

# #NotAWhore! A Computational Linguistic Perspective of Rape Culture and Victimization on Social Media

Ashima Suvarna\*

Delhi Technological University  
Department of Computer Engineering  
asuvarna31@gmail.com

Grusha Bhalla\*

Delhi Technological University  
Department of Computer Engineering  
grushabhalla@gmail.com

## Abstract

The recent surge in online forums and movements supporting sexual assault survivors has led to the emergence of a ‘virtual bubble’ where survivors can recount their stories. However, this also makes the survivors vulnerable to bullying, trolling and victim blaming. Specifically, victim blaming has been shown to have acute psychological effects on the survivors and further discourage formal reporting of such crimes. Therefore, it is important to devise computationally relevant methods to identify and prevent victim blaming to protect the victims. In our work, we discuss the drastic effects of victim blaming through a short case study and then propose a single step transfer-learning based classification method to identify victim blaming language on Twitter. Finally, we compare the performance of our proposed model against various deep learning and machine learning models on a manually annotated domain-specific dataset.

## 1 Introduction

Global statistics indicate that 35% of women worldwide have experienced sexual violence at some point in their lives<sup>1</sup>. Popular hashtags like ‘#metoo’, ‘#sexualharassment’ on Twitter have encouraged victims to share their stories of sexual assault and formally report them. However, the backlash faced by the victims has been staggering. Victims of sexual assault are often held culpable for the assault, and are attacked on social forums by extremists. With the rise of such crimes, it is important to devise a computational framework that can identify and prevent online victimization of sexual assault survivors who choose to report the crime. Ambiguous interpretations of rape culture and victim blaming makes manually sorting and identifying such information an arduous task.

\*Authors contributed equally

<sup>1</sup><http://worldpopulationreview.com/countries/rape-statistics-by-country>

Hence, in our work we have attempted to identify an objectively grounded definition of victim blaming for further research in this domain.

Victim blaming occurs when the victim of a crime or any wrongful act is held entirely or partially at fault for the harm that befell them (Coates et al., 2006). Additionally, “slut shaming” is a popular form of victim blaming which refers to attacking a person’s character on the basis of sexual activity, real or perceived (Ringrose and Renold, 2012). Victims of sexual assault are initially hesitant to make a sexual assault complaint and often encounter victim blaming and slut shaming attitudes when they finally do (Ahrens, 2006). This blame can appear in the form of toxic social responses from medical professionals, the media, the judiciary or a growing majority of online activists on social media platforms (Campbell et al., 2009).

Social platforms like Twitter and Facebook provide victims with a ‘virtual bubble’ to recollect the assault stories and seek emotional help. The victim blaming faced by these victims, however, discourages them from disclosing their personal stories and further seeking medical help (Verdun-Jones and Rossiter, 2010). Therefore, in this work, we propose a method to identify such language on Twitter and protect the victims who choose to disclose their plight. We propose a Twitter-specific classification model which can exclusively identify victim blaming tweets. The key contributions of this work are:

- Our work is the first attempt in devising a computational framework for identifying victim blaming language.
- We provide a manually annotated dataset that contains 5,070 tweets for further research in this domain.
- We propose a single step transfer learning based classification method that identifies victim blaming language and labels it. It obtained

superior results to many deep learning and machine learning based approaches.

## 2 Related Work

Prior research has shown sexual assault is a crime that women are most afraid of (Koss, 1993). Oftentimes, victims of sexual assault are subjected to humiliation, blaming because of which, reasonable doubt is created about their credibility (Ullman, 2000). Popular theories such as the “just world” (Lerner, 1980) theory and the “invulnerability” (Andrew et al., 2003) theory explain the psychological motivation behind victim blaming. The “just world” theory states that people get what they deserve and deserve what they get, thereby shifting the blame of the crime to the victim while the “invulnerability” theory states that people blame the victim to project their own sense of invulnerability.

Victim Blaming, therefore, stems from an individual’s personal sense of insecurity and acts as a silencing function for most victims who are discouraged to disclose their personal stories or seek any help online (Ahrens, 2006). Since victims fail to obtain the required medical assistance, they become highly susceptible to emotional difficulties that manifest as depression in the short term and acute psychological difficulties in the long run (Verdun-Jones and Rossiter, 2010). In addition to this, chronosystem factors like past instances of victimization and sexual revictimization affect the mental health outcomes of the survivor often leading to suicidal behaviour, substance use, depression etc. (Campbell et al., 2009).

With the advent of social media, a new medium has presented itself for victim blaming to occur. Social platforms like Twitter, Facebook and Reddit provide a space to publicly post comments and present an insight into community opinions for researchers and social scientists. Due to its increasing popularity, Twitter is being used for research in opinion mining (Andleeb et al., 2019), keyword extraction (Biswas, 2019), hate speech detection (Badjatiya et al., 2017) etc. It has been widely used for research on sexual violence (Wekerle et al., 2018) as well as suicidal ideation using linear and ensemble classifiers (Sawhney et al., 2018). Research has also been focused on hate speech detection for Twitter using deep learning techniques, classifying tweets as sexist, racist or neither (Badjatiya et al., 2017). Balakrishnan et al. (2020) have used Naive Bayes, Random Forest and

J48 for detection of cyberbullying. Due to the similarities between victim blaming and cyberbullying, we have used Naive Bayes as one of our baseline models. Schrading (2015) analyzes discussions on domestic abuse across social media, using LSTM, Naive Bayes, Logistic Regression, and SVM which has been used as a baseline against our proposed method because of its good performance.

### 2.1 Motivation: Weibo Victim Scandal

*Liu Jingyao, a 21 years old student at the University of Minnesota accused Liu Qiangdong the founder of Chinas largest company JD.com, of raping her. She did not report her case immediately as she was afraid that she would be blamed. After the case became popular in China, people commented things like The woman looks disgusting, She is a slut etc. on Weibo, the Chinese equivalent of Twitter. She suffered from post traumatic stress disorder and insomnia because of this.*<sup>2</sup>

The Weibo Case study is a classic example of the drastic effects of victim blaming on the victim and it’s prevalence in our society. Reporting of instances of sexual violence has shown to pre-empt blame in the talk of women reporting blame which further shows that victim blaming itself is marked by specific topics and framing of sentences that shifts the blame onto the victim. Parameters like location of incident, state of victim etc. can be used for identifying such instances (Stubbs-Richardson et al., 2018). In lieu of these specific markers of victim blaming language which can further lead to biased reporting in offline media as well, we feel it is imperative to study about it in detail. Previous works in hate speech detection classified tweets as racist or sexist only (Badjatiya et al., 2017). These works generalize all instances of sexism under one classification. However, recent research has shown important sub-classifications of sexism that may be important for online media research. (Parikh et al., 2019) classifies tweets into 14 sub categories of sexism and we identify that not all 14 categories may have as drastic effects as victim blaming and slut shaming. Victim Blaming on online media directly leads to psychological disturbances for the victim and biased responses from authorities seeking legal action for such crimes (Gruenewald et al., 2004). This does not undermine the severity of other categories but rather establishes why victim blaming

<sup>2</sup><https://www.nytimes.com/2019/12/13/business/liu-jingyao-interview-richard-liu.html>

Clothing, makeup of victim (short skirt, skimpy clothes, v-neck shirts)
Victim's physiological state at time of incident (drunk, sad, depressed)
Victim's former/current job as a prostitute
Victim's sexual history or promiscuity
Victim's upbringing as explanation for behaviour (raised by lesbians, rich)
Locations that suggest victim culpability (bar, road, club)
Use of loaded terms to describe rape (alleges, accuses, she says)
self-reporting (Victim chooses to report crime on online media first)

Table 1: Coding Instrument to Identify Victim Blaming Language

should be studied separately and not as a specific instance of hate speech due to sensitivity and specific topics related to this issue. Models specific to each sub-category may not seem feasible and scalable but viewing at the issue of victim blaming from a psychological perspective we feel research in this domain is essential to devise computational frameworks to identify and prevent victim blaming on social media extensible to offline media reporting.

### 3 Dataset Construction

Creating our victim-blaming dataset entails a two step process: collection of data and data annotation. Due to the lack of prior work, we create a custom dataset by crawling English tweets from Twitter using the Twitter API<sup>3</sup> that mentioned major hashtags related to sexual harassment. A total of 4,242 tweets were scrapped from November 6, 2019 to November 19, 2019 which contained 'metoo. Also, 413 tweets were scrapped containing 'sexualharassment. We further observed that victim blaming tweets contained derogatory terms like 'whore' therefore, we used common words related to sexual harassment (rape, slut, whore) for increasing the number of positive samples. A total of 732 tweets were added after scrapping tweets from November 16, 2019 to November 29, 2019. Tweets are unstructured and noisy in nature due to the use of informal language prevalent on social media (Zappavigna, 2012). For pre-processing them, stop words are removed along with non alphanumeric text as they provide little information. URL's, emojis like ':' and the symbol '#' are removed using regex. Usernames are also removed for anonymity. The tweets are then tokenized to run the various models. Tokenization and removal of stop words is done using NLTK. The average

number of characters in a tweet is 33<sup>4</sup>. After pre-processing, we remove tweets which have less than 33 characters. The final dataset contains 5,070 tweets out of which 1562 were classified as positive samples. Under the guidance of a social scientist, a coding scheme to identify victim blaming language is devised, taking into account gender related discourse and campaigns as well as psychological analysis of rape culture and victimization from social media. Table 1 summarizes the key identifiers of victim-blaming language.

We follow a two-phase annotation process for the classification of each tweet in the final dataset. In Phase 1, two annotators, a psychology student and a social science student identify tweets which contain features listed in Table 1 and subsequently classify them by marking victim blaming tweets as '1' and other tweets as '0'. The inter annotator agreement, measured by the average of the Cohens Kappa (Cohen, 1960) is 0.712. In case of a disagreement between the annotators, a third annotator who is a social science student classifies the ambiguous tweet. Phase 2 involves checking that all the tweets have been classified and correct symbol that is, either '1' or '0' have been used for it. Table 2 shows some samples from the dataset that were labeled '1' by the annotators.

### 4 The Victim Blaming Language Model

NLP problems are solved using shallow machine learning models or deep learning models along with manually crafted features which fail to capture the linguistic information properly due to dimensionality problems (Khatua et al., 2018). Twitter follows a unique semantic and linguistic format (Hu et al., 2013) that makes it imperative to devise frameworks that incorporate the specific linguistic

<sup>3</sup><https://developer.twitter.com/>

<sup>4</sup><https://techcrunch.com/2018/10/30/twitters-doubling-of-character-count-from-140-to-280-had-little-impact-on-length-of-tweets/>

Tweets	Label
She went with them and got drunk. She slept with them and when they called her a whore accused them of rape	1
She is a dumb British whore who filed fake claim	1
Harvey Weinstein is a very evil man, but some of the celebrity women apart of the #Metoo movement that were “sexually assaulted” knew exactly what they were doing.	1
Another perverted Democratic! Let’s see if Hollywood will say something or if #MeToo is only for Republicans.	0
They definitely raped her. Why else would they call her a whore?	0
Women don’t admit they lost their power to men via rape for the fun of it.	0
If a woman says she was raped, believe her.	0

Table 2: Examples from Annotated Dataset(Paraphrased to preserve anonymity)

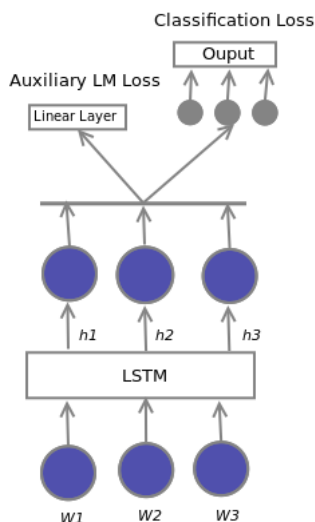


Figure 1: High Level Overview of Victim Blaming Classification Method. Blue indicates transferred layers and grey indicates randomly initialized layers

styles used on Twitter. Recent advancements in using transfer learning for tweet stance classification shows that enriching models with Twitter linguistics can improve performance (Zarrella and Marsh, 2016) Additionally, victim blaming language exhibits specific topics and syntax as shown in the coding instrument in Table 1. However, popular text classification models have failed to incorporate the subtle nuances of victim blaming language on social media specifically Twitter. Therefore, we propose a transfer learning based model that addresses this issue.

We propose a simple yet effective classification method based on single step transfer learning (Chronopoulou et al., 2019). State-of-the-art transfer learning methods employ language models (LM) trained on generic corpora with additional fine tuning of LMs for task specification. In our

method, we combine the task-specification and language modelling with the help of an auxiliary loss function that is adjusted during training for task and linguistic adaptation. This prevents catastrophic forgetting and allows our model to be trained on a social media specific large corpora for e.g.: Twitter or Reddit standard datasets and then be adapted to target tasks on domain specific smaller corpora. (Chronopoulou et al., 2019) theorizes that the prevention of catastrophic forgetting for machine translation tasks is because the language model objective acts as the regularizer that limits the loss of generalizable features and evidence for the same is presented in their work. We intuitively inferred that a similar approach would be effective in capturing the subtle topics of victim blaming on Twitter due to the additional language modelling step that guides the training across the text classification task.

**LM Pretraining:** We train a word-level LM which consists of an embedding LSTM layer, 2 hidden LSTM layers and a linear layer.

**Transfer learning and Auxiliary Loss:** We transfer the weights of the pre-trained model and add an additional LSTM layer.

We introduce an auxiliary LM loss during training to incorporate the contribution of the pre-trained language model in the classification method. The joint loss is the sum of classification loss,  $L_{CLF}$  and auxiliary LM loss,  $L_{LM}$ .

$$L = L_{CLF} + L_{LM}$$

We consider equal contribution of both the loss values to effectively capture language modelling information and classification information specific



Parameter	Value
Activation function	ReLU
Dropout	0.4
Batch size	64
Epochs	20
Optimizer	Adam
Learning rate	0.0005

Table 3: Parameters for CNN Architectures

to the nature of our dataset. A High level overview of our method is shown in Figure 1.

## 5 Experiments

### 5.1 Baselines

#### Traditional Machine Learning (ML) Approaches

We have used two machine learning models, Support Vector Machine (SVM) and Naive Bayes (NB).

**SVM:** For feature extraction, TF-IDF has been used on word unigrams that is fed to the SVM

**NB:** Similar to SVM, TF-IDF has been used for feature extraction for classification

#### LSTM-Based Architectures

**LSTM:** The word embeddings for all the words in a post are fed to a vanilla LSTM

**TextbiRNN:** This is an improvement on a vanilla RNN. The word embeddings for all the words in a post are fed to a bi-directional LSTM.

#### CNN-Based Architectures

**TextCNN:** Convolutional filters we applied to the word vectors of a post followed by max-pooling layers as described by (Kim, 2014)

**CharCNN:** A sequence of encoded characters are fed into a CNN as described by (Zhang et al., 2015)

**fasttext:** fasttext classifier is used for text classification which takes into account n-grams of words to incorporate local word order (Grave et al., 2016)

### 5.2 Implementation of Victim-Blaming Classification Method

To pretrain the language model we create a dataset<sup>5</sup> of 1 million English tweets scraped from Twitter, including approximately 1M unique tokens. We use 50K most frequent tokens as our vocabulary. We then use our Victim-Blaming dataset for classification. To pre-process the tweets we use regex

<sup>5</sup><https://www.kaggle.com/paoloripamonti/twitter-sentiment-analysis>

to remove usernames, urls and emojis. In addition to this, we use NLTK for stop word removal and tokenization of the tweets. For neural models, we use an LM with embedding size of 300, 2 hidden layers, dropout of 0.3 and batch size of 64. We add an LSTM of size 100 with a softmax classification layer on top of the transferred LM. In pretraining, and pretrained layers (of transferred model), Adam was used with a learning rate of 0.0001. For the newly added LSTM and classification layers, Adam with learning rate of 0.0005 was used. For developing our models, we use Pytorch and Sci-kit learn.

### 5.3 Results

Table 4 describes the performance of the baseline models in comparison to our proposed approach across the accuracy metric. The models were trained over 60% of the dataset while 20% was held out for test and 20% was used as dev split to optimize the parameters across all the models tested. The proposed approach outperforms all baselines including RNNs, CNNs, LSTMs and traditional ML approaches SVM and NB. Fasttext model is able to generate domain specific embeddings due to the nature of embedding construction that benefits the unpredictable and unstructured Twitter semantics. CharCNNs usually have a high perplexity due to the character-by-character prediction, however, they presented similar results to the fasttext model which are better than the other baseline models. Our method shows better results when compared with the baseline models. Since we do not have a lot of data, the baseline models fail to identify linguistic features of twitter language which are significantly different from normal conversational language. Our method takes care of this using the auxiliary loss function. The language model trained on generic corpora is successfully able to capture these features and can therefore perform better when retrained for classification. In comparison to our baselines, our model architecture is simpler and computationally inexpensive.

### 5.4 Error Analysis

It has been observed that sometimes incidents of victim blaming are either self reported or reported by a third person. Some tweets may cite previous instances of victim blaming to speak against victim blaming. Since, these tweets consist of the marked topics and linguistic framing encoded in the coding instrument in Table 1, the proposed model classifies

Approach	Accuracy
NB	0.60
SVM	0.73
LSTM	0.74
TextbiRNN	0.74
TextCNN	0.75
CharCNN	0.77
fasttext	0.78
Proposed Method	<b>0.82</b>

Table 4: Performance Comparisons on Victim Blaming Dataset

such tweets as positive examples. This is a typical form of error encountered in even hate speech detection tasks (MacAvaney et al., 2019) where keywords marked for positive examples leads to classification errors. It should further be discussed whether such examples should be classified as positive or negative during the annotation process and requires extensive social and psychological research. Some systemic errors we explored during our experiments:

- *She may tweet against you in #MeToo if you are not careful*: This tweet was annotated as 1, that is, victim blaming tweet since it contains implicit victim blaming. The proposed model wrongly classifies it as 0 as it lacks specific topics and keywords that the model has learned during training. This error arises due to the model failing to effectively capture sarcasm.
- *All the desi feminists....using woman card for personal gains and abusing #Metoo*: This tweet is not classified as victim blaming by the model, however, it is annotated as victim blaming due to implicit gaslighting of victims who choose to report it. It was mis-classified since it is not directly threatening or accusing a victim. The researchers also feel this might be an oversight between the annotators while data annotation as this may be a case of general sexism and not victim blaming. To address this type of error we plan to extend our work into a multi-label categorization task which considers sub-categories of victimization, that is, secondary, primary and gender based victimization in rape cases.
- *Even if I agree, most of what it would take for that to be a valid viewpoint, you still mean*

*that "you can't rape a whore." Justice should be principle based not tribe based.*: This tweet is not victim blaming but citing an instance of victim blaming directly which leads the model to classify it as victim. This is an instance which is very common in hate speech tasks as well where citing or using such phrase and words to talk against the hate or victim blaming language leads to false positives during classification.

- *I'll be a good boy and take it silently if you rape my cunt*: This tweet contains vulgar language that is identified by the model as victim blaming erroneously. This tweet is annotated as 0, not victim blaming but the specific words like 'cunt' or 'take it' has clearly confused the model as it is failing to capture long sequences here and decipher the meaning of the tweet wholly.

## 6 Conclusion and Future Work

In this work, we established the need to devise a computationally effective method to identify victim blaming language on Twitter. To achieve this, we proposed a single step transfer learning based classification method that effectively captures the unique linguistic structures of twitter data and victim blaming language. On a manually annotated dataset, our proposed approach could achieve significant improvement over existing methods that rely on custom textual features and popular deep learning based methods. The prevalence of rape culture and the subsequent victim blaming on unsolicited social media forums like Twitter has not been studied from a computational linguistic perspective before. Our work, therefore presents an extensive study of popular text classification methods on a niche' dataset with victim blaming semantics and further presents the significance of using a simple transfer learning approach to capture Twitter semantics on a limited dataset. We anticipate that this study encourages further research on how victims of sexual assault are portrayed on social media. Our future agenda includes further bifurcating and exploring the specific types of victim blaming and the efficacy of the proposed approach on such a multi label classification task. We plan to explore the different weighting factors for the language modelling loss and classification loss described in section 4 to determine if weighting factors can help customize the auxiliary loss for different tasks.

## References

- Courtney E. Ahrens. 2006. [Being silenced: The impact of negative social reactions on the disclosure of rape](#). In *American Journal of Community Psychology*, pages 23(2), 31–34.
- Aslam Andleeb, Qamar Usman, Ayesha Khan Reda, Pakizah Saqib, Aleema Aleena Ahmad, and Aiman Qadeer. 2019. Opinion mining using live twitter data. In *IEEE International Conference on Computational Science and Engineering (CSE) and IEEE International Conference on Embedded and Ubiquitous Computing (EUC)*.
- Bernice Andrew, Chris R. Brewin, and Suzanna Rose. 2003. Gender, social support, and ptsd in victims of violent crime. In *Journal of Traumatic Stress*, pages (4), 16, 421–427.
- Pinkesh Badjatiya, Shashank Gupta, Manish Gupta, and Vasudeva Varma. 2017. Deep learning for hate speech detection in tweets. In *Proceedings of the 26th International Conference on World Wide Web Companion*, page 759760.
- Vimala Balakrishnan, Shahzaib Khan, and Hamid R. Arabia. 2020. Improving cyberbullying detection using twitter users psychological features and machine learning. In *Computers Security*, page Volume 90.
- Saroj Kumar Biswas. 2019. Keyword extraction from tweets using weighted graph. In *Advances in Intelligent Systems and Computing*.
- Rebecca Campbell, Emily Dworkin, and Giannina Cabral. 2009. [An ecological model of the impact of sexual assault on women’s mental health](#). In *Trauma Violence Abuse*, pages 10(3), 225–46.
- Alexandra Chronopoulou, Christos Baziotis, and Alexandros Potamianos. 2019. [An embarrassingly simple approach for transfer learning from pre-trained language models](#). In *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers)*, pages 2089–2095, Minneapolis, Minnesota. Association for Computational Linguistics.
- Linda Coates, Cathy Richardson, and Allan Wade. 2006. Reshaping responses to victims of violent crime. In *Cowichan Bay, B.C., Canada*.
- Jacob Cohen. 1960. A coefficient of agreement for nominal scales. *Educational and psychological measurement*, 20(1):37–46.
- Edouard Grave, Tomas Mikolov, Armand Joulin, and Piotr Bojanowski. 2016. Bag of tricks for efficient text classification. In *EACL*.
- Tara L Gruenewald, Margaret E Kemeny, Najib Aziz, and John L Fahey. 2004. Acute threat to the social self: Shame, social self-esteem, and cortisol activity. *Psychosomatic medicine*, 66(6):915–924.
- Y. Hu, K. Talamadupula, and S. Kambhampati. 2013. [Dude, srsly?: The surprisingly formal nature of twitter’s language](#). *Proceedings of the 7th International Conference on Weblogs and Social Media, ICWSM 2013*, pages 244–253.
- Aparup Khatua, Erik Cambria, and Apalak Khatua. 2018. [Sounds of silence breakers: Exploring sexual violence on twitter](#). In *2018 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM)*, pages 397–400.
- Yoon Kim. 2014. [Convolutional neural networks for sentence classification](#). In *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pages 1746–1751, Doha, Qatar. Association for Computational Linguistics.
- M. P Koss. 1993. [Rape: Scope, impact, interventions, and public policy responses](#). In *American Psychologist*, pages 48(10), 10621069. American Psychological Association.
- Melvin J. Lerner. 1980. The belief in a just world: A fundamental delusion.
- Sean MacAvaney, Hao-Ren Yao, Eugene Yang, Katina Russell, Nazli Goharian, and Ophir Frieder. 2019. Hate speech detection: Challenges and solutions. *PloS one*, 14(8).
- Pulkrit Parikh, Harika Abburi, Pinkesh Badjatiya, Radhika Krishnan, Niyati Chhaya, Manish Gupta, and Vasudeva Varma. 2019. Multi-label categorization of accounts of sexism using a neural framework. *arXiv preprint arXiv:1910.04602*.
- Jessica Ringrose and Emma Renold. 2012. [Slut-shaming, girl power and sexualisation: Thinking through the politics of the international slutwalks with teen girls](#). In *Gender and Education*.
- Ramit Sawhney, Prachi Manchanda, Raj Singh, and Swati Aggarwal. 2018. [A computational approach to feature extraction for identification of suicidal ideation in tweets](#). In *Proceedings of ACL 2018, Student Research Workshop*, pages 91–98, Melbourne, Australia. Association for Computational Linguistics.
- Nicolas J. Schrading. 2015. Analyzing domestic abuse using natural language processing on social media data.
- Megan Stubbs-Richardson, Nicole E Rader, and Arthur G Cosby. 2018. Tweeting rape culture: Examining portrayals of victim blaming in discussions of sexual assault cases on twitter. *Feminism & Psychology*, 28(1):90–108.
- S.E Ullman. 2000. Psychometric characteristics of the social reactions questionnaire. In *Psychology of Women Quarterly*.

- Simon N. Verdun-Jones and Katherine Rossiter. 2010. The psychological impact of victimization: Mental health outcomes and psychological, legal, and restorative interventions.
- Christine Wekerle, Negar Vakili, Sherry Stewart, and Tara Black. 2018. The utility of twitter as a tool for increasing reach of research on sexual violence. In *Child abuse neglect*, page 85.
- Michele Zappavigna. 2012. *Discourse of Twitter and social media: How we use language to create affiliation on the web*, volume 6. A&C Black.
- Guido Zarrella and Amy Marsh. 2016. Mitre at semeval-2016 task 6: Transfer learning for stance detection. *arXiv preprint arXiv:1606.03784*.
- Xiang Zhang, Junbo Zhao, and Yann LeCun. 2015. [Character-level convolutional networks for text classification](#). In C. Cortes, N. D. Lawrence, D. D. Lee, M. Sugiyama, and R. Garnett, editors, *Advances in Neural Information Processing Systems 28*, pages 649–657. Curran Associates, Inc.