Metadiscourse: Exploring interaction in writing*. NY: Bloomsbury Academic.
- Nguyen, C. (2018, May 30). *Reddit Is the Third-Most-Popular Destination on the Internet*. Retrieved from <https://www.digitaltrends.com/computing/reddit-more-popular-than-facebook-in-2018/>
- Pennebaker, J. W. (2013). *The secret life of pronouns: What our words say about us*. New York: Bloomsbury Press.
- Pennebaker, J. W., Mehl, M. R., & Niederhoffer, K. G. (2003). Psychological Aspects of Natural Language Use: Our Words, Our Selves. *Annual Review of Psychology*, 54, 547-577. <http://dx.doi.org/10.1146/annurev.psych.54.101601.145041>
- Washburn, J. J., Richardt, S. L., Styer, D. M., Gebhardt, M., Juzwin, K. R., Yourek, A., & Aldridge, D. (2012). Psychotherapeutic approaches to non-suicidal self-injury in adolescents. *Child and Adolescents Psychiatry and Mental Health*, 6(14). <https://doi.org/10.1186/1753-2000-6-14>

Williams, J. M. G., & Pollock, L. R. (2000). The psychology of suicidal behaviour. In K. Hawton and K. van Heeringen K (Eds.), *The international handbook of suicide and attempted suicide* (pp. 79-93). Chichester, UK: John Wiley & Sons. <https://doi.org/10.1002/9780470698976.ch5>

Bibliography

- Aalai, A. (2018, June 8). *The social contagion of suicide*. Retrieved from <https://www.psychologytoday.com/us/blog/the-first-impression/201806/the-social-contagion-suicide>
- Adrian, M., Zeman, J., Erdley, C., Lisa, L., & Sim, L. (2010). Emotional dysregulation and interpersonal difficulties as risk factors for non-suicidal self-injury in adolescent girls. *Journal of Abnormal Child Psychology*, *39*, 389-400. <https://dx.doi.org/10.1007/s10802-010-9465-3>
- Allen, J. E. (2011, November 17). 1 in 12 Teens cut, deliberately hurt themselves. *ABC News*, ABC News Medical Unit. Retrieved from <https://abcnews.go.com/Health/MindMoodNews/12-teens-cut-harm-themselves/story?id=14969232>
- Arensman, E., Larkin, C., Corcoran, P., Reulbach, U., & Perry, I. J. (2013). Factors associated with self-cutting as a method of self-harm: Findings from the Irish National Registry of Deliberate Self-Harm. *The European Journal of Public Health*, *24*(2), 292-297. doi:10.1093/eurpub/ckt087
- Baetens, I., Claes, L., Muehlenkamp, J., Grietens, H., & Onghena, P. (2011). Non-suicidal and suicidal self-injurious behavior among Flemish adolescents: A Web-Survey. *Archives of Suicide Research*, *15*, 56-67. <https://dx.doi.org/10.1080/13811118.2011.540467>
- Brown, R. C., Fischer, T., Goldwisch, A. D., Keller, F., Young, R., & Plener, P. L. (2018). #cutting: Non-Suicidal Self-Injury (NSSI) on Instagram.

Psychological Medicine, 48, 337-346.

<https://dx.doi.org/10.1017/S0033291717001751>

Calmenson, B., & Sullivan, P. (2018, February 06). *Why social media challenges are spiraling out of control*. Retrieved March 31, 2018, from

<https://www.washingtonexaminer.com/why-social-media-challenges-are-spiraling-out-of-control>

Camarota, S. A., & Zeigler, K. (2014, October 3). One in five U.S. residents speaks foreign language at home, record 61.8 million. Retrieved from

<https://cis.org/One-Five-US-Residents-Speaks-Foreign-Language-Home-Record-618-million>

Chancellor, S., Lin, Z., Goodman, E. L., Zerwas, S., & Choudhury, M. D. (2016).

Quantifying and predicting mental illness severity in online pro-eating disorder communities. *Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work & Social Computing - CSCW 16*, 1171-1184. <https://dx.doi.org/10.1145/2818048.2819973>

Choudhury, M. D., Counts, S., & Horvitz, E. (2013). Major life changes and

behavioral markers in social media. *Proceedings of the 2013 Conference on Computer Supported Cooperative Work - CSCW 13*.

<https://dx.doi.org/10.1145/2441776.2441937>

Corpus Linguistics. (n.d.). Retrieved from

https://www1.essex.ac.uk/linguistics/external/clmt/w3c/corpus_ling/content/glossary.html

- Cressie, N., & Read, T. R. (1989). Pearson's χ^2 and the loglikelihood ratio statistic G2: a comparative review. *International Statistical Review/Revue Internationale de Statistique*, 57, 19-43. <http://dx.doi.org/10.2307/1403582>
- Debate.org. (n.d.). *Tumblr or Reddit; which is better?* Retrieved from <https://www.debate.org/opinions/tumblr-yes-or-reddit-no-which-is-better>
- Dekmezian, G. (2016, September 23). Why do people blog? The benefits of blogging. Retrieved from https://www.huffingtonpost.com/gary-dekmezian/why-do-people-blog-the-be_b_8178624.html
- Deutsch, M., & Gerard, H. B. (1955). A study of normative and informational social influences upon individual judgment. *The Journal of Abnormal and Social Psychology*, 51, 629-636. <http://dx.doi.org/10.1037/h0046408>
- Ernhout, C., & Whitlock, J. (2014). *Bringing up self-injury with your clients*. The Fact Sheet Series, Cornell Research Program on Self-Injury and Recovery. Cornell University, Ithaca, NY.
- Eveleth, R. (2013, July 22). *What happened to 'self-harm blogs' after Tumblr banned them?* Retrieved from <https://www.smithsonianmag.com/smart-news/what-happened-to-self-harm-blogs-after-tumblr-banned-them-15883320/>
- Explainer: What is Tumblr? (2018, August 08). Retrieved from <https://www.webwise.ie/parents/explainer-what-is-tumblr-2/>
- Eysenck, H. J., & Eysenck, M. W. (1985). *Personality and individual differences: a natural science approach*. NY: Plenum Press.

- Favazza, A. R. (1998). The coming of age of self-mutilation. *The Journal of Nervous & Mental Disease*, 186, 259-268. <https://dx.doi.org/10.1097/00005053-199805000-00001>
- Gilbert, P. (2010). *Compassion focused therapy*. NY: Routledge.
- Gordon, K. (n.d.). *Topic: Tumblr- Statistics and Facts*. Retrieved from <https://www.statista.com/topics/2463/tumblr/>
- Han, B., & Baldwin, T. (2011). Lexical normalisation of short text messages: Makn sens a #twitter. In *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies-Volume 1* (pp. 368-378). Stroudsburg PA: Association for Computational Linguistics. Retrieved from: <https://dl.acm.org/citation.cfm?id=2002520>
- Han, B., & Baldwin, T. (2011). Lexical normalisation of short text messages: Makn sens a #twitter. In *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies-Volume 1* (pp. 368-378). Stroudsburg PA: Association for Computational Linguistics. Retrieved from: <https://dl.acm.org/citation.cfm?id=2002520>
- Hardie, A. (2014). *Log Ratio – an informal introduction*. Lancaster, UK: Centre for Corpus Approaches to Social Science. Retrieved from <http://cass.lancs.ac.uk/log-ratio-an-informal-introduction/>
- Harriagian, K. (2018, October 07). *Geocoding without geotags: A text-based approach for reddit*. Retrieved from <https://export.arxiv.org/abs/1810.03067>

- Harris, I. M., & Roberts, L. M. (2013). Exploring the use and effects of deliberate self-harm Websites: An Internet-Based Study. *Journal of Medical Internet Research, 15*. <http://dx.doi.org/10.2196/jmir.2802>
- Hashtag. (n.d.). Retrieved from <https://www.merriam-webster.com/dictionary/hashtag>
- Healthyplace.com. (2016). *Self-injurious behavior, self-injury treatment*. San Antonio, TX: Author. Retrieved from <https://www.healthyplace.com/abuse/articles/self-injurious-behavior-self-injury-treatment>
- Hoffman, R. M., & Kress, V. E. (2008). Narrative therapy and non-suicidal-self-injurious behavior: Externalizing the problem and internalizing personal agency. *The Journal of Humanistic Counseling, Education and Development, 47*, 157-171. <https://dx.doi.org/10.1002/j.2161-1939.2008.tb00055.x>
- Huh, Y. E., Vosgerau, J., & Morewedge, C. K. (2014). Social defaults: Observed choices become choice defaults. *Journal of Consumer Research, 41*, 746-760. <https://dx.doi.org/10.1086/677315>
- Hyland, K. (2019). *Metadiscourse: Exploring interaction in writing*. NY: Bloomsbury Academic.
- Jarvi, S., Jackson, B., Swenson, L., & Crawford, H. (2013). The impact of social contagion on non-suicidal self-injury: A review of the literature. *Archives of Suicide Research, 17*, 1-19. <https://dx.doi.org/10.1080/13811118.2013.748404>

- Kamarudin, N. S., Rakesh, V., Beigi, G., Manikouda, L., & Liu, H. (2018). A study of Reddit-users response to rape. *2018 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM)*.
<https://dx.doi.org/10.1109/asonam.2018.8508855>
- Kane, A. (2017, October 07). *What you can say if you see my self-harm scars*. Retrieved from <https://www.intrusivethoughts.org/blog/can-say-see-self-harm-scars/>
- Lewis, S. P., & Seko, Y. (2015). A double-edged sword: A review of benefits and risks of online non-suicidal self-injury Activities. *Journal of Clinical Psychology, 72*, 249-262. <http://dx.doi.org/10.1002/jclp.22242>
- Li, F., & Du, T. C. (2014). Listen to me — Evaluating the influence of micro-blogs. *Decision Support Systems, 62*, 119-130.
<https://dx.doi.org/10.1016/j.dss.2014.03.008>
- Marchant, A., Hawton, K., Stewart, A., Montgomery, P., Singaravelu, V., Lloyd, K., . . . John, A. (2018). A systematic review of the relationship between internet use, self-harm and suicidal behaviour in young people: The good, the bad and the unknown. *Plos One, 13*. <https://dx.doi.org/10.1371/journal.pone.0193937>
- Mars, B., Heron, J., Biddle, L., Donovan, J. L., Holley, R., Piper, M., . . . Gunnell, D. (2015). Exposure to, and searching for, information about suicide and self-harm on the Internet: Prevalence and predictors in a population based cohort of young adults. *Journal of Affective Disorders, 185*, 239-245.
<https://dx.doi.org/10.1016/j.jad.2015.06.001>

- McEnery, T., & Hardie, A. (2012). *CORPUS LINGUISTICS: Method, theory and practice*. Cambridge, UK: Cambridge University Press.
- McGlashan, M. (Ed.). (2018). Corpus: Some key terms. *CASS: Briefings*. Retrieved from <http://cass.lancs.ac.uk/wp-content/uploads/2013/12/CASS-Gloss-final1.pdf>
- Moreno, M. A., Ton, A., Selkie, E., & Evans, Y. (2016). Secret Society 123: understanding the language of self-harm on Instagram. *Journal of Adolescent Health, 58*, 78-84. <https://doi.org/10.1016/j.jadohealth.2015.09.015>
- Mueller, A. S. (2017). Does the media matter to suicide?: Examining the social dynamics surrounding media reporting on suicide in a suicide-prone community. *Social Science & Medicine, 180*, 152-159. <http://dx.doi.org/10.1016/j.socscimed.2017.03.019>
- Neff, K. D. (2009). The role of self-compassion in development: A healthier way to relate to oneself. *Human Development, 52*, 211-214. <http://dx.doi.org/10.1159/000215071>
- Neuroticism. (n.d.). Retrieved from <https://www.psychologytoday.com/us/basics/neuroticism>
- Nguyen, C. (2018, May 30). *Reddit Is the Third-Most-Popular Destination on the Internet*. Retrieved from <https://www.digitaltrends.com/computing/reddit-more-popular-than-facebook-in-2018/>
- Nicol, W. (2018, July 20). What is Reddit? A Beginner's Guide to the Front Page of the Internet. Retrieved from <https://www.digitaltrends.com/social-media/what-is-reddit/>

- O'Reilly, M., Dogra, N., Whiteman, N., Hughes, J., Eruyar, S., & Reilly, P. (2018). Is social media bad for mental health and wellbeing? Exploring the perspectives of adolescents. *Clinical Child Psychology and Psychiatry*, 23, 601-613. <https://dx.doi.org/10.1177/1359104518775154>
- Pennebaker, J. W. (2013). *The secret life of pronouns: What our words say about us*. New York: Bloomsbury Press.
- Pennebaker, J. W., Mehl, M. R., & Niederhoffer, K. G. (2003). Psychological Aspects of Natural Language Use: Our Words, Our Selves. *Annual Review of Psychology*, 54, 547-577. <https://dx.doi.org/10.1146/annurev.psych.54.101601.145041>
- Pennebaker, J.W., Boyd, R.L., Jordan, K., & Blackburn, K. (2015). The development and psychometric properties of LIWC2015. Austin, TX: University of Texas at Austin. <https://dx.doi.org/10.15781/T29G6Z>
- Pennebaker, J.W., Chung, C.K., Ireland, M., Gonzales, A., & Booth, R.J., (2007). *The development and psychometric properties of LIWC2007*. Austin, TX: University of Texas at Austin.
- Portmann, J. (2004). *Bad for us: The lure of self-harm*. Beacon Press.
- Rayson, P., & Garside, R. (2000). Comparing corpora using frequency profiling. In *Proceedings of the workshop on comparing corpora-Volume 9* (pp. 1-6). Stroudsburg, PA: Association for Computational Linguistics. <http://dx.doi.org/10.3115/1117729.1117730>

- Rosenthal, R. J., Rinzler, C., Wallsh, R., & Klausner, E. (1972). Wrist-cutting syndrome: The meaning of a gesture. *American Journal of Psychiatry*, *128*, 1363-1368. <https://doi.org/10.1176/ajp.128.11.1363>
- Self-injurious Behavior, Self-injury Treatment. (2010, February 22). Retrieved from <https://www.healthyplace.com/abuse/articles/self-injurious-behavior-self-injury-treatment>
- Setzekorn, M. (2018, January 25). *Peer pressure factor in social media Tide Pod Challenge*. Retrieved March 31, 2018, from http://www.dailytoreador.com/news/peer-pressure-factor-in-social-media-tide-pod-challenge/article_8bbcd9ba-0188-11e8-9ec6-bb3b7faa067f.html
- Shatz, S. M. (2004). The relationship between Horney's three neurotic types and Eysenck's PEN model of personality. *Personality and Individual Differences*, *37*, 1255-1261. <https://doi.org/10.1016/j.paid.2003.12.009>
- Skegg, K. (2005). Self-harm. *Lancet*, *366*, 1471-1483. [https://doi.org/10.1016/S0140-6736\(05\)67600-3](https://doi.org/10.1016/S0140-6736(05)67600-3)
- Smith, A., & Anderson, M. (2018, September 19). *Social Media Use 2018: Demographics and Statistics*. Washington, DC: Pew Research Center. Retrieved from <http://www.pewinternet.org/2018/03/01/social-media-use-in-2018/>
- Spiegel, A. (2005, June 10). *The history and mentality of self-mutilation*. Retrieved from <http://www.npr.org/templates/story/story.php?storyId=4697319>

- Tausczik, Y. R., & Pennebaker, J. W. (2009). The psychological meaning of words: LIWC and computerized text analysis methods. *Journal of Language and Social Psychology, 29*, 24-54. <https://dx.doi.org/10.1177/0261927x09351676>
- Textese. (n.d.). Retrieved from <https://www.dictionary.com/browse/textese>
- The Conversation Starts Here. (n.d.). Retrieved from <https://www.redditinc.com/>
- Van Vliet, K., & Kalnins, G. (2011). A compassion-focused approach to nonsuicidal self-injury. *Journal of Mental Health Counseling, 33*, 295-311. <https://dx.doi.org/10.17744/mehc.33.4.j7540338q223t417>
- Washburn, J. J., Richardt, S. L., Styer, D. M., Gebhardt, M., Juzwin, K. R., Yourek, A., & Aldridge, D. (2012). Psychotherapeutic approaches to non-suicidal self-injury in adolescents. *Child and Adolescents Psychiatry and Mental Health, 6*(14). <https://doi.org/10.1186/1753-2000-6-14>
- Weisser, M. (2016). *Practical corpus linguistics: An introduction to corpus-based language analysis*. NY: John Wiley & Sons.
- Weisser, M. (2016). *Practical corpus linguistics: An introduction to corpus-based language analysis*. NY: John Wiley & Sons.
- Williams, J. M. G., & Pollock, L. R. (2000). The psychology of suicidal behaviour. In K. Hawton and K. van Heeringen K (Eds.), *The international handbook of suicide and attempted suicide* (pp. 79-93). Chichester, UK: John Wiley & Sons. <https://doi.org/10.1002/9780470698976.ch5>
- Zeng, M. L. (2005). *4. Semantic relationships used in controlled vocabularies*. Retrieved from <http://marciazeng.slis.kent.edu/Z3919/42equivalency.htm>

Appendices

Appendix A

IRB Determination of Non-Human Subjects



Oregon State University
Research Office

Human Research Protection Program
& Institutional Review Board
B308 Kerr Administration Bldg, Corvallis OR 97331
(541) 737-8008
IRB@oregonstate.edu
<http://research.oregonstate.edu/irb>

Date of Notification	05/22/2018	Study Number	8618
Notification Type	Oversight Determination		
Principal Investigator	Cass Dykeman		
Study Team Members	Mandy Greaves		
Study Title	A corpus linguistic analysis of public Tumblr blog posts on non-suicidal self-injury		
Funding Source	None	Cayuse Number	N/A

DETERMINATION: RESEARCH, BUT NO HUMAN SUBJECTS

It has been determined that your project, as submitted, does meet the definition of research but **does not** involve human subjects under the regulations set forth by the Department of Health and Human Services 45 CFR 46.

Additional review is not required for this study.

Please do not include HRPP contact information on any of your study materials.

Note that amendments to this project may impact this determination.

The federal definitions and guidance used to make this determination may be found at the following link: [Human Subject](#)

Appendix B:
Greaves Non-Lethal Suicide Word Dictionary

Author Note

In this dictionary, an asterisk marks where a word stem was used.

Methods of NSSI

bang*

bite

biting

bleed*

break*

bruise*

burn*

carve*

cut

cutt*

embed*

erase

erasing

excoriate*

hit*

hit*

ingest*

insert*

pick*

pick*

pierc*

pinch*

poison*

pull*

rip*

rub*

scrape*

scraping

scratch*

stab*

stick*

tattoo*

tear*

Cutting-Specific Terms

cut

cutt*

NSSI Terms

non-suicidal

nssi

para-suicidal

self-abuse

self-harm

self-injury

self-mutilation

self-wounding

Instruments Used

bleach

eraser

fingernail

knife

needle

razor

Reasons for NSSI

angry

attention

bad

bore*

control*

copy

desperate

empty

escape

feel*

help*

notice

numb*

pain

punish*

reaction

relax*

relief

understand

Hidden Hashtag Terms

ehtilb

mauled

mysecretfamily,

selfharmmmm,

selfinjuryyy,

secretsociety_123

