Discovering Intimate Partner Violence from Web Search History

Anis Zaman[a], Henry Kautz[a], Vincent Silenzio[c], Md. Ehsan Hoque[a], Corey Nichols[b], Catherine Cerulli[d]

[a] *Computer Science Department, University of Rochester, United States*
[b] *School of Medicine and Dentistry, University of Rochester, United States*
[c] *Department of Urban-Global Public Health, Rutgers University, United States*
[d] *Department of Psychiatry, University of Rochester Medical Center United States*

Abstract

Intimate partner violence is a public health problem with increasing prevalence and harmful influence to both individuals and society. Automated screening for intimate partner violence is still an unsolved problem in academic research and practical applications. Current detection methods use self-reporting scales and in-person interviews, which are laborious, expensive, and often lack precision and sensitivity, making it essential to develop new approaches. This paper proposes a scalable and lightweight ubiquitous screening technique, validated via ground truth data collected through self assessment survey, for detecting signs of intimate partner violence by analyzing individual-level Google search histories. Initial analysis shows that there are temporal, textual, contextual differences in search behavior between individuals who have/haven't experienced intimate partner violence. Using these differentiating signals, we were able to build a model that can detect violence in intimate relationships with an F1 score of 0.80. Though preliminary, we hope our findings pave the way for the AI community to address this important public health problem.

Keywords: Public health, Intimate partner violence, Online search history, Ubiquitous sensing from search engine logs

1. Introduction

Intimate partner violence (IPV) is a major public health problem with serious consequences for victims' physical and mental health. There have been numerous research studies focused on the prevalence of IPV [1, 2, 3, 4, 5, 6, 7, 8] and its influence on mental health [9, 10, 11, 12, 13, 14]. A study in 2002 showed that 10% to 52% of women from all over the world had experienced physical violence by an intimate partner, and 10% to 30% of them had been sexually abused by an intimate partner at some point in their lives [15]. According to a WHO study in 2006, this number is rising, with up to 71% of ever-partnered women have reported experiencing IPV [16]. Every minute nearly 20 people are physically abused by an intimate partner in the United States and both men (1 in 7) and women (1 in 4) experience severe physical violence by an intimate partner during their lifetime [17].

Despite such high prevalence of IPV, the development of screening tools is difficult due to the unclear definition of IPV and how it is recognized in different cultures. Most frequently, married heterosexual women are studied primarily for signs of physical violence [18]. Violence can be physical, emotional, and/or sexual, and individuals with different gender and sexual orientation can be victims or perpetrators. For instance, prior research demonstrated that both men and women perpetrate IPV at near-equal rates [19, 20], and that violence is prevalent in gay male and lesbian relationships with rates comparable to heterosexual relationships [21, 22]. In addition, the psychological abuse aspect of violence is often overlooked in different cultures when interpreting IPV. Furthermore, due to the stigma of being labeled as "victims" or "perpetrators", people tend to hide information about their situation during interviews or self-report scales [23]. For example, Latina women are more likely to experience IPV for longer period of time and due to the cultural pressure to remain in a violent relationship, and may choose to avoid exposing their IPV situation [24].

Email addresses: azaman2@cs.rochester.edu (Anis Zaman), kautz@cs.rochester.edu (Henry Kautz), vincent.silenzio@rutgers.edu (Vincent Silenzio), mehoque@cs.rochester.edu (Md. Ehsan Hoque), corey_nichols@urmc.rochester.edu (Corey Nichols), corey_nichols@urmc.rochester.edu (Catherine Cerulli)

In addition, victims of IPV can suffer from various chronic conditions and acute complications, such as paranoia, stress, recurring pain, post-traumatic stress disorder (PTSD), depression, alcohol and drug abuse, suicide ideation, committing murder, and being the victim of stalking [7, 5, 3, 25, 26, 27, 28, 29].

Designing interventions to limit each of the adverse effects of IPV will not be sufficient to ameliorate the situation. Instead, more emphasis is needed on identifying and addressing early stage IPV situations. Possible venues in which to implement such detection tools include healthcare settings, counselling services, social service agencies, and legal and criminal justice services.

In this paper we investigate observable textual and contextual signatures related to IPV conditions using both individual level Google search histories and answers to a validated IPV assessment survey. For search queries, as opposed to other uses of digital media, users are less likely to self-select or face the same temptations to self-censor, and there is no fear of the social stigma associated with expressing IPV. The nature and volume of search queries also provide temporal behavioral information about a user. Prior researchers have analyzed search logs of large aggregate populations using Google Trends [30, 31]; however no prior work explored the prevalence of signs of IPV in search history.

In summary, given the shortcomings of current IPV detection instruments and approaches, we propose a new screening method based on individual level search history. Our contributions are: (i) IPV has rarely been addressed in the AI literature; (ii) introducing a novel approach for detecting IPV and relationship distress using web search history instead of traditional self-rating scales or interviews; (iii) in addition to maintaining user anonymity and privacy, to the best of our knowledge, we are the first to explore such a screening tool; (iv) the proposed method is scalable, lightweight and can be deployed at any location to use immediately available web search history data to provide instantly actionable interventions.

2. Related Work

IPV is a complex phenomenon related to an array of contributing factors associated with cultural and social norms, financial and interpersonal stressors, environmental experiences, etc. [32]. In addition to the physical injury and illness that can result from IPV, it is also commonly associated with various mental health outcomes [33].

A data source that is quite popular among researchers is search engine logs from Google Trends [34]. However, data from Google Trends is population level and researchers have leveraged this to build predictive models for accessing various mental and physical health conditions. For example, early diagnosis of lung carcinoma [35], detecting neurodegenerative disorders [36], surveillance of influenza outbreaks [37, 38]; monitoring toxicological outbreaks [39]; identifying seasonality in seeking mental health services [30]; and monitoring suicide related terms to help predict rates of self-injury and death within the population [40]. Unlike the previous works, [41] used individual level search history to detect incidence of low self-esteem among college students. [42] showed that mouse tracking on search engine can detect users at early states of Parkinson's disease.

Another type of data most closely related to the data used in our study is social media data. Social media platforms have become an avenue for health personnel to reach out to populations at risk to mitigate the crisis situation [43]. It is important to note that digital technologies such as online platforms occupy an important position in the IPV ecosystem [44]. In particular, prior work shows that abusers use technology to impact survivors even after separation [45, 46]. To address this, researchers have proposed various privacy and security preserving technologies for IPV survivors [47]. Fahdah et al. presented a comprehensive review of how digital platforms can be used to collect data for tackling public health related problems [48]. To better understand IPV, researchers have conducted studies where they used social media platforms, such as Facebook, to recruit participants [49]. Watkins et al. recruited participants through online platforms and have proposed scales to measure aggression and violence in intimate partner relationships [50]. Despite these efforts, there are significant challenges in trying to detect abuse on digital platforms. In particular, recognizing abuse on popular social platforms is often subtle and may only be noticeable to the victim [46]. Rabin et al. provided comprehensive review of the existing IPV screening techniques [51]. Existing screening methodologies include computer-assisted self-administered, self-administered written method, and face-to-face interview sessions designed for people from different background [52, 53, 54, 55]. All of the above screening methods are laborious, slow, require human supervision, and come with the fear of being stigmatized. None of these studies considered interactions in digital platforms as a way to screen potential IPV.

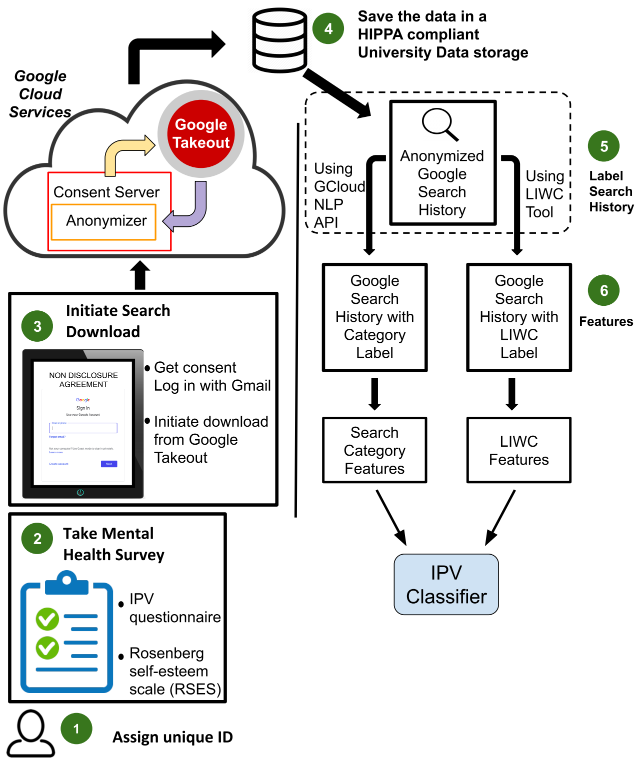


Figure 1: Pipeline for the entire IPV detection framework, from data collection to modeling.

3. Data

We adapted the data collection process for this work from [41]. This study was reviewed and approved by our Institutional Review Board. Data were collected during a 6-month recruitment period from three different locations, namely a college campus, an out-patient behavioral clinic of a reputed hospital, and a local Family Court[1] to investigate how online services such as Google's search engine can be used to detect signs of IPV. As described below, we developed a Human Subjects Review Board-approved, cloud-based application to document the subject's informed consent to participate in the study, download and anonymize their search histories, and link anonymized histories to corresponding survey responses.

3.1. Recruitment Procedure

Participation was voluntary and the wording of our study advertisements and consent materials were very generic, along the lines of "help us learn about health using online data", without specific mention of IPV. This was done to limit participation bias, although we acknowledge that it is virtually impossible to completely eliminate it.

In total, 96 individuals expressed interest in taking part in the study. However, only 56 qualified to take part, representing a 58% participation rate, because one had to be at least 18 years old, able to read and write in English, and had an active Google services account. This rate is consistent with our research team's prior experiences of gathering such proprietary, sensitive data, and the fact that 42% either declined or did not qualify to participate was not unexpected. The most common reason given for declining to participate for recruitment in Family Court was anxiety about sharing personal information. By contrast, few potential subjects at the clinic recruitment site expressed such anxiety, so the most common reason for non-participation was lack of a Google account (i.e. Gmail account).

[1] A place where hearings related to family law take place. In order to capture as diverse population as possible, we opted to select for these three locations. For example, divorce, orders of protection, guardianship, child custody, etc.

Demographic Variable		Value
	Age	35.15(mean) ($\sigma = 9.13$)
Gender	*Male*	49% (24)
	Female	51% (25)
Sexual Orientation	*Heterosexual*	66.1% (37)
	Bisexual	17.9% (10)
	Homosexual	12.5% (7)
	Other	3.6% (2)
Race	*American Indian*	1.8% (1)
	White	55.4% (31)
	Black or African American	28.6% (16)
	Other	14.3% (8)
Marital Status	*Single*	53.6% (30)
	Married	32.1% (18)
	Divorced	14.3% (8)
Recruitment Sites	*Family Court*	41% (23)
	Clinic	36% (20)
	On-campus	23% (13)

Table 1: Participants demographics ($N = 56$)

All of the respondents to the advertisement for the campus study were willing and able to participate. Qualified participants were compensated with $10 for taking part in the study.

3.2. Google Search History Download Process

We leveraged the data collection pipeline pioneered by [41]. There are two parts to the data collection process (see Fig. 1). First, a study recruiter asked participants to enter their Google login credentials into a web-based app running on a tablet. Second, study participants were asked to take the Promote Health survey, a brief self assessment online questionnaire developed and used in previous studies by our research team. Table 1 presents a summary of the participants in the pilot study.

Recently, Google launched a platform called Google Takeout[2] that enables anyone with Google account to access his or her entire personal search history data at anytime, for free. To make life easy, Google consolidates the entire online search history from different devices (mobile & desktop) and browsers and make it available through Google Takeout. Using the Google Oauth authentication protocol, participants authorized our search download application to access their search history data. Once a user authorized the application to obtain a copy of the search history, a one-time log in authorization was used to automatically transfer the search data from the user's Google drive to a secured server in the cloud.

Before any data were transferred from the cloud and given to the research team to store as part of this study, the data were fully de-identified via Google's Cloud Data Loss Prevention (DLP) API to assure removal of personally identifying information. The API scans for personal identifiers such as names, addresses, phone numbers, etc., that can potentially identify individual study participants, and this information was automatically removed from the search history data prior to being transferred to the research team and saved as research data.

During the survey process, participants did not provide their names, but instead entered a random ID number (specific to each user) that was used to link their responses to their search history data. To assure confidentiality, no records connecting subject names to ID numbers were compiled or maintained. Once the search history data was

[2] `https://takeout.google.com/settings/takeout`

successfully de-identified and transferred to the research team, the one-time log in authorization credentials were permanently deleted from the tablet device and our servers. The research team was then no longer able to use this authorization token to access any of the participant's online data again. Fig. 1 presents the entire data collection process.

3.3. The Promote Health Survey

Promote Health is a computer assisted survey instrument (CASI), designed at a fifth grade reading level. It was used to collect ground truth data about the participants. In addition to collecting age, gender, mental and sexual health measures, and demographic information, it includes the following three scales to score the presence of IPV: (i) the George Washington University Universal Violence Prevention Screening Protocol (UVPSP): a screening tool for physical violence, threat of violence, sexual violence, and emotional violence in the past year. A positive response to any question yields a positive screen for IPV [56]; (ii) the Computer Based IPV Questionnaire (CIPVQ), list of four questions used to recognize if the subject is the IPV perpetrator in the relationship; and (iii) the Danger Assessment (DA) questionnaire, used to detect if an individual was threatened by intimate partner in the form of physical, psychological and sexual over the past year, and measures the danger level of the violence [57].

In addition, IPV has many mental health consequences including depression, anxiety, substance abuse, low self-esteem, etc. [58, 59, 60, 61, 62]. Tackling all of these mental health conditions and their association to IPV is beyond the scope of this paper. However, prior research showed that there is high prevalence of low self-esteem among IPV victims [63, 64] which may lead to a vicious cycle where victims with low self-esteem are less likely to take steps to come out of an abusive relationship [65, 66]. This motivated us to ask questions for accessing self-esteem of an individual via the Rosenberg self-esteem (ROS) [67] scale. The 30 point ROS scale is a commonly used tool to quantify levels of self-esteem among young adults [68]. A ROS score of below 15 suggests prevalence of low self-esteem. Although we did not use the self-esteem information for building our IPV model, we noticed that the distribution of ROS score varied significantly between IPV and NIPV individuals.

3.4. Identifying IPV

Subjects were defined as having experienced IPV based on their responses to a series of questions adapted from the UVPSP, CIPVQ, and DA questionnaires. In order to capture temporal context, we prefixed each of these questions with the following text: *Over the last 15 months, how often have you experienced any of the following?* In total there were 30 questions and the number of "Yes" answer to these questions was used to measure the extent of the presence of IPV. For the scope of this work, single "Yes" answer to any of the questions is used to label the participant as having experienced IPV. Greater the number of "Yes" higher the extent of IPV, however we did not explore the degree of IPV in the work. Based on these criteria, 24 out of 56 participants were found to have some form of violence in their relationships, and the remaining 32 subjects had no sign of IPV (NIPV).

3.5. Data Breakdown

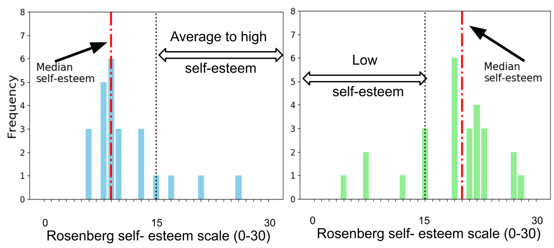
Before any analysis and modeling, we separate the dataset into two parts, namely analysis set and test set. In order to align with the survey questions timeline, we considered only the most recent 15 months of online data for this study. For the IPV detection task, the analysis set consisted of data from 34 participants while there were 22 subjects in the test set. This is illustrated in Table 2.

4. Observation

By analyzing the Promote Health survey responses, in particular the the Rosenberg self-esteem scale, we observed that the overall distribution of self-esteem score was distinctly different between individuals with and without IPV. The affect of abusive life experiences and relationships on one's self-esteem has been well studied by the community [69, 70, 71]. Based on the survey responses, we found that the median self-esteem score of individuals with IPV was 9 where as the median score was 20 for NIPV individuals, see Fig.2. One possible explanation for the high prevalence of low self-esteem among the IPV individuals can be that their emotions and judgements are predominantly engulfed by feelings of self-doubt and fear which may have resulted from years of abusive relationship.

N = 56	Analysis or Train Set	Test Set
Intimate Partner Violence (IPV)	16	8
No Intimate Partner Violence (NIPV)	18	14
Total	34	22

Table 2: Search dataset breakdown. Prior to any analysis, we separated 60% of the data for training and roughly 30% as the test set. Given the small dataset, we allocated 2/3 of IPV participants for training and 1/3 for testing the models.



(a) IPV (b) NIPV

Figure 2: Comparison of self-esteem score between victims of IPV and NIPV. Score < 15 implies evidence of low self-esteem (a) Distribution of self-esteem score among victims of IPV. Their median self-esteem score is 9 which is less than the threshold score of 15. (b) Distribution of self-esteem score among the NIPV individuals, their median self-esteem score is 20 which is within the normal range (15 to 25 is normal, > 25 represents high self-esteem.)

Since Google search phrases are short and recorded in the moment they are good indicators of what may be going through an individual's mind at an instant of time. We identify distinctive search characteristics and behaviors between IPV and NIPV participants from the perspective of 1) linguistic attributes of the search queries, 2) search times, and 3) the categories and content of search queries.

4.1. Differentiating linguistic attributes in search logs

Previous researchers have performed psycholinguistic analysis on text data to uncover signals of abuse (emotional, sexual, or physical) and domestic violence using the Linguistic Inquiry and Word Count (LIWC) toolkit [72, 73, 74]. LIWC is a text analysis software package that outputs the proportion of words in a given text that fall into one or more of over 80 linguistic categories (such as prepositions, adverbs, first-person singular pronouns, and conjunctions), psychological (happy, anger, achievement, etc.), and topical (leisure, money, etc.) [75].

The search phrases for each IPV subject from the analysis set were analyzed using LIWC, and the output for each subject was a 92-dimension vector. The result can be treated as a table with 16 rows and 92 columns. A similar procedure was repeated for all participants with NIPV condition, resulting in a table with 18 rows and 92 columns.

Next, we explored whether there were any statistically significant LIWC categories between the two populations. We analyzed the LIWC variables from the IPV and NIPV groups using the Mann–Whitney U test [76]. We found 12 LIWC attributes that were statistically significantly different (with p-value < 0.05) between participants with IPV and NIPV, as shown in Fig. 3.

During our analysis we found that individuals with IPV had elevated usage of words related to linguistic variables that capture emotions such as sadness (*sad*) and cognitive processing (*cognitive processes*). This supports previously established findings, where researchers have shown that LIWC attributes scores related to cognition(psychological process) and emotions were higher (in terms of how they express themselves) for people exposed to violence and abuse [73]. Prevalence of pronouns in text can provide interesting insights about one's personality and thought patterns [77]. On analyzing the search histories, we observed that the *pronoun* and *i* category scores were statistically

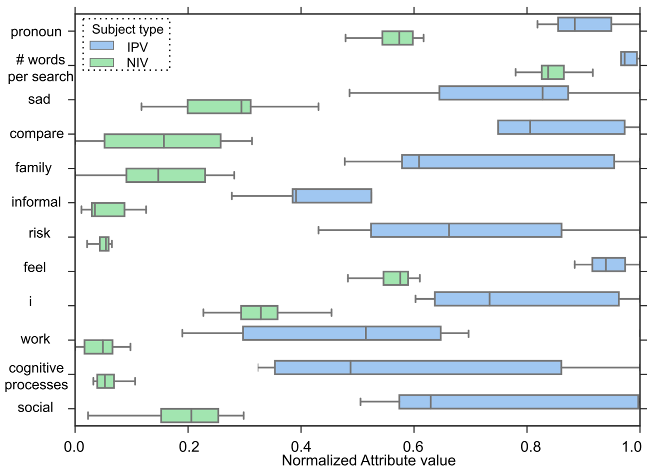


Figure 3: LIWC attributes that were statistically significantly different(Mann–Whitney U test, $p < 0.05$) between the subjects with IPV and NIPV condition. All of these attributes score were higher among subjects with IPV condition (despite fewer number of IPV in the dataset).

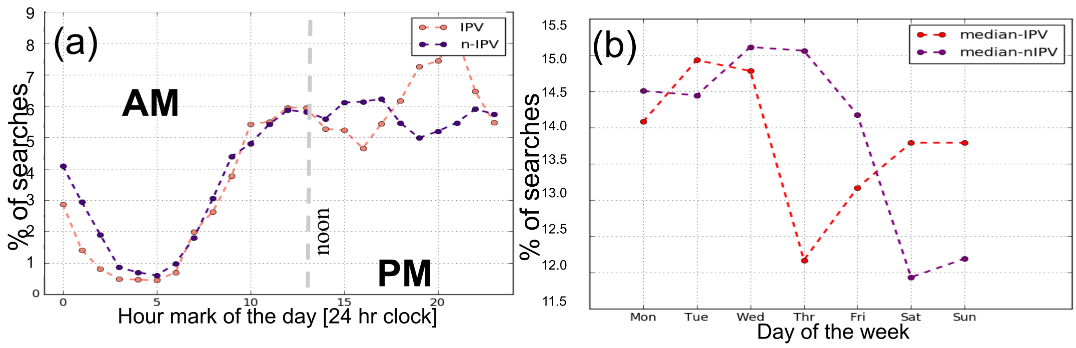


Figure 4: Comparison of searching frequency between individuals with IPV and NIPV. (a) Distribution of Google searches over 24 hour period. (b) Weekly distribution of Google searches.

significantly higher for victims of IPV than NIPV. Prior research showed that increased use of first-person singular pronouns is associated with depression, a condition that IPV victims are likely to be living with [78].

Unlike other popular data sources such as social media, blogs, and interviews that are commonly analyzed using LIWC toolkit, search history data have the major advantage in that when someone searches, there is no fear of any judgment and stigma, and that what he/she is typing reflects his or her state of mind at that moment.

4.2. Differences in search activity times

Since every search query is time tagged in the data provided by Takeout, were able to explore whether potential differences might exist between IPV and NIPV subjects. We observed that individuals with IPV were more active (i.e. made more Google searches) on the internet on certain days of the week compared to individuals with NIPV. For instance, during weekends, individuals with IPV search more than individuals with NIPV, as shown in Fig. 4(b). We also observed that between 2pm and midnight, the distribution of search volume was reversed between individuals with and without IPV, as shown in Fig. 4(a). From 2pm to 5pm, the volume of Google searches is higher for the NIPV population, and from 5pm to midnight it decreases and then gradually increases. Among IPV participants however, this searching behavior was reversed compared to those with NIPV. In addition, the overall number of search queries among NIPV participants peaked during the beginning through midweek, and declined on weekend days. Among participants with IPV, however, we found a sharp decline after the midweek, and a brisk resurgence in search activity over the weekend.

Search Category	p value (following Mann-Whitney U test)
Home & Gardening	0.035
Health	0.026
Online Communities	0.002
Sensitive Subjects	0.030
Science	0.041
Adult	0.039
Real Estate	0.017
References	0.028

Table 3: Search Categories that were statistically significantly different IPV and NIPV subjects. Mann-Whitney U test was done to compute the respective p-values

This pattern represents an interesting finding. Should additional research confirm the validity of the finding, it raises intriguing questions as to potential explanations and interpretations. For example, it could be proposed that individuals who experience IPV are not able to engage in online activities at the same times as NIPV individuals. Regardless of the underlying mechanism that might produce such a pattern, if further research confirms this preliminary finding, it would offer a potentially important signal in models for IPV classification.

4.3. Differences in search categories

Using the content classification feature of the Google Cloud NLP API[3] we were able to classify the search queries of participating subjects. The API returns a hierarchical category label for every search query. We treat the broad category within the hierarchy as the 'category' label for the query. For instance, for a query q, if the label from the API is "Arts & Entertainment/Humor/Funny Pictures & Videos", we consider "Arts & Entertainment" as the category for q. A comprehensive list of all the categories can be found here[4].

4.4. Rationale for using the Google NLP API

We did not adapt data-driven topic modeling to label the search logs because we want an *unified* set of topics that can be applied to consistently label and compare searches from all individuals. Consider two searches, "how to garnish salad" (s1) and "Texas BBQ recipe" (s2). Google API will classify both as *Food* whereas data-driven topic modeling may put these in two different clusters.

In total, there were 27 such broad categories. We observed that, at any time of the day, individuals exposed to IPV condition had more searches relating to Home & Gardening, Arts & Entertainment, Hobbies, and Electronic Gadgets, and References but fewer searches related to Finance, Adult, and Community/Societal Matters. Table 3 presents the categories that were statistically significantly different (Mann–Whitney U test with $p < 0.05$) between the IPV and NIPV population.

We found that victims of IPV more frequently search about topics related to References, Home & Gardening, Real Estate, Electronics, etc. From the survey responses we found that individuals with IPV condition had low self-esteem (median score < 15) when compared with someone with NIPV. It is arguable, but one plausible explanation can be that individuals with low self-esteem constantly feel insecure about things around them and are more reliant on search engines to seek reassurance about everyday things such school, office works, etc. We think that this may be, at least in part, a reflection of individual search acts as a form of validation. Prior work has shown that individuals with high self-esteem do not need continual validation [79].

On the contrary, our analysis showed that victims of IPV had fewer number of searches related to Adult, Finances, Community and Social matters. A potential explanation for fewer number of Adult related searches among the victims of IPV can be that partners may feel drained out because of the never ending conflicts in the relationship. Eventually, persistence of this situation for longer period of time may make them lose interest in their respective significant others

[3] `https://cloud.google.com/natural-language/docs/classifying-textt`
[4] `https://cloud.google.com/natural-language/docs/categories`

(a) Using LIWC attributes as features

	precision	recall	f1-score
NIPV	0.71 (LR) 0.79 (SVM)	0.71 (LR) 0.79 (SVM)	0.71 (LR) 0.79 (SVM)
IPV	0.50 (LR) 0.62 (SVM)	0.50 (LR) 0.62 (SVM)	0.50 (LR) 0.62 (SVM)
macro avg	0.61 (LR) 0.71(SVM)	0.61 (LR) 0.71 (SVM)	**0.61** (LR) **0.71** (SVM)

(b) Using Search categories as features

	precision	recall	f1-score
NIPV	0.77 (LR) 0.86 (SVM)	0.71(LR) 0.86 (SVM)	0.74 (LR) 0.86(SVM)
IPV	0.56 (LR) 0.75 (SVM)	0.62 (LR) 0.75 (SVM)	0.59 (LR) 0.75 (SVM)
macro avg	0.66 (LR) 0.80 (SVM)	0.67 (LR) 0.80 (SVM)	**0.66** (LR) **0.80** (SVM)

Figure 5: Performance of SVM and LR on test set. (a) Features used: 12 linguistic features that were statistically significantly different among the IPV and NIPV individuals. (b) Features used: Search categories that were statistically significantly different between the two population

and stop seeking intimacy. In addition, if someone is going through sexual abuse, especially those who do not realize that he/she is a victim, it may be less likely for him/her to search about contents related to sex. For someone who recognize sexual abuse in the relationship, he/she may want to hide the situation due to shame and stigma associated to this issue. However, we acknowledge that in a large sample size sex-related searchers may be more prevalent among IPV victims.

Lack of financial independence, social support and outside social activities can be some of the potential circumstances associated to abusive relationship [80, 81, 82]. Victims may often feel paranoid, isolated, and compelled to stay indoor; hence, it may not be surprising that they can seek help from the internet for ways to feel more secure at home and outside. In particular, during our analysis we observed frequent occurrences of alarming search queries such as, “change home computer password”, “kids mother abusing us”, “how to change locks without telling bf”, and “signs husband cheating”.

5. Modeling & Results

For the IPV detection task, we ask the question: *Given the search logs of an individual, can we identify whether the person is in an abusive relationship?* Following from the previous section, we have engineered 12 linguistic features (Fig. 3) and 8 search categories features (Table 3). For the classification task, we trained a Logistic Regression (LR) and Support Vector Machine (SVM). Prior to training, we have separated search logs from 22 individuals (8 IPV and 14 NIPV) as test set and trained the models on data from 34 individuals (16 IPV and 18 NIPV). We did not employ a decision tree based classifiers because they are prone to over-fit, especially since our data-set is small.

The performance of the IPV detection classifier on the test set is presented in Fig. 5. We considered F1 score as the metric to evaluate the models performance. Overall, we observed SVM, with F1[5] score of 0.80 (Fig.5 (b)), performed better than LR, with F1 score of 0.66, for the classification task. SVM maximizes the margin between the two classes whereas LR tries to fit a hyper-plane that best fit the training data. Our findings suggested that features based on search categories were more powerful that the features extracted using the linguistic attributes. This was evident by the fact the SVM consistently outperformed LR regardless of feature type.

[5]F1 = harmonic mean of precision and recall = $2\times\frac{precision \times recall}{precision + recall}$

N = 22 (NIPV: 14 IPV: 8)	Pred NIPV	Pred IPV
True NIPV	<10> (11)	<4> (3)
True IPV	<4> (3)	<4> (5)

(a)
Feature Used: LIWC attributes that were statistically significantly different

N = 22 (NIPV: 14 IPV: 8)	Pred NIPV	Pred IPV
True NIPV	<10> (12)	<4> (2)
True IPV	<3> (2)	<5> (6)

(b)
Feature Used: Search categories that were statistically significantly different

Figure 6: Confusion Matrix. A common metric used to quantify the performance of a classification model. Numbers within <> are for the Logistic Regression and numbers within () are for the Support Vector Machine

We also compared the performance of the models on the test set by analyzing the confusion matrix tables presented in Fig 6. Compared to LR, SVM had fewer misclassification for the IPV detection task. When linguistic features were used (Fig 6(a)), LR made a total of 8 misclassifications where as SVM had 6 miscategorizations. On the contrary, SVM made 4 miclassifications compared to LR's 7 when search category features were used (Fig.6(b)).

6. Discussion

Intimate partner violence is a serious public health concern that can negatively affect individuals and families, irrespective of race, age, gender, sexual orientation, or socioeconomic status. It remains one of the oldest problems confronting society. Exposure to IPV can lead to both physical and mental health related problems, including depression, low self-esteem, psychological distress, and post-traumatic stress disorder (PTSD).

While IPV is preventable, early detection is difficult. Current methods for detecting IPV by self-assessment surveys and in-person interviews are time consuming, expensive, lack precision, and are hampered by factors such as fear and stigma associated with the problem. We propose a less intrusive, cost effective, instant, scalable, and lightweight technique to screen for IPV using search engine logs. To the best of our knowledge, we are the first to propose such a data driven empirical screening technique for detecting the presence of violence and abuse in one's personal life.

Our analysis showed that there are statistically significant differences in searching behavior, in terms of temporal, textual (i.e. linguistic), and contextual (search categories), between individuals who have and who have not experienced IPV. Using these signals from search logs we were able to build models that can detect incidence of IPV in one's relationship with high accuracy. The advantage of search history based screening, over the in-person interview or survey based questionnaires, is that it is scalable, lightweight, does not require human supervision, maintains anonymity, and can be readily deployed at any location, such as counselling and health care centers. Finally, our work opens a way for therapists, clinicians, care providers, and other concerned agencies to better understand potential troubling situations in their clients/patients personal life with high fidelity and accuracy far beyond the limits of traditional self-reported data. For example, knowing that a client is suspicious of his/her significant other may help a therapist to appropriately intervene at the right moment, preventing the situation from escalating further.

6.1. Limitations

The major limitation of this pilot study is its limited sample size. While our findings are consistent with what might be expected based on prior, 'traditional' IPV research, they can not be considered definitive. However, our results clearly underscore the need for further research, and highlight the potential contributions the AI community may bring to the field.

Secondly, traditional survey based IPV detection tools, although clinically proven and state-of-the-art, can not guarantee that a participant may (i) conceal information or choose the more socially acceptable answer rather than being honest; (ii) not be able to access themselves accurately; (iii) have grown up in different geographical location and have varying understanding and interpretation of intimate partner violence; (iv) have varying subjective interpretation of the survey questions; (v) need to recall something from the past (to answer survey questions) which may result in inaccurate response.

Thirdly, while training the IPV models we used the linguistic and search categories features individually. Incorporating both feature sets might have significantly increased the dimension of the data which could eventually negatively affect our model performance.

Finally, the most important limitation is related to the degree to which individuals engage in searching for information online. And, while this type of activity is rapidly approaching universality, the approach to composing search queries can vary substantially between individuals. Further work is necessary to determine how best to account for such differences in practice.

6.2. *Practical Implication*

Using our lightweight screening method for detecting evidence of IPV, we believe counselling services, mental and physical health professionals, and other relevant agencies can have capabilities to recognize clinically relevant events in their patients/clients personal lives which may not otherwise be possible.

The sensitive and subtle consequences of abusive relationship may not be easily evident during in-person interview sessions. However, analysis of search history may help identify concerns early enough in a relationship before it become more serious. It is quite possible to build a mobile application that can constantly monitor one's daily online activities and look for markers that may identify one to someone who may be experiencing violent or troubling relationships. The application may have a dashboard that aggregates user activities and generates report that can be shared, electronically, with anyone. This app can be used to direct resources related to educating individuals and loved ones about how to identify abusive/unhealthy relationships, ways to overcome the stigma associated to IPV, how to improve quality of intimate relationships, consider self-care and how to protect oneself from abusive relationships etc. Furthermore, a user of such an app can link his/her account with their care provider. Hence, at the time of detecting alarming signals through the search logs, appropriate medical, social, and emotional support can be offered to the individual at risk.

6.3. *Privacy & Ethical Considerations*

There are unique ethical dimensions related to work in this area, especially in terms of the importance of privacy and personal safety. As outlined above, we designed this pilot study with rigorous protections in place for subjects, approved by the Institutional Review Board of our institution. In terms of building and deploying automated systems along the lines we have demonstrated to be potentially feasible here, we want to further acknowledge the ethical challenges that would pose. For instance, how could such an automated screening process, which could be prone to errors, identify sufficiently accurate signals for IPV, and ensure one's privacy? How could one address the potential consequences of misinterpreting or misrepresenting any of the temporal, linguistic, or contextual cues that may indicate IPV? After all, just because an individual searches for things relating to partner violence or domestic abuse, it may not mean that he or she has *directly* experienced IPV personally.

There are significant ethical challenges at the micro (individual), meso (family and local community), and macro (regional, state, and national) levels posed by this line of work. A detailed discussion of these myriad concerns is beyond the scope of the current paper. We focus instead on the main ethical challenges posted by this work at the micro level. In particular, there are potential threats to individual privacy and safety when data such as search logs are used for secondary purposes. Users likely never intended such secondary uses of their personal data, and may often refuse to allow it, if given direct control over data they generate or contribute to. The traditional notion of informed consent remains a fundamental component of current research ethics. The concept refers to an individual's right to self-determination based on clear, accurate, and timely information before, during, and after participation in research.

To address this concern, in part, we use an exclusively opt-in model for study participation. In the commercial world, web search data is commonly collected at scale with little semblance of a process of informed consent that would satisfy even the most lenient standards of human subjects protections. Many users of online search services are unaware that the data and metadata they are generating is being collected or stored, much less used for any purpose beyond the primary action of conducting an online search. Even among those who are aware that their data is being collected, few users of search services would ever contemplate that their data would ever be used for purposes far beyond their initial intent, namely to find what they are seeking online quickly, efficiently, and precisely.

7. Future Work

We acknowledge that our findings are based on a small dataset collected from a very specific population group. In order to establish stronger generalizable signals of IPV, based on search history, we need to conduct the study on larger general population. Following that, it is quite feasible to build robust predictive models for identifying presence of IPV through internet usage data. In addition, we aim to leverage data such as Youtube and Google map usage history to uncover one's quotidian online activities.

The accessibility of individual level search log data opens the door to studying signals for other phenomena,such as anxiety, stress, PTSD, panic disorder, etc., associated to IPV. We are working with other researchers and clinicians at our university's medical center to include this work as a collaboration opportunity between mental and public health researchers.

8. Conclusions

In this paper, we show that individual level search histories can be used for identifying abusive relationship. Our work is an empirical demonstration that daily online search history has the potential to capture intimate partner violence among couples. We employed both the linguistic and contextual aspects of search logs to build model for identifying individuals in abusive relationships. We look forward to expanding our analysis across different forms of online activities and eagerly explore signals of other forms of related phenomena.

References

[1] L. K. Hamberger, D. G. Saunders, M. Hovey, Prevalence of domestic violence in community practice and rate of physician inquiry., Family medicine 24 (4) (1992) 283–287.

[2] N. E. Gin, L. Rucker, S. Frayne, R. Cygan, F. A. Hubbell, Prevalence of domestic violence among patients in three ambulatory care internal medicine clinics, Journal of General Internal Medicine 6 (4) (1991) 317–322.

[3] S. R. Dearwater, J. H. Coben, J. C. Campbell, G. Nah, N. Glass, E. McLoughlin, B. Bekemeier, Prevalence of intimate partner abuse in women treated at community hospital emergency departments, Jama 280 (5) (1998) 433–438.

[4] A. Browne, B. Miller, E. Maguin, Prevalence and severity of lifetime physical and sexual victimization among incarcerated women., International journal of law and psychiatry (1999).

[5] G. T. Hotaling, D. B. Sugarman, An analysis of risk markers in husband to wife violence: The current state of knowledge, Violence and victims 1 (2) (1986) 101–124.

[6] R. M. Cate, J. M. Henton, J. Koval, F. S. Christopher, S. Lloyd, Premarital abuse: A social psychological perspective, Journal of Family Issues 3 (1) (1982) 79–90.

[7] M. A. Straus, R. J. Gelles, S. K. Steinmetz, Behind closed doors: Violence in the American family, Routledge, 2017.

[8] J. E. Deal, K. S. Wampler, Dating violence: The primacy of previous experience, Journal of Social and Personal relationships 3 (4) (1986) 457–471.

[9] P. Mullen, V. Walton, S. Romans-Clarkson, G. P. Herbison, Impact of sexual and physical abuse on women's mental health, The Lancet 331 (8590) (1988) 841–845.

[10] S. Plichta, The effects of woman abuse on health care utilization and health status: a literature review, Women's Health Issues 2 (3) (1992) 154–163.

[11] J. C. Campbell, L. A. Lewandowski, Mental and physical health effects of intimate partner violence on women and children, Psychiatric Clinics of North America 20 (2) (1997) 353–374.

[12] A. L. Coker, P. H. Smith, L. Bethea, M. R. King, R. E. McKeown, Physical health consequences of physical and psychological intimate partner violence, Archives of family medicine 9 (5) (2000) 451.

[13] A. L. Coker, K. E. Davis, I. Arias, S. Desai, M. Sanderson, H. M. Brandt, P. H. Smith, Physical and mental health effects of intimate partner violence for men and women, American journal of preventive medicine 23 (4) (2002) 260–268.

[14] J. C. Campbell, Health consequences of intimate partner violence, The lancet 359 (9314) (2002) 1331–1336.

[15] L. Heise, C. Garcia-Moreno, Violence by intimate partners. (2002).

[16] C. Garcia-Moreno, H. A. Jansen, M. Ellsberg, L. Heise, C. H. Watts, et al., Prevalence of intimate partner violence: findings from the who multi-country study on women's health and domestic violence, The lancet 368 (9543) (2006) 1260–1269.

[17] M. C. Black, K. C. Basile, M. J. Breiding, S. G. Smith, M. L. Walters, M. T. Merrick, M. R. Stevens, et al., The national intimate partner and sexual violence survey: 2010 summary report, Atlanta, GA: National Center for Injury Prevention and Control, Centers for Disease Control and Prevention 19 (2011) 39–40.

[18] R. Jewkes, Intimate partner violence: causes and prevention, The lancet 359 (9315) (2002) 1423–1429.

[19] J. Archer, Sex differences in aggression between heterosexual partners: a meta-analytic review., Psychological bulletin 126 (5) (2000) 651.

[20] M. K. Ehrensaft, T. E. Moffitt, A. Caspi, Clinically abusive relationships in an unselected birth cohort: men's and women's participation and developmental antecedents., Journal of abnormal psychology 113 (2) (2004) 258.

[21] P. A. Brand, A. H. Kidd, Frequency of physical aggression in heterosexual and female homosexual dyads, Psychological Reports 59 (3) (1986) 1307–1313.

[22] C. M. Renzetti, Violence in lesbian relationships: A preliminary analysis of causal factors, Journal of Interpersonal violence 3 (4) (1988) 381–399.
[23] M. C. McHugh, I. H. Frieze, Intimate partner violence, Annals of the New York Academy of Sciences 1087 (1) (2006) 121–141.
[24] E. W. Gondolf, E. Fisher, J. R. McFerron, Racial differences among shelter residents: A comparison of anglo, black, and hispanic battered, Journal of Family Violence 3 (1) (1988) 39–51.
[25] P. Tjaden, N. Thoennes, Prevalence and consequences of male-to-female and female-to-male intimate partner violence as measured by the national violence against women survey, Violence against women 6 (2) (2000) 142–161.
[26] M. C. Ellsberg, R. Pena, A. Herrera, J. Liljestrand, A. Winkvist, Wife abuse among women of childbearing age in nicaragua., American journal of public health 89 (2) (1999) 241–244.
[27] K. A. Miczek, J. F. DeBold, M. Haney, J. Tidey, J. Vivian, E. M. Weerts, Alcohol, drugs of abuse, aggression, and violence, Understanding and preventing violence 3 (1994).
[28] P. Harvey, P. Gow, Sex and violence: issues in representation and experience, Routledge, 2013.
[29] M. Ellsberg, L. Heise, R. Pena, S. Agurto, A. Winkvist, Researching domestic violence against women: methodological and ethical considerations, Studies in family planning 32 (1) (2001) 1–16.
[30] J. W. Ayers, B. M. Althouse, J.-P. Allem, J. N. Rosenquist, D. E. Ford, Seasonality in seeking mental health information on google, American journal of preventive medicine 44 (5) (2013) 520–525.
[31] E. Yom-Tov, S. H. Fischer, The werther effect revisited: Measuring the effect of news items on user behavior, in: Proceedings of the 26th International Conference on World Wide Web Companion, International World Wide Web Conferences Steering Committee, 2017, pp. 1561–1566.
[32] W. H. Organization, Changing cultural and social norms that support violence (2009).
[33] H. A. Beydoun, M. Williams, M. A. Beydoun, S. M. Eid, A. B. Zonderman, Relationship of physical intimate partner violence with mental health diagnoses in the nationwide emergency department sample, Journal of Women's Health 26 (2) (2017) 141–151.
[34] Google, Google trends (2017).
URL `https://trends.google.com/trends/`
[35] R. W. White, E. Horvitz, Evaluation of the feasibility of screening patients for early signs of lung carcinoma in web search logs, JAMA oncology 3 (3) (2017) 398–401.
[36] R. W. White, P. M. Doraiswamy, E. Horvitz, Detecting neurodegenerative disorders from web search signals, NPJ digital medicine 1 (1) (2018) 1–4.
[37] J. Ginsberg, M. H. Mohebbi, R. S. Patel, L. Brammer, M. S. Smolinski, L. Brilliant, Detecting influenza epidemics using search engine query data, Nature 457 (7232) (2009) 1012.
[38] H. A. Carneiro, E. Mylonakis, Google trends: a web-based tool for real-time surveillance of disease outbreaks, Clinical infectious diseases 49 (10) (2009) 1557–1564.
[39] S. Yin, M. Ho, Monitoring a toxicological outbreak using internet search query data, Clinical toxicology 50 (9) (2012) 818–822.
[40] M. J. McCarthy, Internet monitoring of suicide risk in the population, Journal of affective disorders 122 (3) (2010) 277–279.
[41] A. Zaman, R. Acharyya, H. Kautz, V. Silenzio, Detecting low self-esteem in youths from web search data, in: The World Wide Web Conference, ACM, 2019, pp. 2270–2280.
[42] L. Allerhand, B. Youngmann, E. Yom-Tov, D. Arkadir, Detecting parkinson's disease from interactions with a search engine: Is expert knowledge sufficient?, in: Proceedings of the 27th ACM International Conference on Information and Knowledge Management, 2018, pp. 1539–1542.
[43] G. P. Cooper Jr, V. Yeager, F. M. Burkle Jr, I. Subbarao, Twitter as a potential disaster risk reduction tool. part ii: descriptive analysis of identified twitter activity during the 2013 hattiesburg f4 tornado, PLoS currents 7 (2015).
[44] D. Freed, J. Palmer, D. E. Minchala, K. Levy, T. Ristenpart, N. Dell, Digital technologies and intimate partner violence: A qualitative analysis with multiple stakeholders, Proceedings of the ACM on Human-Computer Interaction 1 (CSCW) (2017) 46.
[45] J. P. Dimond, C. Fiesler, A. S. Bruckman, Domestic violence and information communication technologies, Interacting with computers 23 (5) (2011) 413–421.
[46] D. Freed, J. Palmer, D. Minchala, K. Levy, T. Ristenpart, N. Dell, "a stalker's paradise": How intimate partner abusers exploit technology, in: Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems, ACM, 2018, p. 667.
[47] T. Matthews, K. O'Leary, A. Turner, M. Sleeper, J. P. Woelfer, M. Shelton, C. Manthorne, E. F. Churchill, S. Consolvo, Stories from survivors: Privacy & security practices when coping with intimate partner abuse, in: Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems, ACM, 2017, pp. 2189–2201.
[48] F. Alshaikh, F. Ramzan, S. Rawaf, A. Majeed, Social network sites as a mode to collect health data: a systematic review, Journal of medical Internet research 16 (7) (2014).
[49] S. M. Strasser, M. Smith, D. Pendrick-Denney, S. Boos-Beddington, K. Chen, F. McCarty, Feasibility study of social media to reduce intimate partner violence among gay men in metro atlanta, georgia, Western journal of emergency medicine 13 (3) (2012) 298.
[50] L. E. Watkins, R. C. Maldonado, D. DiLillo, The cyber aggression in relationships scale: a new multidimensional measure of technology-based intimate partner aggression, Assessment (2016) 1073191116665696.
[51] R. F. Rabin, J. M. Jennings, J. C. Campbell, M. H. Bair-Merritt, Intimate partner violence screening tools: a systematic review, American journal of preventive medicine 36 (5) (2009) 439–445.
[52] N. Hussain, S. Sprague, K. Madden, F. N. Hussain, B. Pindiprolu, M. Bhandari, A comparison of the types of screening tool administration methods used for the detection of intimate partner violence: a systematic review and meta-analysis, Trauma, Violence, & Abuse 16 (1) (2015) 60–69.
[53] K. M. Iverson, M. W. King, M. R. Gerber, P. A. Resick, R. Kimerling, A. E. Street, D. Vogt, Accuracy of an intimate partner violence screening tool for female vha patients: A replication and extension, Journal of Traumatic Stress 28 (1) (2015) 79–82.
[54] S. R. Beach, C. R. Carpenter, T. Rosen, P. Sharps, R. Gelles, Screening and detection of elder abuse: Research opportunities and lessons learned from emergency geriatric care, intimate partner violence, and child abuse, Journal of elder abuse & neglect 28 (4-5) (2016) 185–216.

[55] M. E. Dichter, T. N. Haywood, A. E. Butler, S. L. Bellamy, K. M. Iverson, Intimate partner violence screening in the veterans health administration: demographic and military service characteristics, American journal of preventive medicine 52 (6) (2017) 761–768.
[56] M. A. Dutton, B. Mitchell, Y. Haywood, The emergency department as a violence prevention center., Journal of the American Medical Women's Association (1972) 51 (3) (1996) 92–5.
[57] J. C. Campbell, D. W. Webster, N. Glass, The danger assessment: validation of a lethality risk assessment instrument for intimate partner femicide, Journal of interpersonal violence 24 (4) (2009) 653–674.
[58] T. O. Afifi, H. MacMillan, B. J. Cox, G. J. Asmundson, M. B. Stein, J. Sareen, Mental health correlates of intimate partner violence in marital relationships in a nationally representative sample of males and females, Journal of interpersonal violence 24 (8) (2009) 1398–1417.
[59] A. E. Bonomi, M. L. Anderson, R. J. Reid, F. P. Rivara, D. Carrell, R. S. Thompson, Medical and psychosocial diagnoses in women with a history of intimate partner violence, Archives of internal medicine 169 (18) (2009) 1692–1697.
[60] D. G. Kilpatrick, What is violence against women: Defining and measuring the problem, Journal of interpersonal violence 19 (11) (2004) 1209–1234.
[61] M. A. Straus, R. J. Gelles, L. M. Asplund, Physical violence in american families: Risk factors and adaptations to violence in 8,145 families (1990).
[62] T. Logan, L. Shannnon, R. Walker, Partner violence and stalking of women: context, consequences and coping (2005).
[63] M. Scheffer Lindgren, B. Renck, 'it is still so deep-seated, the fear': psychological stress reactions as consequences of intimate partner violence, Journal of Psychiatric and Mental Health Nursing 15 (3) (2008) 219–228.
[64] T. L. Weaver, G. A. Clum, Psychological distress associated with interpersonal violence: A meta-analysis, Clinical psychology review 15 (2) (1995) 115–140.
[65] C. M. Clements, D. K. Sawhney, Coping with domestic violence: Control attributions, dysphoria, and hopelessness, Journal of traumatic stress 13 (2) (2000) 219–240.
[66] D. Umberson, K. Anderson, J. Glick, A. Shapiro, Domestic violence, personal control, and gender, Journal of Marriage and the Family (1998) 442–452.
[67] M. Rosenberg, Rosenberg self-esteem scale (rse), Acceptance and commitment therapy. Measures package 61 (1965) 52.
[68] C. R. Pineiro, Social media use and self-esteem in undergraduate students, Rowan University, 2016.
[69] D. R. Jezl, C. E. Molidor, T. L. Wright, Physical, sexual and psychological abuse in high school dating relationships: Prevalence rates and self-esteem issues, Child and adolescent social work journal 13 (1) (1996) 69–87.
[70] R. J. Aguilar, N. N. Nightingale, The impact of specific battering experiences on the self-esteem of abused women, Journal of family violence 9 (1) (1994) 35–45.
[71] C. Liu, X. Chen, P. Song, A. Lu, L. Wang, X. Zhang, Z. Huang, D. Zheng, Relationship between childhood emotional abuse and self-esteem: a dual mediation model of attachment, Social Behavior and Personality: an international journal 46 (5) (2018) 793–800.
[72] M. Burmeister, Linguistic analysis of intimate partner violence (2012).
[73] J. S. Parker, G. S. Stewart, C. Gantt, Research and intervention with adolescents exposed to domestic violence, Family Therapy 33 (1) (2006) 45.
[74] Y. R. Tausczik, J. W. Pennebaker, The psychological meaning of words: Liwc and computerized text analysis methods, Journal of language and social psychology 29 (1) (2010) 24–54.
[75] J. Pennebaker, R. Booth, R. Boyd, M. Francis, Linguistic inquiry and word count: Liwc2015. austin, tx: Pennebaker conglomerates (2015).
[76] S. Siegel, The mann-whitney u test, Nonparametric statistics for the behavioral sciences (1956) 116–127.
[77] J. Nerbonne, The secret life of pronouns. what our words say about us, Literary and Linguistic Computing 29 (1) (2014) 139–142.
[78] J. W. Pennebaker, M. R. Mehl, K. G. Niederhoffer, Psychological aspects of natural language use: Our words, our selves, Annual review of psychology 54 (1) (2003) 547–577.
[79] M. H. Kernis, Substitute needs and the distinction between fragile and secure high self-esteem, Psychological Inquiry 11 (4) (2000) 298–300.
[80] S. R. Schuler, S. M. Hashemi, A. P. Riley, S. Akhter, Credit programs, patriarchy and men's violence against women in rural bangladesh, Social science & medicine 43 (12) (1996) 1729–1742.
[81] B. Burton, N. Duvvury, N. Varia, Domestic violence in india: a summary report of a multi-site household survey. (2000).
[82] I. Kawachi, B. P. Kennedy, Socioeconomic determinants of health: Health and social cohesion: why care about income inequality?, Bmj 314 (7086) (1997) 1037.