The Language of Minority Stress Experiences of Sexual Minority Communities on Reddit

A Thesis
Presented to
The Academic Faculty

By

Sang-Chan Kim

In Partial Fulfillment
of the Requirements for the Degree
Bachelor of Science in Computer Science from the
College of Computing

Georgia Institute of Technology

May 2019

The Language of Minority Stress Experiences of Sexual Minority Communities on Reddit

Approved by:

Dr. Munmun De Choudhury, Advisor
School of Interactive Computing
Georgia Institute of Technology

Dr. Rosa Arriaga
School of Interactive Computing
Georgia Institute of Technology

Date Approved:
April 30, 2019

Acknowledgement

Introduction

Evidence from the decades of research have shown deeply concerning mental health disparities surrounding sexual minorities, specifically the lesbian, gay, bisexual, transgender, and additional sexual and gender minority (also known as LGBTQ+) community (Schneider, 1989; Diaz, 2001; Remafedi, 1998; Tebbe, 2016). The evidence in these studies suggest heightened risk of mental illness and suicide in sexual minorities. Schneider et al. conducted questionnaire of gay youth to explore relationship between sexual orientation and suicidal behavior, discovering alarming rates of such behavior in their interviewees (Schneider, 1989). Another study found depression as the primary mediator of minority stress processes such as discrimination and internal/external anti-transgender stigma with risk of suicide (Tebbe, 2016). It also suggests that substance abuse and suicide risk are correlated, strengthening previous argument for proliferation of suicide risk and substance abuse in sexual minority communities. In addition, a study found high proportion of suicide ideation among gay Latino men from 3 different cities with 17 percent of the subjects thinking about suicide once or twice in the last 6 months (Diaz, 2001). Furthermore, lifetime proportion of attempted suicides was four times higher in gay and bisexual males than in heterosexual males.

In his 1995 study, Ilan H Meyer introduced the minority stress theory, a theoretical framework to conceptualize the unique mental health challenges of sexual minorities. In his theoretical framework, Meyer defined minority stress as "psychosocial stress derived from minority status" and introduced three minority stressors: perceived stigma, prejudice events, and internalized homophobia (Meyer, 1995). Research aiming to help improve support interventions and counseling efforts of sexual minorities have been hindered by the limitations of existing

quantitative methods, such as surveys and self-reports, of collecting data on these minority

stressors. Survey based approaches toward measuring minority stress suffer from many

challenges. Due to the large temporal gaps in which they are administered, it is difficult to gather

near real-time information on the mental health challenges experienced by LGBTQ+ individuals.

Moreover, responses are prompted by the experimenter and typically comprise recollection of

(sometimes subjective) health facts.

Online activities and communities have become popular amongst LGBTQ+ individuals

(Pascoe, 2011). Many are utilizing the Internet to find peers and to seek for help and social

support (McDermott, 2015; McDermott, 2012). Furthermore, recent social computing research

has made strides in showing the benefits of online communities (De Choudhury, 2014; Andalibi,

2016; De Choudhury; 2016). Specifically, online communities that afford users a level of

anonymity can provide a safe space for them to disclose personal experiences, feelings, and

issues that would normally be stigmatized by society. These online communities provide us a

with a new medium to further explore the minority stress theory, helping overcome the

hindrances of traditional survey or self-report based methods. This can hopefully provide insight

into how to better serve the sexual minority community through improved mental health

interventions, therapy and counseling methods, and public policy.

Our research achieves three primary objectives. First, we constructed a codebook, drawn

heavily from Meyer's minority stress theory (Meyer, 1995), in order to evaluate the presence of

minority stressors in posts made on social media. Specifically, we utilized posts made to r/lgbt, a

public Reddit community for LGBTQ+ individuals. We collected around 12.6K posts from the

r/lgbt community and hand-annotated 350 randomly sampled posts for absence/presence of

minority stressors using the codebook. Second, we developed a *minority stress classifier* that could identify the presence of minority stress in text using the 350 annotated posts from our first objective. We used popular social media analysis and natural language techniques to build a rich set of features and used them to train a classifier that could automatically identify posts for minority stress. Our classifier performed considerably well, scoring a AUC of 0.80 and F1 of 0.75. We utilized this classifier to produce a much larger labelled dataset of social media posts. Third, we analyzed the linguistic cues of these newly machine-labelled posts, built a lexicon of minority stress markers, and related our findings back to minority stress theory.

Background

Minority Stress

Meyer defines minority stress as "psychosocial stress derived from minority status" (Meyer, 1995). We refer to the complex social situations that cause minority stress as "stressors". Furthermore, stressors can be considered "distal" or "proximal". External and interpersonal factors like prejudice, discrimination, and rejection are considered distal stressors while internal processes like hiding one's minority identity or internalizing homophobia are considered proximal stressors. These stressors, over time, can lead to adverse mental health conditions in the individual. The Office of Disease Prevention and Health Promotions has identified a widespread of health disparaties that the LGBTQ community suffers from (Fisher, 2014). The National Alliance on Mental Illness states that LGBTQ indvidiuals are three more times likely to undergo mental health problems, such as depression or generalized anxiety disorder, compared to the rest of the population (Mustanski, 2010). Compared to heterosexuals, LGBTQ individuals experience higher levels of psychological processes (Meyer, 2003). These processes consist of emotion regulation, and social/interpersonal and cognitive processes, and they increase the risk for adverse mental health. Circumstances like poor social support, interpersonal prejudice and discrimination, socio-economic status are some of the many factors have shown to cause minority stress. Prejudice and discriminations act as distal stressors. When exposed to distal stressors, individuals began to experience proximal stressors in the form of internalized homophobia, for example. There has been numerous studies that conduct surveys in order to collect data on the prevalence of minority stress and its impact. One such study discovered

increased risk of substance abuse and psychiatric disorder in gay/lesbian and bisexual men and women (Kerridge, 2017). Another study found links between negative family treatment and depressive symptoms, heightened risk of suicide, and lower life satisfaction (Nguyen, 2016). A comprehensive analysis of recent population-based studies on LGBT and suicidal risk discovered strong evidence of higher rates of suicide attempts in the LGBT population in the United States (Haas, 2010). One such study found that gay and bisexual males are four times more likely to attempt suicide than heterosexuals (King, 2008). Another analysis suggests the increased suicide risk is more prominent in adolescents and young adults (De Graaf, 2006; Paul, 2002; Remafedi, 1991). Despite studies that support a strong correlation between poor mental health and suicide risk, one study found that even after controlling for mental health disparity, rate of attempted suicides were still substantially higher in LGB, suggesting that mental health is not the sole factor for heightened risk for attempted suicides (Belik, 2010).

Self-Disclosure on Social Media

Disclosure of one's own experiences to others can be a positive method to deal with one's emotional suffering (Rimé, 2007). Studies show that those people who disclose on social media sites often receive helpful feedback through social support as well forming meaningful connection with other members of the community (Haimson, 2015; Newman, 2011; Andalibi, 2019; Andalibi, 2017). Stigmatized experiences can also become less stigmatizing. Social media websites like Tumblr and Reddit provide a level of anonymity that is appropriate for self disclosure of personal matter. In fact, proliferation of anonymous "throwaway" accounts, which are accounts made temporarily in order to hide one's true identity on the platform, does not reduce the quality of social support (Andalibi, 2016; Andalibi, 2018). Research by Andalibi et al.

further suggest that throwaway accounts can actually help men facing societal expectation of masculinity comfortably disclose past experiences of sexual abuse (Andalibi, 2016). De Choudhury et al. utilizes online mental health communities (OHMCs) as a medium to study stigmatized populations suffering from mental health problems. One study analyzed OHMCs on Reddit like r/SuicideWatch and r/depression and found that self-disclosures on these communities aligned with clinical literature on mental health (De Choudhury, 2014). Responses to self-disclosure were predominately informational, instructional, emotional, and generally of high quality. Another study by De Choudhury et al. has shown successful reduction in depression after receiving social support following self-disclosure (De Choudhury, 2016).

Data

This paper studies the language shared by self-disclosing LGBTQ+ individuals on social media. In our study, we utilized data from Reddit, a widely used semi-anonymous online social forum. The platform organizes discussions into a variety of sub-communities called subreddits. Reddit harbors a variety of subreddits (or communities) that focus on LGBTQ+ issues that can be surfaced based on Reddit's subreddit search feature, such as r/lgbt, r/LGBTPolitics, r/lgbtnews, r/dixiequeer, r/lgbtaww, r/ainbow and others. Since our work focuses on mental health in particular, we first manually examined all of these subreddits to identify those where such issues are most frequently discussed. Subreddits like r/lgbtnews were deemed as not relevant due to their primary focus on recent news about LGBTQ+ issues. Additionally factoring in the size of the community into account, we directed our attention to the subreddit r/lgbt only. The r/lgbt community defines itself as a safe space for members of GSRM (Gender, Sexual, and Romantic Minority) to "discuss their lives, issues, interests, and passions" and welcomes all GSRM members "beyond lesbian, gay, bisexual, and transgender people." It has been active for the last nine years, has considerable traffic, and is heavily moderated by a group of 25 moderators who ensure access to high quality content. As of March 2019, the subreddit has about 295,000 subscribers. We collected two and half years of posts published on the r/lgbt subreddit from a Google BigQuery database that hosts the entire corpus of Reddit data. We filtered posts for ones that were "text-only" and were not removed or deleted by the author or moderator. The filtered dataset contained a total of 12,645 authored posts. The timestamps of the posts ranged from January 2016 to May 2018. Fig. 1 provides the top keywords in the dataset as well as

distribution of post lengths. We randomly sampled 350 posts from this dataset for hand
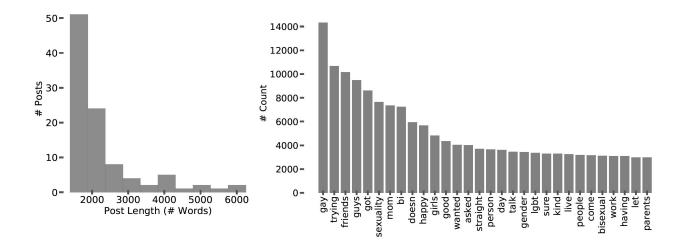
annotation for signals of minority stress.

Figure 1. Left: histogram of lengths of posts (i.e. number of words in a post); Right: top 30 keywords found in the collected corpus

Methods

Meyer's minority stress model identifies three categories that minority stress can fall under - prejudice, perceived stigma, and internalized homophobia. Our collaborators at University of Michigan used a directed coding methodology (Hsieh, 2005), deriving heavily from the existing Meyer minority stress model theory (Meyer, 1995), to develop our own codebook for identifying minority stress from social media language. Our collaborators' codebook differed from Meyer's model in some ways. In contrast to Meyer's 1995 study that utilized closed-end scale ratings from gay men who were asked questions about minority stress, our codebook was developed using a recent dataset of posts authored by a diverse community of LGBTQ+ individuals seeking support and advice in a computer-mediated, social-media context. Developing the codebook helped our collaborators consistently and systematically annotate a training set of 350 Reddit posts for minority stress.

Each collaborators initially annotated the same 20 posts sampled randomly from our main dataset. They discussed any differences in their individual annotations and revised their initial codebook to reflect on agreeable themes and concepts. Next, they annotated 30 more randomly selected posts and repeated the aforementioned process to produce a more refined codebook. Our collaborators were able to achieve high inter rater agreement, achieving an overall Cohen's k coefficient of 0.91. Using this refined codebook, the collaborators then annotated 300 more posts. Of the 350 total annotated posts, nearly half were identified to contain minority stress. Our collaborators found that the minority stress categories (prejudice events, perceived stigma, and internalized homophobia) had major overlap in many of the posts. However, this was a

negligible concern since the end goal of the annotation task is to find overall signs of minority stress.

Prejudice Events: Actions of rejection towards an individual
    1. Violence towards and individual
        a. Verbal Violence
            i. Directing slurs towards minority individuals (before or after disclosure)
            ii. Rejecting sexuality
        b. Physical Violence
    2. Legal discrimination in housing, employment, entitlements, and basic civil rights
        a. Excluded based on identity
        b. Lack of sexuality education in schools
    3. Anxiety/stress about discriminatory political powers
    4. Includes prejudice events from those in other subsets of the LGBTQ+ population (e.g., violence perpetuated from cisgender LGBQ people to trans people; violence perpetuated from gay people to bisexualpeople)

Perceived Stigma: fears and anxiety that come from self about how others perceive them
    1. Individuals constantly monitoring their behavior: how one dresses, speaks, walks, and talks in contrast to expected social norms.
    2. Lying to cover up identity
    3. Thinking others do not or will not "accept" them
    4. Feelings of not "fitting in"
    5. Anxiety/stress about potential future discriminatory political powers
    6. Includes perceived stigma from other subsets of the LGBTQ+ population (e.g., stigma felt by trans people from cisgender LGBQ people; stigma felt by bisexual people from gay people)

Internalized LGBTphobia: internalization of negative societal attitudes, and negative feelings that come from self about self
    1. Self-hate for being a part of the LGBTQ+ community
    2. Applying negative attitudes/dislike to themselves related to LGBTQ+ identity
    3. Feelings of negativity rooted from LGBTQ+ identity
    4. Feelings of isolation or of ending up alone
    5. Hate/distaste for others within LGBTQ+ community

Table 1. The minority stress codebook for social media

Next, we leveraged social media analysis and natural language processing techniques on our annotated dataset to build a machine learning model that could identify presence of minority stress. We refer to this as the *minority stress classifier*. First, we performed standard tokenization and removal of stop words from the corpus. Next, we built word embeddings to capture latent semantics. Word embeddings, which are essentially vectorized representation of tokenized texts, have proven to be useful tools for solving problems in the natural language processing space

(Chandrasekharan, 2018; Mikolov, 2013). We then utilized Linguistic Inquiry and Word Count or "LIWC" to capture the psycholinguistic cues from textual data. Given a piece of text, LIWC outputs the number of words reflecting various psycholinguistic categories. Each psycholinguistic category comes with a predetermined vocabulary drawn from dominant theories in psychology, business, medicine, and common sense. LIWC has been shown to be useful in previous research assessing social media and psychological well being (De Choudhury, 2013; Saha, 2017; Schwartz, 2013). The categories used in this study are

1. Affective attributes - anger, anxiety, negative and positive affect, sadness, swear

2. Cognitive attributes - causation, inhibition, cognitive mechanics, discrepancies, negation, tentativeness, certainty

3. Perception attributes - feel, hear, insight, see

4. Interpersonal focus - first person singular, second person plural, third person plural, indefinite pronoun

5. Temporal reference - future tense, past tense, present tense

6. Lexical density and awareness - adverbs, verbs, article, exclusive, inclusive, preposition, quantifier

7. Biological concerns - bio, body, death, health, sexual

8. Personal concerns - achievement, home, money, religion

9. Social concerns - family, friends, humans, social

In order to capture the use of offensive language used against LGBTQ+ community, we utilized lexicon developed from research on social media and hate speech language to assess the presence or absence of gender and sexual orientation related hate speech (Mondal, 2017). We

also incorporated top 500 n-gram (n ≤ 3) features as well as sentiment analysis tool from the Stanford CoreNLP library to compute how positive, negative, or neutral each post in the corpus was (Manning, 2014). As noted before, minority stress literature indicates heightened risk of mental health problems in the community we are studying. We utilized classification output from a previously developed machine learning classifier to identify presence of mental health symptoms as additional set of features (Saha, 2017; Saha, 2018). Given a post, the classifier outputs the presence or absence of each of four mental health conditions - depression, anxiety, sadness, and suicidal ideation, lending the name *DASS classifier*. Using a corpus with a newly generated set of rich features (659 in total), we trained a variety of common machine learning models such as Naive Bayes, Logistic Regression, Random Forest, Support Vector Machine, and Multi-layer Perceptron using stratified five fold cross validation to find optimal hyperparameters.

Results and Discussion

The multi-layer perceptron (MLP) classifier achieved the highest performance metric, scoring median AUC (area under the curve) of 0.8, median precision of 0.75, and median recall of 0.74. Performance between folds deviated little (σ=0.03). Listed below are the model performances for each model. All trained models performed better than the baseline accuracy of 58% which was achieved using a model based on chance.

Model	Precision	Recall	F1	AUC
Naive Bayes	0.70	0.49	0.53	0.54
Logistic Regression	0.73	0.72	0.72	0.76
Linear SVM	0.74	0.74	0.74	0.77
Random Forest	0.76	0.67	0.70	0.75
Adaboost	0.73	0.72	0.72	0.77
MLP	0.75	0.74	0.75	0.80

Table 2. the various performances metrics of the trained machine learning models

In order to compute the relative importance of each feature, we applied a univariate feature selection method using mutual information to obtain K highest scoring features. We used ANOVA to establish statistical significance. Non-interpretable features like word embeddings were excluded from this analysis. Large portions of features associated with the affective, perceptive, and cognitive attributes from LIWC as well as features associated with DASS scored high in importance. The results from the feature importance test strengthen key discoveries from prior literature on minority stress and mental health (Ernala, 2017; Pennebaker, 2003; Saha, 2017). High scoring affective attributes like *anxiety*, *negative affect*, etc align with minority

stress categories like perceived stigma and prejudice events. Perceptive attributes like *insight*,

see, and *feel* emphasize the prevalence of communication, especially when used with social

attributes like *friends* and *family*, in posts expressing minority stress. Prior studies show high

utilization of cognitive attributes by those describing first-hand experiences and events, aligning

with the high relative importance of cognitive attributes in our study (Boals, 2005; Brubaker,

2012). High relative importance of DASS can be explained by prior literature that show

LGBTQ+ communities are prone to heightened mental healths risks.

Feature	Score	Feature	Score
DASS - Depression	0.14***	LIWC - Past Tense	0.09***
DASS - Anxiety	0.13***	LIWC - Family	0.09***
LIWC - Tentativeness	0.13***	LIWC - Article	0.09***
LIWC - Negative Affect	0.12***	LIWC - Present Tense	0.08***
DASS - Suicidal Ideation	0.12	LIWC - Cognitive Mechanism	0.08***
LIWC - Positive Affect	0.12***	LIWC - Indefinite Pronouns	0.08***
LIWC - Adverbs	0.11***	LIWC - 1st Person Singular	0.08***
LIWC - Sexual	0.10**	LIWC - Humans	0.08***
LIWC - Discrepancies	0.11***	LIWC - Negation	0.08***
DASS - Stress	0.11**	LIWC - Future Tense	0.08***
LIWC - Causation	0.10***	LIWC - Conjunction	0.08***
LIWC - Anxiety	0.10***	LIWC - Social	0.08***
LIWC - Verbs	0.10***	LIWC - Percept	0.08***
LIWC - Certainty	0.10***	LIWC - Inhibition	0.08***
LIWC - Quantifier	0.10***	LIWC - Health	0.08***
LIWC - Achievement	0.10***	LIWC - Friends	0.08***
LIWC - Insight	0.10***	n-gram - card	0.08***
LIWC - Exclusive	0.09***	ngram - lady	0.08***

LIWC - Sadness	0.09***	Hate Lexicon: "Homophobic"	0.08***
LIWC - Bio	0.09**	LIWC - Swear	0.05***
LIWC - Preposition	0.09***	n-gram - appreciated	0.04***
LIWC - Inclusive	0.09***	Sentiment: Negative	0.04***
LIWC - Work	0.09***		

Table 3. Top 45 features. The p-values reported after Bonferroni correction following ANOVA
(*** p <0.0001, **p<0.001, *p<0.1)

Using this minority stress classifier, we labelled all 12.6K posts in our dataset to dive

further into the language of minority stress posts. To obtain discriminating n-grams between

posts that exhibit minority stress and posts that do not, we supplied labelled posts to SAGE

(Sparse Additive Generative Model) algorithm (Eisenstein, 2011). This allowed us to get a

ranking of tokens that best differentiated language with a presence of minority stress versus

language with an absence of minority stress. Some of the top keywords for minority stress

related to communication align with literature about self-disclosure and communication about

ones LGBTQ identity and the difficulty associated with it. However, among top n-grams

frequent in posts that do not exhibit minority stress, we see high frequency of phrases like *thanks

in advance*, *really appreciate*, *thanks for reading*, *lgbtq community*, and *lgbt friendly* which

denote neither self-disclosure nor communication of one's personal experiences. Instead, these

keywords express the lack of distress from one's minority status and signals desire for

community, diversity, and inclusiveness. This suggests that posts without minority stress lean

towards topics surrounding generic, non-personal matters surrounding the LGBTQ+ community.

Posts with minority stress				Posts without minority stress			
token	SAGE	token	SAGE	token	SAGE	token	SAGE

like shit	2.09	started talking	0.82	feel free	-1.46	year old male	-0.79
tell mom	1.74	months ago	0.78	thanks advance	-1.37	thank reading	-0.77
didnt want	1.53	feel bad	0.76	just wondering	-1.35	past year	-0.77
im afraid	1.52	didn care	0.76	recently came	-1.16	wanted know	-0.74
acting like	1.10	started feel	0.71	let know	-1.15	old male	-0.71
really bad	1.10	life just	0.69	really appreciate	-1.15	just looking	-0.68
want live	1.07	point life	0.69	gender fluid	-1.05	attracted women	-0.68
telling people	1.02	just said	0.68	gay lesbian	-0.97	support lgbt	-0.67
like time	1.02	couple months	0.67	greatly appreciated	-0.92	long term	-0.67
got really	1.02	mom dad	0.67	really confused	-0.88	guys think	-0.66
point just	0.99	im sorry	0.65	hey guys	-0.88	lgbt community	-0.66
eye contact	0.99	people life	0.64	need help	-0.86	love hear	-0.64
days later	0.94	felt like	0.63	lgbtq community	-0.85	ask questions	-0.64
come home	0.94	wouldn want	0.63	sexually attracted	-0.84	questioning sexuality	-0.64
didn feel	0.94	time life	0.62	gay bisexual	-0.83	sexual orientation	-0.63
like maybe	0.93	act like	0.62	sexual attraction	-0.82	male female	-0.63
living life	0.93	ago started	0.60	wanted share	-0.82	lgbt friendly	-0.62
people come	0.83	dont understand	0.59	just phase	-0.80	sex men	-0.62

Table 4. Top discriminating keywords found in posts with minority stress vs posts without minority stress

We measure a minority stress post's similarity to one of the three minority stress categories - prejudice events, perceived stigma, internalized LGBTphobia - using a simple word embedding comparison methodology. Specifically, given a post, we obtain vectorized form using a pretrained Word2Vec model and compute its cosine similarity to the vectorized descriptions in the codebook under each minority stressor. If a post measured above a 0.80 threshold in cosine similarity with the description of one of the three minority stressor, it was labelled to exhibit that particular stressor (Rekabsaz, 2017). In posts most semantically close to the prejudice event

category, phrases like *didn't want*, *didn't feel*, and *didn't say* occur frequently. Frequent usage of

negation with a verb like *want*, *feel*, and *say* indicate common confrontation with negative

experiences associated with gender and sexual identity. Perceived stigma is defined as

"internalized fears and anxiety of expected rejection from others" by our codebook and in prior

literature. We also see high use of negations in posts categorized as "perceived stigma". Frequent

use of negations indicate inhibition, a behavior that can take in the form of hiding one's identity,

fear of rejection, worrying about potential discrimination due to unfavorable politics and policy,

or even stigma from other members in the LGBTQ+ community. Prior literature even indicates

high correlation between negation and inhibition (Creswell, 2007; Hancock, 2007). In addition,

high frequency of temporal phrases like *started talking*, *months after*, and *started feel* indicate

motivation to disclose their experiences and feelings. We also observe that posts falling under

"Internalized LGBTphobia" contain high use of phrases conveying expression of one's feelings

and emotions like *want to live* and *feel bad*. Combined with uses of phrases like *im gay* and

thought gay, it reflects internalization of LGBTphobia and heightened focus on one's own

emotions and behaviors.

Our research presents many useful applications within social media platforms and

beyond. Our classifier can be used to automate moderation of forums on social media with

significant LGBTQ+ members. Specifically, a system deployed on social media that incorporates

this classifier could automatically flag posts or users that exhibit signs of minority stressors and

its manifestations, allowing moderators to quickly connect users with professional individuals

who can provide the right assistance. As noted previously, disclosing one's personal experiences

and thoughts have also shown to be have positive effect on oneself. Our research suggests that

this this is true for the LGBTQ+ community. Journaling tools that encourage one to log their personal thoughts, emotions, struggles, etc can provide positive therapy while alleviating the harmful effects of minority stress. We also believe our research can provide valuable information for public health researchers studying this area as well as provide support for building mental health interventions adapted for LGBTQ+ individuals who utilize online communities like Reddit.

References

Schneider, S. G., Farberow, N. L., & Kruks, G. N. (1989). Suicidal behavior in adolescent and young adult gay men. *Suicide and Life‑Threatening Behavior*, *19*(4), 381-394.

Diaz, R. M., Ayala, G., Bein, E., Henne, J., & Marin, B. V. (2001). The impact of homophobia, poverty, and racism on the mental health of gay and bisexual Latino men: findings from 3 US cities. *American journal of public health*, *91*(6), 927.

Remafedi, G., French, S., Story, M., Resnick, M. D., & Blum, R. (1998). The relationship between suicide risk and sexual orientation: results of a population-based study. *American journal of public health*, *88*(1), 57-60.

Tebbe, E. A., & Moradi, B. (2016). Suicide risk in trans populations: An application of minority stress theory. *Journal of Counseling Psychology*, *63*(5), 520.

Meyer, I. H. (1995). Minority stress and mental health in gay men. *Journal of health and social behavior*, 38-56.

McDermott, E. (2015). Asking for help online: Lesbian, gay, bisexual and trans youth, self-harm and articulating the 'failed'self. *Health:*, *19*(6), 561-577.

McDermott, E., & Roen, K. (2012). Youth on the virtual edge: Researching marginalized sexualities and genders online. *Qualitative Health Research*, *22*(4), 560-570.

Pascoe, C. J. (2011). Resource and risk: Youth sexuality and new media use. *Sexuality Research and Social Policy*, *8*(1), 5-17.

De Choudhury, M., & De, S. (2014, May). Mental health discourse on reddit: Self-disclosure, social support, and anonymity. In *Eighth International AAAI Conference on Weblogs and Social Media*.

Andalibi, N., Haimson, O. L., De Choudhury, M., & Forte, A. (2016, May). Understanding social media disclosures of sexual abuse through the lenses of support seeking and anonymity. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems* (pp. 3906-3918). ACM.

Andalibi, N., Haimson, O. L., Choudhury, M. D., & Forte, A. (2018). Social support, reciprocity, and anonymity in responses to sexual abuse disclosures on social media. *ACM Transactions on Computer-Human Interaction (TOCHI)*, *25*(5), 28.

De Choudhury, M., Kiciman, E., Dredze, M., Coppersmith, G., & Kumar, M. (2016, May). Discovering shifts to suicidal ideation from mental health content in social media. In *Proceedings of the 2016 CHI conference on human factors in computing systems* (pp. 2098-2110). ACM.

Fisher, C. B., & Mustanski, B. (2014). Reducing health disparities and enhancing the responsible conduct of research involving LGBT youth. *Hastings Center Report*, *44*(s4), S28-S31.

Mustanski, B. S., Garofalo, R., & Emerson, E. M. (2010). Mental health disorders, psychological distress, and suicidality in a diverse sample of lesbian, gay, bisexual, and transgender youths. *American journal of public health*, *100*(12), 2426-2432.

Meyer, I. H. (2003). Prejudice, social stress, and mental health in lesbian, gay, and bisexual populations: conceptual issues and research evidence. *Psychological bulletin*, *129*(5), 674.

Kerridge, B. T., Pickering, R. P., Saha, T. D., Ruan, W. J., Chou, S. P., Zhang, H., ... & Hasin, D. S. (2017). Prevalence, sociodemographic correlates and DSM-5 substance use disorders and other psychiatric disorders among sexual minorities in the United States. *Drug and alcohol dependence*, *170*, 82-92.

Nguyen, Trang Quynh, et al. "Negative treatment by family as a predictor of depressive symptoms, life satisfaction, suicidality, and tobacco/alcohol use in Vietnamese sexual minority women." *LGBT health* 3.5 (2016): 357-365.

Haas, A. P., Eliason, M., Mays, V. M., Mathy, R. M., Cochran, S. D., D'Augelli, A. R., ... & Russell, S. T. (2010). Suicide and suicide risk in lesbian, gay, bisexual, and transgender populations: Review and recommendations. *Journal of homosexuality*, *58*(1), 10-51.

King, M., Semlyen, J., Tai, S. S., Killaspy, H., Osborn, D., Popelyuk, D., & Nazareth, I. (2008). A systematic review of mental disorder, suicide, and deliberate self harm in lesbian, gay and bisexual people. *BMC psychiatry*, *8*(1), 70.

De Graaf, R., Sandfort, T. G., & ten Have, M. (2006). Suicidality and sexual orientation: Differences between men and women in a general population-based sample from the Netherlands. *Archives of sexual behavior*, *35*(3), 253-262.

Paul, J. P., Catania, J., Pollack, L., Moskowitz, J., Canchola, J., Mills, T., ... & Stall, R. (2002). Suicide attempts among gay and bisexual men: Lifetime prevalence and antecedents. *American journal of public health*, *92*(8), 1338-1345.

Remafedi, G., Farrow, J. A., & Deisher, R. W. (1991). Risk factors for attempted suicide in gay and bisexual youth. *Pediatrics*, *87*(6), 869-875.

Belik, S. L., & Sareen, J. (2010). Sexual orientation and its relationship to mental disorders and suicidal behavior. In *annual conference of the American Association of Suicidology, Orlando, FL*.

Rimé, B. (2007). The social sharing of emotion as an interface between individual and collective processes in the construction of emotional climates. *Journal of social issues*, *63*(2), 307-322.

Andalibi, N. (2019). What Happens After Disclosing Stigmatized Experiences on Identified Social Media: Individual, Dyadic, and Social/Network Outcomes.

Andalibi, N., Ozturk, P., & Forte, A. (2017, February). Sensitive Self-disclosures, Responses, and Social Support on Instagram: the case of# depression. In *Proceedings of the 2017 ACM conference on computer supported cooperative work and social computing* (pp. 1485-1500). ACM.

Haimson, O. L., Brubaker, J. R., Dombrowski, L., & Hayes, G. R. (2015, February). Disclosure, stress, and support during gender transition on Facebook. In *Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing* (pp. 1176-1190). ACM.

Newman, M. W., Lauterbach, D., Munson, S. A., Resnick, P., & Morris, M. E. (2011, March). It's not that i don't have problems, i'm just not putting them on facebook: challenges and opportunities in using online social networks for health. In *Proceedings of the ACM 2011 conference on Computer supported cooperative work* (pp. 341-350). ACM.

Hsieh, H. F., & Shannon, S. E. (2005). Three approaches to qualitative content analysis. *Qualitative health research*, *15*(9), 1277-1288.

Chandrasekharan, E., Samory, M., Jhaver, S., Charvat, H., Bruckman, A., Lampe, C., ... & Gilbert, E. (2018). The Internet's Hidden Rules: An Empirical Study of Reddit Norm Violations

at Micro, Meso, and Macro Scales. *Proceedings of the ACM on Human-Computer Interaction*, *2*(CSCW), 32.

Mikolov, T., Chen, K., Corrado, G., & Dean, J. (2013). Efficient estimation of word representations in vector space. *arXiv preprint arXiv:1301.3781*.

De Choudhury, M., Gamon, M., Counts, S., & Horvitz, E. (2013, June). Predicting depression via social media. In *Seventh international AAAI conference on weblogs and social media*.

Saha, K., & De Choudhury, M. (2017). Modeling stress with social media around incidents of gun violence on college campuses. *Proceedings of the ACM on Human-Computer Interaction*, *1*(CSCW), 92.

Schwartz, H. A., Eichstaedt, J. C., Kern, M. L., Dziurzynski, L., Ramones, S. M., Agrawal, M., ... & Ungar, L. H. (2013). Personality, gender, and age in the language of social media: The open-vocabulary approach. *PloS one*, *8*(9), e73791.

Mondal, M., Silva, L. A., & Benevenuto, F. (2017, July). A measurement study of hate speech in social media. In *Proceedings of the 28th ACM Conference on Hypertext and Social Media* (pp. 85-94). ACM.

Manning, C., Surdeanu, M., Bauer, J., Finkel, J., Bethard, S., & McClosky, D. (2014). The Stanford CoreNLP natural language processing toolkit. In *Proceedings of 52nd annual meeting of the association for computational linguistics: system demonstrations* (pp. 55-60).

Ernala, S. K., Rizvi, A. F., Birnbaum, M. L., Kane, J. M., & De Choudhury, M. (2017). Linguistic markers indicating therapeutic outcomes of social media disclosures of schizophrenia. *Proceedings of the ACM on Human-Computer Interaction*, *1*(CSCW), 43.

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*(1), 547-577.

Boals, A., & Klein, K. (2005). Word use in emotional narratives about failed romantic relationships and subsequent mental health. *Journal of Language and Social Psychology*, *24*(3), 252-268.

Brubaker, J. R., Kivran-Swaine, F., Taber, L., & Hayes, G. R. (2012, May). Grief-stricken in a crowd: The language of bereavement and distress in social media. In *Sixth International AAAI Conference on Weblogs and Social Media*.

Eisenstein, J., Ahmed, A., & Xing, E. P. (2011). Sparse additive generative models of text.

Rekabsaz, N., Lupu, M., & Hanbury, A. (2017, April). Exploration of a threshold for similarity based on uncertainty in word embedding. In *European Conference on Information Retrieval* (pp. 396-409). Springer, Cham.

Creswell, J. D., Lam, S., Stanton, A. L., Taylor, S. E., Bower, J. E., & Sherman, D. K. (2007). Does self-affirmation, cognitive processing, or discovery of meaning explain cancer-related health benefits of expressive writing?. *Personality and Social Psychology Bulletin*, *33*(2), 238-250.

Hancock, J. T., Landrigan, C., & Silver, C. (2007, April). Expressing emotion in text-based communication. In *Proceedings of the SIGCHI conference on Human factors in computing systems* (pp. 929-932). ACM.

Saha, K., Weber, I., & De Choudhury, M. (2018, June). A Social Media Based Examination of the Effects of Counseling Recommendations After Student Deaths on College Campuses. In *Twelfth International AAAI Conference on Web and Social Media*.

Saha, K., & De Choudhury, M. (2017). Modeling stress with social media around incidents of gun violence on college campuses. *Proceedings of the ACM on Human-Computer Interaction*, *1*(CSCW), 92.