

Nonsuicidal Self-Injury and Substance Use Disorders: A Shared Language of Addiction

Salvatore Giorgi^{1,2}, McKenzie Himelein-Wachowiak², Daniel Habib²,
Lyle H. Ungar¹, Brenda Curtis²

¹University of Pennsylvania, ²National Institute on Drug Abuse
sggiorgi@sas.upenn.edu, brenda.curtis@nih.gov

Abstract

Nonsuicidal self-injury (NSSI), or the deliberate injuring of one's body without intending to die, has been shown to exhibit many similarities to substance use disorders (SUDs), including population-level characteristics, impulsivity traits, and comorbidity with other mental disorders. Research has further shown that people who self-injure adopt language common in SUD recovery communities (e.g., "clean", "relapse", "addiction," and celebratory language about sobriety milestones). In this study, we investigate the shared language of NSSI and SUD by comparing discussions on public Reddit forums related to self-injury and drug addiction. To this end, we build a set of LDA topics across both NSSI and SUD Reddit users and show that shared language across the two domains includes SUD recovery language in addition to other themes common to support forums (e.g., requests for help and gratitude). Next, we examine Reddit-wide posting activity and note that users posting in *r/selfharm* also post in many mental health-related subreddits, while users of drug addiction related subreddits do not, despite high comorbidity between NSSI and SUDs. These results show that while people who self-injure may contextualize their disorder as an addiction, their posting habits demonstrate comorbidities with other mental disorders more so than their counterparts in recovery from SUDs. These observations have clinical implications for people who self-injure and seek support by sharing their experiences online.

1 Introduction

Nonsuicidal self-injury (NSSI), or the intentional injuring of one's body without aiming to die for reasons outside of social norms, causes significant morbidity (Nock, 2010). Lifetime prevalence of NSSI is estimated to range from 5-6% among adults to 17-18% among adolescents (Swannell et al., 2014) while the prevalence of NSSI among

adolescents with psychiatric disorders is thought to be much higher (Nock and Prinstein, 2004; Glenn and Klonsky, 2013).

Qualitative, quantitative, mixed-methods, and psychometric studies have pointed towards the addictive features of NSSI and shared characteristics between NSSI and substance use disorders (SUDs; Brown and Kimball, 2013; Davis and Lewis, 2019). Populations who self-injure may resemble populations with SUDs in personality (MacLaren and Best, 2010) or impulsivity traits (Dir et al., 2013). NSSI and SUD are often comorbid with anxiety, depressive, and psychotic disorders (Guvendeger Doksat et al., 2017) and with each other, with one study reporting that approximately 60% of people who self-injure met criteria for a SUD (Nock et al., 2006). Both substance use and NSSI are used to avoid and/or cope with feelings of psychological distress (Chawla and Ostafin, 2007), especially among adolescents (Peterson et al., 2008). Although there is debate as to whether or not NSSI is an addiction in a clinical sense or experienced as intensely as SUD (Victor et al., 2012), addiction models of NSSI have been proposed (Faye, 1995; Buser and Buser, 2013; Blasco-Fontecilla et al., 2016). There is also evidence that those who endorse more addictive features of self-injury harm themselves more frequently and more severely (Martin et al., 2013), have higher levels of internalized anger (Nixon et al., 2002), and are at increased risk for accidentally harming to a greater extent than intended (Buser et al., 2017) and attempting suicide (Csorba et al., 2009). Some have urged clinicians to consider addictive features of NSSI when treating people who self-injure (Blasco-Fontecilla et al., 2016), making the addictive aspects of NSSI a valuable research target. Previous work has also identified the adoption of language used in SUD recovery circles Alcoholics Anonymous (AA) and Narcotics Anonymous (NA) in NSSI communities. This "addiction

language" (i.e., phrases such as "relapse", "recovery" and celebration of time without self-injury) has been found on NSSI message boards (Whitlock et al., 2006), Facebook groups (Niwa and Mandrusiak, 2012), and LiveJournal (Davis and Lewis, 2019).

This study builds on the work of Himelein-Wachowiak et al. (2022), who investigated the use of "addiction language" and experiences of addiction in the *r/selfharm* subreddit. This was done through a text-based annotation process where experts in addiction and recovery adapted the Diagnostic and Statistical Manual of Mental Disorders Fifth Edition (DSM-5) criteria for SUD to NSSI as a method for measuring the symptoms and severity of addiction to NSSI for Reddit users. Results showed that over three-quarters of the sample met the criteria for addiction and 86% used "addiction language". This work also builds on a large body of research using Reddit as a tool for mental health applications (De Choudhury and De, 2014), which includes depression (Pirina and Çöltekin, 2018), anxiety (Shen and Rudzicz, 2017), suicide (Zirikly et al., 2019), substance use (Lu et al., 2019), schizophrenia (Zomick et al., 2019), and DSM-5 evaluations (Gaur et al., 2018).

The purpose of this study is to further evaluate experiences of addiction by examining shared "addiction language" between NSSI and SUD communities using automated methods. We begin by directly comparing NSSI and SUD subreddits through a set of LDA topics estimated over a corpus of Reddit comments and examine how themes of addiction and recovery are used across both communities. While previous studies have identified "addiction language" in NSSI communities, to our knowledge, none have directly compared NSSI and SUD communities. We also identify where users of these subreddits are posting across Reddit in order to identify common communities. We end by discussing the clinical implications of our findings.

2 Data

Self-injury We looked at posts from the *r/selfharm* subreddit from the Pushshift Reddit Data set (Baumgartner et al., 2020). We focused on *r/selfharm* based on the fact that it had the highest number of posts and users and the most diverse discussion around NSSI. See Supplemental Materials for descriptions of other self-injury related subreddits as well as temporal trends in post histories.

Substance Use In order to compare the *r/selfharm* users to those posting in SUD recovery communities, we gather the posting activity from all users who have posted in the following subreddits: *r/addiction*, *r/alcoholism*, *r/opiatesrecovery*, *r/leaves*, *r/stopdrinking*, and *r/redditorsinrecovery*. These were manually selected due to high post volume and focus on recovery (vs. drug use itself).

Our data set consisted of both comments and submissions (i.e., the first post in a Reddit thread) across 2019 in order to match the temporal span of Himelein-Wachowiak et al. (2022). Across both data sets, we removed any accounts with the word "bot" in the user handle, after manually inspecting the account to confirm that the account is indeed a bot, as well as deleted posts, deleted accounts, and moderators. We also removed any redditor who posted in both *r/selfharm* and one or more SUD subreddits, in order to remove the possibility that common users will drive shared language. To identify redditors who are active in their respective communities, we remove any redditor with less than 10 comments, resulting in 2,470 *r/selfharm* who together posted 77,414 comments. We then identified a matched sample of 2,470 SUD redditors (posting 77,424 comments), approximately matched on both comment and submission counts.

3 Methods

Task 1: Shared Language To examine shared language across NSSI and SUD subreddits, we estimate a set of Content Specific LDA topics (Zamani et al., 2020). Content Specific LDA (CSLDA) is a method for estimating LDA topics across a thematically concise corpus and has been previously used to model conversations around excessive drinking, diabetes, and Black Lives Matter tweets (Giorgi et al., 2020; Griffis et al., 2020; Giorgi et al., 2022). CSLDA contains a preprocessing pipeline that identifies words related to the theme in question (i.e., NSSI and SUD) by comparing this data to a background corpus of general language (i.e., data from *r/AskReddit*). This removes language that is specific to Reddit as opposed to being NSSI or SUD related. We create CSLDA topics across a combined corpus of comments from NSSI and SUD subreddits. See Supplemental Materials for full details of the CSLDA pipeline.

We use the Mallet Java software wrapper within the DLATK Python package (McCallum, 2002; Schwartz et al., 2017). All default settings are used,

alpha	25 Topics		50 Topics		75 Topics		100 Topics	
	TU	Coh.	TU	Coh.	TU	Coh.	TU	Coh.
1	0.98	0.46	0.85	0.38	0.72	0.33	0.61	0.28
3	0.99	0.31	0.90	0.37	0.72	0.32	0.65	0.34
5	0.99	0.41	0.90	0.36	0.78	0.37	0.66	0.36

Table 1: Topic quality as measured through Topic Uniqueness (TU) and Coherence (Coh.).

and we evaluate a range of $\alpha \in \{1, 3, 5\}$, a prior on the number of topics per document, and topic set sizes $K \in \{25, 50, 75, 100\}$. All topics are quantitatively and qualitatively evaluated. Quantitative evaluation consists of two metrics: coherence and topic uniqueness. Coherence measures semantic similarity between the words in the topic using Normalized Pointwise Mutual Information (NPMI; Syed and Spruit, 2017). This is calculated for each topic and then averaged across all topics. Topic uniqueness (TU), a measure of topic diversity, is inversely proportional to the number of times a set of L keywords is repeated across a set of K topics (Nan et al., 2019). Thus, a topic set with high TU means that the representative keywords are rarely repeated across topics. While past research has used a value of $L = 10$ (Nan et al., 2019), we set $L = 30$ to be more conservative (with a large L , the probability of a given word appearing in more than one topic will increase, thus decreasing TU). TU ranges between 1 and $1/K$, and therefore we normalize TU to be between 0 and 1, since we are evaluating topic sets of sizes K .

Qualitative evaluation consisted of manually inspecting topics for three criteria: (1) breadth of themes, (2) minimal thematic overlap, and (3) a single topic contains a single theme. Note that (2) and (3) are similar to TU and coherence.

Task 2: Posting Activity Here, we look at all posts (i.e., submissions and comments) across the whole of Reddit in 2019 from our disjoint samples of NSSI and SUD redditors. For the NSSI redditors, we gather 1,019,796 of their posts in subreddits other than *r/selfharm*. For the SUD redditors, we gather 927,733 posts to subreddits other than the 6 addiction subreddits used to collect the sample. We then reported the most frequently visited subreddits for both the NSSI and SUD samples and calculated the percentage of users posting in each.

4 Results

Task 1: Shared Language In Table 1, we evaluate the CSLDA topics. Here we see Topic Uniqueness (TU) decrease as the number of topics grows. This is to be expected since as the number of topics grows one can expect words to be shared across a larger number of topics. TU also increases with α within a fixed topic set (i.e., column-wise). Coherence shows no clear pattern across α or topic set size. Through the qualitative evaluation, $K = 50$ topics with $\alpha = 5$ were chosen as the most interpretable. Since TU is high across the $K = 50$ topic sets and coherence is reasonable with $\alpha = 5$ (i.e., neither the highest nor the lowest value across all topic sets), we proceed with this topic set.

Figure 1 shows the average user-level topic usage across all 50 topics, ordered by the difference between the NSSI (green) and SUD (blue) groups. In total, we see 9 out of 50 topics include addiction or recovery-related keywords within the top 10 highest weighted words in the topic. We do not include topics that contain both “clean” and “cut” since “clean” most likely does not refer to “staying clean” in the recovery sense. Additionally, we see that the most similar topic is addiction-related (“addiction”, “addicted”), as well as the 6th (“recovery”, “relapsed”, “clean”). Notably, the remaining addiction topics are the least similar (i.e., towards the

NSSI users		SUD users	
Subreddit	% Users	Subreddit	% Users
AskReddit	52.8	AskReddit	47.7
depression \diamond	37.8	pics	21.9
MadeOfStyrofoam \dagger	34.4	funny	21.8
SuicideWatch \diamond	33.9	aww	21.3
SelfHarmScars \dagger	29.0	todayilearned	17.0
teenagers	26.2	AmItheAsshole	16.3
memes	21.1	Showerthoughts	16.2
Showerthoughts	18.7	mildlyinteresting	15.1
aww	16.8	relationship_advice	14.9
AmItheAsshole	15.5	news	14.9
unpopularopinion	15.3	worldnews	14.5
funny	14.4	politics	14.4
wholesomememes	14.0	gifs	13.6
dankmemes	13.6	trashy	12.8
mentalhealth \diamond	13.5	unpopularopinion	12.8
2meirl4meirl	13.0	interestingasfuck	12.6
Anxiety \diamond	12.6	gaming	12.6
mildlyinteresting	12.6	PublicFreakout	12.4
offmychest \diamond	12.0	videos	12.1
relationship_advice	11.8	tifu	11.2

Table 2: Most popular subreddits, defined as the percentage of users within each group that post in a given subreddit (% Users). \dagger and \diamond are self-injury and mental health related subreddits, respectively.

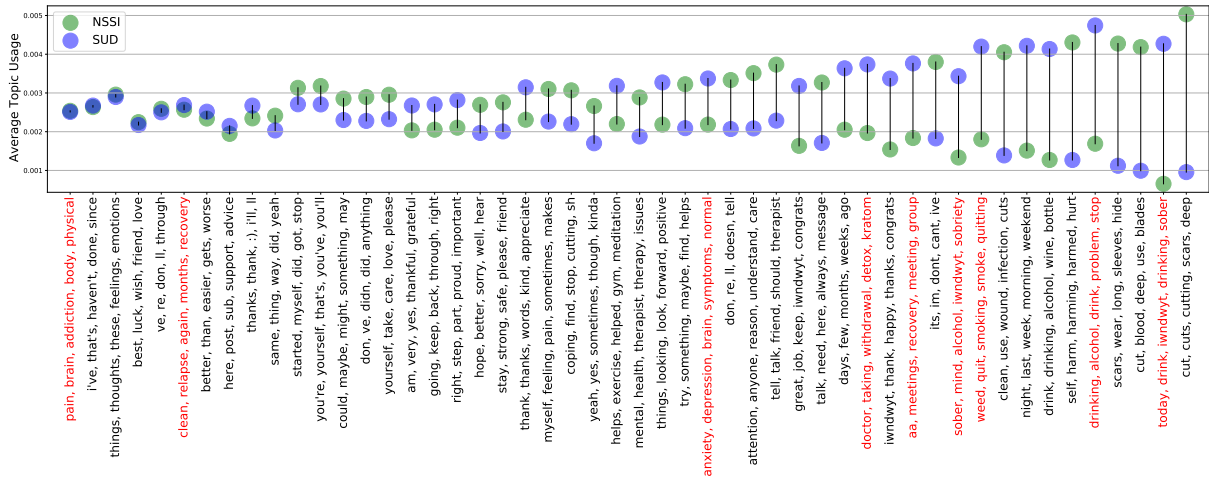


Figure 1: Differences in average user-level topic frequency. The top 5 weighted words from each topic are visualized with red topics containing an addiction keyword (e.g., “sober” and “relapse”) in the top 10 weighted words.

right side of the figure). These later topics contain either substance use keywords (and are, thus, used more often by SUD users) or contain words like “withdrawal” which was one of the least prevalent DSM-5 symptoms (6% of users) found in Himelein-Wachowiak et al. (2022).

There are also a number of topics that do not contain recovery keywords but seem to be related to addiction and recovery such as mentions of starting/stopping (“started”, “stop”) and getting easier (“gets”, “easier”). While not addiction or recovery related, we do see shared language of gratitude (“thanks” and “thank”), support (“better”, “easier”, “gets” as well as “support”, “advice”), emotions (“thoughts”, “feelings”, “emotions”), and coping.

Task 3: Posting Activity Table 2 shows the 20 subreddits in which the highest percentage of NSSI and SUD redditors are also posting. Not surprisingly, we see that NSSI redditors are posting in other NSSI related subreddits: *r/MadeOfStyrofoam* and *r/SelfHarmScars*. On the other hand, SUD redditors are not posting in other substance-related subreddits outside of the six used to collect data. We also note that there are a number of mental health related subreddits in which NSSI redditors are posting, which is in line with common NSSI comorbidities: *r/depression*, *r/SuicideWatch*, and *r/Anxiety*. We do not see similar posting activity in mental health subreddits among SUD redditors despite high comorbidity between SUD and both depression and anxiety (Conway et al., 2006).

5 Conclusion

In this work, we directly compared language across large online communities dedicated to discussing NSSI and SUD recovery. We showed that there is indeed a shared language of addiction between these two communities, evidenced by equivalent usage of topics related to addiction (“addiction”, “addicted”) and recovery (“recovery”, “relapsed”, “clean”). To our knowledge, this is the first study using automated methods to quantify addiction language in NSSI communities, as well as the first to directly compare language between online NSSI and SUD recovery forums. We also examined Reddit-wide posting activity and showed that, while NSSI redditors posted in a number of mental health related subreddits, SUD redditors did not even though both NSSI and SUD are comorbid with many of the same mental health disorders.

One limitation of our study is the high comorbidity between NSSI and SUD: 60% of adolescents engaging in NSSI meeting criteria for SUDs (Nock et al., 2006). Thus, the shared addiction language may be a result of NSSI redditors also having and discussing SUDs. We attempted to control for this by excluding redditors who are posting in both NSSI and SUD subreddits. Additionally, Himelein-Wachowiak et al. (2022) noted that only 2% of NSSI redditors explicitly mentioned having a SUD.

Despite this limitation, our results suggest that the adoption of addiction and recovery language in NSSI communities may provide psychological benefit to the users and help them cope with self-injury. Himelein-Wachowiak et al. (2022) posits that alignment with SUD may buffer against self-

stigma and encourage adoption of common SUD recovery strategies. In a similar study, Pritchard et al. (2021) suggest that people who self-injure use addiction messages to convey the difficulty in stopping, as well as to caution those considering NSSI as a coping strategy. Our results also suggest that NSSI redditors seek support in similar communities for other mental health concerns, perhaps with the goals of broadening their support network or seeking specific advice for a separate mental disorder. The lack of mental health cross posting among SUD redditors may imply a need for more discussion of comorbid mental disorders among SUD recovery communities as well as greater engagement with people dealing with other mental disorders. Regardless, NSSI and SUD recovery communities share similar language of support ("yourself", "take", "care", "love") and encouragement ("better", "easier", "gets"), illustrating the broad psychological benefits of sharing intimate experiences with empathetic others online, regardless of the particular mental health concern.

6 Ethical Considerations

NSSI communities and their members tend to refer to NSSI as "self harm." In this paper, we use "self-injury" or the acronym NSSI as it is more specific to the behavior in question ("self harm" could also include suicide attempts) as well as the term most frequently found in recent literature. However, papers we cite may use the terms "deliberate self harm" (DSH) or "self-mutilation." For a review and discussion of the most appropriate language to use when referring to people who self-injure, see (Hasking et al., 2021).

There are a number of ethical considerations when using sensitive data. Since Reddit data is publicly available, this study was deemed non-human subjects research and exempt from approval of an Institutional Review Board. Despite this official classification, the data used throughout is indeed human generated and reflects the lived experiences, intimate feelings, and personal struggles of the authors. Related, there are issues regarding informed consent when using public data. Online communities such as *r/selfharm* are intimate and personal spaces, where consensual sharing happens between community members and not with researchers who collect the data. For a full discussion of related issues, we recommend the work of Chancellor et al. (2019), which identifies conflicting representations

of humans in "human centered machine learning." There are also issues of privacy; while Reddit is anonymous, there are risks of revealing sensitive information or the identities of the accounts used in the study (Proferes et al., 2021). As such, we only report aggregate information throughout the manuscript, and we have chosen to not publicly release any of the data used in this study. Finally, there are some egregious use cases with this data. For example, given the cross posting between NSSI and mental health forums, one could imagine ads for anti-depressants being targeted to the redditors in this study.

One must also consider researchers' well-being when working with data of this type. Spending time with sensitive and potentially triggering data can be emotionally challenging for researchers. As such, researchers should also consent to working with this type of data and continue to consent throughout the life of the project. To help with these issues, our research group held one-on-one and group sessions to discuss triggering content and mental and emotional fatigue experienced while working on this and similar projects.

Acknowledgements

This research was supported by the Intramural Research Program of the NIH, NIDA. We would like to thank Elise Bragard, Amy Kwarteng, Destiny Schriefer, Chase Smitterberg, and Kenna Yadeta for their help in the annotation process, as well as Garrick Sherman, Amanda Devoto, and the CLPsych reviewers for their thoughtful feedback.

References

- Jason Baumgartner, Savvas Zannettou, Brian Keegan, Megan Squire, and Jeremy Blackburn. 2020. The pushshift reddit dataset. In *Proceedings of the International AAAI Conference on Web and Social Media*, volume 14, pages 830–839.
- Hilario Blasco-Fontecilla, Roberto Fernández-Fernández, Laura Colino, Lourdes Fajardo, Rosa Perteguer-Barrio, and Jose De Leon. 2016. The additive model of self-harming (non-suicidal and suicidal) behavior. *Frontiers in Psychiatry*, 7:8.
- Tiffany B Brown and Thomas Kimball. 2013. Cutting to live: A phenomenology of self-harm. *Journal of Marital and Family Therapy*, 39(2):195–208.
- Trevor J Buser and Juleen K Buser. 2013. Conceptualizing nonsuicidal self-injury as a process addiction: Review of research and implications for counselor

- training and practice. *Journal of Addictions & Offender Counseling*, 34(1):16–29.
- Trevor J Buser, Juleen K Buser, and Corrine C Rutt. 2017. Predictors of unintentionally severe harm during nonsuicidal self-injury. *Journal of Counseling & Development*, 95(1):14–23.
- Stevie Chancellor, Eric PS Baumer, and Munmun De Choudhury. 2019. Who is the "human" in human-centered machine learning: The case of predicting mental health from social media. *Proceedings of the ACM on Human-Computer Interaction*, 3(CSCW):1–32.
- Neharika Chawla and Brian Ostafin. 2007. Experiential avoidance as a functional dimensional approach to psychopathology: An empirical review. *Journal of clinical psychology*, 63(9):871–890.
- Kevin P Conway, Wilson Compton, Frederick S Stinson, and Bridget F Grant. 2006. Lifetime comorbidity of dsm-iv mood and anxiety disorders and specific drug use disorders: results from the national epidemiologic survey on alcohol and related conditions. *The Journal of clinical psychiatry*, 67(2):10343.
- Janos Csorba, Elek Dinya, Paul Plener, Edit Nagy, and Eszter Páli. 2009. Clinical diagnoses, characteristics of risk behaviour, differences between suicidal and non-suicidal subgroups of hungarian adolescent outpatients practising self-injury. *European Child & Adolescent Psychiatry*, 18(5):309–320.
- Sarah Davis and Christopher Alan Lewis. 2019. Addiction to self-harm? the case of online postings on self-harm message boards. *International journal of mental health and addiction*, 17(4):1020–1035.
- Munmun De Choudhury and Sushovan De. 2014. Mental health discourse on reddit: Self-disclosure, social support, and anonymity. In *Eighth international AAAI conference on weblogs and social media*.
- Allyson L Dir, Kenny Karyadi, and Melissa A Cyders. 2013. The uniqueness of negative urgency as a common risk factor for self-harm behaviors, alcohol consumption, and eating problems. *Addictive behaviors*, 38(5):2158–2162.
- Pamela Faye. 1995. Addictive characteristics of the behavior of self-mutilation. *Journal of psychosocial nursing and mental health services*, 33(6):36–39.
- Manas Gaur, Ugur Kursuncu, Amanuel Alambo, Amit Sheth, Raminta Daniulaityte, Krishnaprasad Thirunarayan, and Jyotishman Pathak. 2018. "let me tell you about your mental health!" contextualized classification of reddit posts to dsm-5 for web-based intervention. In *Proceedings of the 27th ACM International Conference on Information and Knowledge Management*, pages 753–762.
- Salvatore Giorgi, Sharath Chandra Guntuku, McKenzie Himelein-Wachowiak, Amy Kwarteng, Sy Hwang, Muhammad Rahman, , and Brenda Curtis. 2022. Twitter corpus of the #blacklivesmatter movement and counter protests: 2013 to 2021. *Proceedings of the International AAAI Conference on Web and Social Media*.
- Salvatore Giorgi, David B Yaden, Johannes C Eichstaedt, Robert D Ashford, Anneke EK Buffone, H Andrew Schwartz, Lyle H Ungar, and Brenda Curtis. 2020. Cultural differences in tweeting about drinking across the us. *International Journal of Environmental Research and Public Health*, 17(4):1125.
- Catherine R Glenn and E David Klonsky. 2013. Nonsuicidal self-injury disorder: an empirical investigation in adolescent psychiatric patients. *Journal of clinical child & adolescent Psychology*, 42(4):496–507.
- Heather Griffis, David A Asch, H Andrew Schwartz, Lyle Ungar, Alison M Bittenheim, Frances K Barg, Nandita Mitra, and Raina M Merchant. 2020. Using social media to track geographic variability in language about diabetes: Infodemiology analysis. *JMIR diabetes*, 5(1):e14431.
- Neslim Guvendeger Doksat, Oguzhan Zahmacioglu, Arzu Ciftci Demirci, Gizem Melissa Kocaman, and Ayten Erdogan. 2017. Association of suicide attempts and non-suicidal self-injury behaviors with substance use and family characteristics among children and adolescents seeking treatment for substance use disorder. *Substance use & misuse*, 52(5):604–613.
- Penelope A Hasking, Mark E Boyes, and Stephen P Lewis. 2021. The language of self-injury: a data-informed commentary. *The Journal of Nervous and Mental Disease*, 209(4):233–236.
- McKenzie Himelein-Wachowiak, Salvatore Giorgi, Amy Kwarteng, Destiny Schriefer, Chase Smittberg, Kenna Yadeta, Elise Bragard, Amanda Devoto, Lyle Ungar, and Brenda Curtis. 2022. Getting "clean" from nonsuicidal self-injury: Experiences of addiction on the subreddit r/selfharm. *Journal of behavioral addictions*, 11(1):128–139.
- John Lu, Sumati Sridhar, Ritika Pandey, Mohammad Al Hasan, and Georege Mohler. 2019. Investigate transitions into drug addiction through text mining of reddit data. In *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, pages 2367–2375.
- Vance V MacLaren and Lisa A Best. 2010. Nonsuicidal self-injury, potentially addictive behaviors, and the five factor model in undergraduates. *Personality and Individual Differences*, 49(5):521–525.
- Jodi Martin, Paula F Cloutier, Christine Levesque, Jean-François Bureau, Marie-France Lafontaine,

- and Mary K Nixon. 2013. Psychometric properties of the functions and addictive features scales of the ottawa self-injury inventory: A preliminary investigation using a university sample. *Psychological assessment*, 25(3):1013.
- Andrew Kachites McCallum. 2002. Mallet: A machine learning for language toolkit (2002).
- Feng Nan, Ran Ding, Ramesh Nallapati, and Bing Xiang. 2019. Topic modeling with wasserstein autoencoders. In *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*, pages 6345–6381.
- Kendra D Niwa and Michael N Mandrusiak. 2012. Self-injury groups on facebook. *Canadian Journal of Counselling and Psychotherapy*, 46(1).
- Mary K Nixon, Paula F Cloutier, and Sanjay Aggarwal. 2002. Affect regulation and addictive aspects of repetitive self-injury in hospitalized adolescents. *Journal of the American Academy of Child & Adolescent Psychiatry*, 41(11):1333–1341.
- Matthew K Nock. 2010. Self-injury. *Annual review of clinical psychology*, 6:339–363.
- Matthew K Nock, Thomas E Joiner Jr, Kathryn H Gordon, Elizabeth Lloyd-Richardson, and Mitchell J Prinstein. 2006. Non-suicidal self-injury among adolescents: Diagnostic correlates and relation to suicide attempts. *Psychiatry research*, 144(1):65–72.
- Matthew K Nock and Mitchell J Prinstein. 2004. A functional approach to the assessment of self-mutilative behavior. *Journal of consulting and clinical psychology*, 72(5):885.
- John Peterson, Stacey Freedenthal, Christopher Sheldon, and Randy Andersen. 2008. Nonsuicidal self injury in adolescents. *Psychiatry (Edgmont)*, 5(11):20.
- Inna Pirina and Çağrı Çöltekin. 2018. [Identifying depression on Reddit: The effect of training data](#). In *Proceedings of the 2018 EMNLP Workshop SMM4H: The 3rd Social Media Mining for Health Applications Workshop & Shared Task*, pages 9–12, Brussels, Belgium. Association for Computational Linguistics.
- Tyler R Pritchard, Chelsey A Fedchenko, and Stephen P Lewis. 2021. Self-injury is my drug: the functions of describing nonsuicidal self-injury as an addiction. *The Journal of Nervous and Mental Disease*, 209(9):628–635.
- Nicholas Proferes, Naiyan Jones, Sarah Gilbert, Casey Fiesler, and Michael Zimmer. 2021. Studying reddit: A systematic overview of disciplines, approaches, methods, and ethics. *Social Media+ Society*, 7(2):20563051211019004.
- H Andrew Schwartz, Salvatore Giorgi, Maarten Sap, Patrick Crutchley, Lyle Ungar, and Johannes Eichstaedt. 2017. Dlatk: Differential language analysis toolkit. In *Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing: System Demonstrations*, pages 55–60.
- Judy Hanwen Shen and Frank Rudzicz. 2017. [Detecting anxiety through Reddit](#). In *Proceedings of the Fourth Workshop on Computational Linguistics and Clinical Psychology — From Linguistic Signal to Clinical Reality*, pages 58–65, Vancouver, BC. Association for Computational Linguistics.
- Sarah V Swannell, Graham E Martin, Andrew Page, Penelope Hasking, and Nathan J St John. 2014. Prevalence of nonsuicidal self-injury in nonclinical samples: Systematic review, meta-analysis and meta-regression. *Suicide and Life-Threatening Behavior*, 44(3):273–303.
- Shaheen Syed and Marco Spruit. 2017. Full-text or abstract? examining topic coherence scores using latent dirichlet allocation. In *2017 IEEE International conference on data science and advanced analytics (DSAA)*, pages 165–174. IEEE.
- Sarah Elizabeth Victor, Catherine Rose Glenn, and Elisha David Klonsky. 2012. Is non-suicidal self-injury an “addiction”? a comparison of craving in substance use and non-suicidal self-injury. *Psychiatry research*, 197(1-2):73–77.
- Janis L Whitlock, Jane L Powers, and John Eckenrode. 2006. The virtual cutting edge: the internet and adolescent self-injury. *Developmental psychology*, 42(3):407.
- Mohammadzaman Zamani, H. Andrew Schwartz, Johannes Eichstaedt, Sharath Chandra Guntuku, Adithya Virinchipuram Ganesan, Sean Clouston, and Salvatore Giorgi. 2020. [Understanding weekly COVID-19 concerns through dynamic content-specific LDA topic modeling](#). In *Proceedings of the Fourth Workshop on Natural Language Processing and Computational Social Science*, pages 193–198, Online. Association for Computational Linguistics.
- Ayah Zirikly, Philip Resnik, Özlem Uzuner, and Kristy Hollingshead. 2019. [CLPsych 2019 shared task: Predicting the degree of suicide risk in Reddit posts](#). In *Proceedings of the Sixth Workshop on Computational Linguistics and Clinical Psychology*, pages 24–33, Minneapolis, Minnesota. Association for Computational Linguistics.
- Jonathan Zomick, Sarah Ita Levitan, and Mark Serper. 2019. [Linguistic analysis of schizophrenia in Reddit posts](#). In *Proceedings of the Sixth Workshop on Computational Linguistics and Clinical Psychology*, pages 74–83, Minneapolis, Minnesota. Association for Computational Linguistics.