Context is Important in Depressive Language: A Study of the Interaction Between the Sentiments and Linguistic Markers in Reddit Discussions

Neha Sharma and **Kairit Sirts**
Institute of Computer Science, University of Tartu, Estonia
neha.sharma@ut.ee, sirts@ut.ee

Abstract

Research exploring linguistic markers in individuals with depression has demonstrated that language usage can serve as an indicator of mental health. This study investigates the impact of discussion topic as context on linguistic markers and emotional expression in depression, using a Reddit dataset to explore interaction effects. Contrary to common findings, our sentiment analysis revealed a broader range of emotional intensity in depressed individuals, with both higher negative and positive sentiments than controls. This pattern was driven by posts containing no emotion words, revealing the limitations of the lexicon based approaches in capturing the full emotional context. We observed several interesting results demonstrating the importance of contextual analyses. For instance, the use of 1st person singular pronouns and words related to anger and sadness correlated with increased positive sentiments, whereas a higher rate of present-focused words was associated with more negative sentiments. Our findings highlight the importance of discussion contexts while interpreting the language used in depression, revealing that the emotional intensity and meaning of linguistic markers can vary based on the topic of discussion.

1 Introduction

Research on linguistic markers of depression aiming to identify people at risk (De Choudhury et al., 2013; Yates et al., 2017; Coppersmith et al., 2018; Chancellor and De Choudhury, 2020) have found that individuals with depression often use more first-person pronouns and negative emotion words and fewer positive emotion words than healthy controls (Trifu et al., 2017; Savekar et al., 2023; Rude et al., 2004; Chung and Pennebaker, 2007). Despite these consistent findings, their effects are relatively modest, with a recent meta-analysis showing small effect sizes (Pearson r of 0.19 for the first-person pronouns, 0.12 for negative emotion words, and -0.21 for positive emotion words)(Tølbøll, 2019).

Thus far, research has concentrated on identifying the main effects, i.e., the overall signifi-cant differences between depression and control groups. Nevertheless, as can be expected, studies have shown that the linguistic markers of the same person can vary depending on the communication context (Mehl et al., 2012).

According to Beck's cognitive model of depression (Beck, 1979; Beck and Alford, 2009), schemas of depressive symptoms develop over time and, when active, shape an individual's perceptions, thoughts, and feelings, which influence the linguistic choices of that person when expressing their thoughts and feelings. It is entirely possible and even likely, especially in the case of mild to moderate depression, that the depressive schemas are variably activated in different contexts. Thus, the schema influence on thoughts and linguistic choices are not the same in every context. Therefore, identifying distinct linguistic markers of depression may greatly rely on the context and the activation of depressive schemas at the moment. In light of these considerations, a more fine-grained analysis focusing on interaction effects is indicated, considering in which linguistic contexts these differences occur more strongly and which linguistic features co-occur consistently.

Only a few studies have explored the contextual effect on language markers for depression. Bernard et al. (2016) found that higher depression levels correlate with increased use of first-person pronouns. However, they noted that the frequent use of negative emotion words is linked more to higher negative affect than to the depression level itself, suggesting that the prevalence of negative emotion words is not depression per se but rather a negative effect on the state, which is presumably more common in individuals with depression than in healthy controls.

Seabrook et al. (2018) and Teodorescu et al. (2023) studied the association between depression and the variability of emotion word rates over time, assuming that people might express different proportions of positive and negative emotion words over time. Both studies found that negative emotion

variability was positively associated with depression. That means, two people might express the same overall proportion of negative emotion words, but the higher variability over time (e.g., a higher rate of negative emotion words on one occasion and a lower rate on another occasion) predicted higher depressive symptomatology (Seabrook et al., 2018) or the diagnostic status (Teodorescu et al., 2023).

From the computational linguistics point of view, Ireland and Iserman (2018) studied the linguistic markers of anxiety in posts originating from anxiety-related and neutral sub-reddits and found that anxious individuals' word choices differed depending on the sub-reddit. However, we are unaware of any computational linguistics study that has researched the contextual aspects of linguistic markers of depression.

In this paper, we study the linguistic variation that relates contexts to differential affective tones. Our focus is specifically on understanding the features of the so-called "depressive language", i.e., the social science aspect, and not on predicting the diagnostic labels based on textual data, which has been the predominant goal in many previous studies (Guntuku et al., 2017). This approach will help us better understand the varied and context-dependent ways depression influences how people express themselves.

We start with the assumption that the affective quality expressed in texts varies between contexts and thus also necessitates for the authors making different linguistic choices. We operationalize affective tone as sentiment and contexts as different discussion topics, assuming that some topics activate the depressive schemas more than others. To study the interactions, we use simple linear regression analysis to understand which psycholinguistic features, extracted with the LIWC tool (Pennebaker et al., 2015), correlate with the differential affective tones between depressed and control users over different topics as context. For our study, we utilize a Reddit-based dataset containing posts from various subreddits of users with and without self-declared depression diagnosis for clinical and control groups, respectively.

Our study centers on the following key research question:

RQ: *Which psycholinguistic features affect the sentiment differences observed between depression and control groups across various topics?*

To address this RQ effectively, we begin with a preliminary analysis to lay the groundwork for a deeper inquiry, i.e., ***Are there differences between depression and control groups regarding the topics discussed and sentiments expressed within our dataset?***

Based on prior research, we expected that individuals with depression will exhibit more negative sentiment and less positive sentiment (Rude et al., 2004; Savekar et al., 2023). Our analysis showed that the posts of people from the depression group showed more negative sentiment. However, contrary to our expectations, we found that the depression group also had more posts with positive sentiment. We then followed up with the analyses investigating the RQ. We found that although overall, the depression group used more negative emotion words than the control group, in the contextual analyses, the usage of anger- and sadness-related words were significantly related to the overall positive sentiment of the posts. A small-scale qualitative analysis suggested that posts containing mixed emotions (i.e., references to both positive and negative aspects) might play a role in this correlation. We conclude the paper with some further analyses investigating this direction.

2 Method

2.1 Data

	Depression	Control	Total
#users	1316	1316	2632
Total #posts	195.2K	364.4K	559.6K
Avg #posts per user (std)	148 (78)	277 (146)	213 (133)
Avg #words per post (std)	35 (38)	25 (30)	28 (33)

Table 1: Statistics of the balanced depression dataset.

We used an existing Reddit-based data set comprising posts from users with self-reported mental health diagnoses (SMHD) (Cohan et al., 2018). Each diagnosed user is matched with nine control users on average. The data set covers nine mental health diagnoses, including depression. For this study, we used the depression part only involving 1316 users with the self-reported depression diagnosis. We randomly sampled an equivalent number of 1316 control users to create a balanced data set. Additionally, we filtered out all posts containing less than three or more than 200 words. Table 1 shows the statistics of the study dataset. More de-

tails about the underlying SMHD dataset can be found in Appendix A.

2.2 Sentiment Analysis

We evaluated two sentiment models for applicability to our dataset: a RoBERTa-based model, trained on Twitter (Barbieri et al., 2020), and the lexicon-based VADER (Hutto and Gilbert, 2014). On comparing both models on a set of manually annotated 200 posts randomly drawn from our dataset, we found that although VADER demonstrated slightly higher accuracy (0.69) than RoBERTa (0.66), it more often confused posts with positive and negative sentiments. Therefore, we chose the RoBERTa-based model for subsequent analyses. For more details of the comparative analysis, refer to the Appendix B.

2.3 Topic Modeling

BERTopic (Grootendorst, 2022) leverages the power of transformer-based language models to capture the contextual information and meaning of words in each document. We used the default topic model setting for our purpose, as according to the BERTopic documentation,[1] the default model works quite well for most use cases. More than 5000 topics were initially derived from the model, with about 50% of the data classified as outliers (documents not fitting any topic, labeled as -1). Utilizing the "reduced outlier" function, which leverages the c-TF-IDF strategy and cosine similarity, the proportion of outliers was reduced to 0.25%. Additionally, the "automatic topic reduction" function was applied to reduce the number of topics, resulting in 4187 topics.

2.4 LIWC Analysis

LIWC (Linguistic Inquiry and Word Count) is a lexicon-based tool that analyzes texts by mapping words to psycho-linguistic attributes, resulting in the proportion of words in various categories. These categories include Summary Variables, Linguistic Dimensions, Psychological Processes, and more, detailed in (Pennebaker et al., 2015) and Appendix D. Among the 110 LIWC attributes, we selected 63 attributes relevant to our research objectives that were not highly correlated to each other to avoid multicollinearity in the subsequent linear regression analysis. We assessed the correlations using Pearson's correlation coefficient with

[1]https://maartengr.github.io/BERTopic/index.html

a threshold of 0.5. Appendix D shows the list of selected attributes.

2.4.1 User-based LIWC Analysis

Because our data is a balanced and length-restricted subset of the SMHD depression dataset, we first analyze the difference in LIWC attributes between the depression and control groups and compare our results to those reported by Cohan et al. (2018). Similarly to the cited study, the group means are aggregated over users, i.e., for computing the feature values for a single user, all their posts were first concatenated. Similarly, we performed Welch's t-test (Welch, 1947) with adjusted p-value using Bonferroni correction. For effect size assessment, we calculated Cohen's d statistics (Cohen, 1987).

2.4.2 Topic-Specific LIWC Analysis

In the main analysis of the paper, we wanted to understand which LIWC attributes affect the sentiments expressed in relation to various topics. Moreover, we wanted to capture the sentiment differences between depression and control groups. For that, we employed linear regression analysis with the features derived from LIWC attributes as independent variables and an overall sentiment polarity difference between groups as the dependent variable.

The dependent variable, i.e., the sentiment difference between depression and control groups per topic, is calculated as the net sentiment score difference denoted as $y = (pos - neg)_{\text{depression}} - (pos - neg)_{\text{control}}$ (see Table 2 for some examples), where pos and neg columns show the percentage of positive and negative sentiments per topic and group, respectively. For instance, in Topic 0 (Animals), for the depression group, 28.0% of posts are labeled as positive, whereas 26.0% have a negative sentiment. The net difference is 2%, showing that overall, the depression group has slightly more positive sentiment towards that topic. In contrast, for the control group, the net difference is $19.1\% - 23.9\% = -4.8\%$, showing that over-

Topics	Depression		Control		Sent diff
	pos	neg	pos	neg	y
0 Animals	28.0	26.0	19.1	23.9	6.8
1 Relationships	14.4	46.4	13.4	31.3	-14.1
2 US elections	20.2	18.5	16.8	73.4	-5.0

Table 2: Dependent variable calculation.

all, the control group has more negative sentiment towards that topic. Subtracting these differences $y = 2 - (-4.8) = 6.8$ yields an outcome value capturing the overall difference between depression and control groups towards that topic. Positive difference refers to more positive sentiment in the depression group posts, while negative value means more positive sentiment in the control group posts. Values close to zero indicate the similarity of positive and negative sentiment proportions in both depression and control groups.

Independent features were computed in two steps. First, we calculated mean aggregated LIWC attribute scores for both groups topic-wise. For instance, consider Topic 0 (Animals) and the Analytic feature (see Table 3). For the depression group, the average score of that attribute for Topic 0 is 43.0, while for the control group it is 49.7. Subsequently, we calculated the difference between these scores, i.e., $f_{\text{Analytic}} = 43.0 - 49.7 = -6.7$. Positive feature values refer to higher proportion of the attribute value in the depression group, while negative feature values mean that control group had more of that attribute. Independent features were computed this way for all topics and 63 selected attributes, resulting in a size 4187×63 (topics $\times$ attributes) matrix. This dataset was used to fit the linear regression model using the ordinary least squares method.

Topics	Analytic			Clout		
	Dep	Ctr	f_A	Dep	Ctr	f_C
0 Animals	43.0	49.7	-6.7	47.3	49.6	-2.3
1 Relationships	25.5	26.5	-1.0	59.3	62.3	-3.0
2 US elections	48.8	57.2	-8.4	32.4	34.2	-1.8

Table 3: Independent feature calculation for the regression analysis (Dep = Depression group, Ctr = Control group, $f_A = f_{\text{Analytic}}$, $f_C = f_{\text{Clout}}$).

3 Results

3.1 User-based LIWC analysis

The analysis, shown in Table 4, revealed more significant differences and larger effect sizes between the groups than those reported by Cohan et al. (2018). This discrepancy may arise firstly because we only tested 63 pre-selected, uncorrelated attributes, which makes the adjusted p-value threshold higher than it would be with the full attribute set used by Cohan et al. (2018). Secondly, although
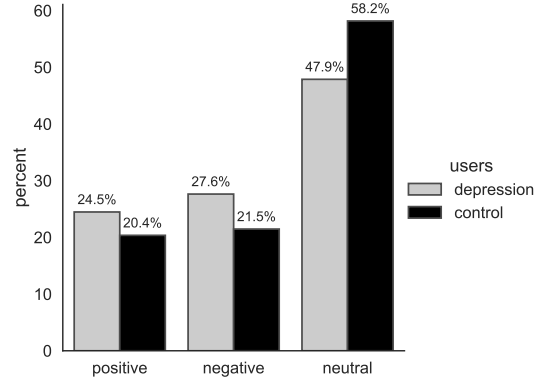

Figure 1: Sentiment distribution of both groups.

Cohan et al. (2018) does not report means and standard deviations, we expect the standard deviations to be smaller in our subset due to restrictions on the post length, which affects both the attribute significance and the magnitude of the effect sizes.

Most attributes that showed statistical significance in both studies (13 out of 14) had consistent effect sizes, except for the present focus. While Cohan et al. (2018) found a small positive effect size, indicating a higher proportion of present-focused words in the depression group, our subsample showed a negative effect size, aligning with studies suggesting that individuals with depression are generally less present-focused (Nolen-Hoeksema et al., 2008; Rodriguez et al., 2010). Thus, our random subsample appears to be representative of the whole depression part of the SMHD data.

3.2 Sentiment Analysis

As shown in Figure 1, the depression group exhibits more negative sentiments (6%) and, unexpectedly, more positive sentiments (4%) than the control group. This variation in sentiment distribution is statistically supported by the Chi-square test ($\chi^2(2) = 5503.79, p < 0.001$). This suggests that the users with depression not only express more negative emotionality but display an overall higher emotionality in their posts.

3.3 Topic Modeling

Topic modeling resulted in 4187 topics discussed by both groups, of which 13 low-frequency topics were unique to one group only. These topics and their posts were removed from the data, resulting in 4174 topics. We analyzed the frequency and sentiment distribution of these topics among the user groups. Despite the general common topic discussions, noticeable differences in topic preva-

LIWC Attributes	Depression		Control		p-value	Cohen's d this study	Cohen's d Cohan et al. (2018)
	mean	std	mean	std			
Word Count	36.1	14.5	24.6	10.9	***	0.90	N/S
Analytic	42.4	9.9	48.8	10.0	***	−0.63	N/S
Clout	37.5	10.8	40.6	9.7	***	−0.30	-0.06
Authentic	56.0	9.7	49.6	9.7	***	0.65	0.2
1st person singular	6.1	1.8	4.7	1.8	***	0.77	0.23
3rd person singular	1.2	0.9	1.0	0.7	***	0.29	0.09
Impersonal pronouns	5.8	0.9	5.5	1.2	***	0.22	0.06
Insight	2.8	0.7	2.4	0.7	***	0.49	0.09
Causation	1.7	0.4	1.6	0.5	**	0.15	N/S
Tentative	3.1	0.7	2.9	0.8	***	0.28	0.07
Certitude	0.9	0.3	0.8	0.3	***	0.34	N/S
Differentiation	3.7	0.7	3.6	0.9	***	0.19	0.08
Negative emotion	0.7	0.3	0.5	0.3	***	0.65	N/S
Anxiety	0.1	0.1	0.1	0.1	***	0.60	0.07
Anger	0.2	0.1	0.1	0.1	**	0.16	N/S
Sadness	0.1	0.1	0.1	0.1	***	0.44	N/S
Swear words	0.4	0.4	0.6	0.7	***	−0.24	N/S
Social behavior	3.7	0.9	3.5	0.9	***	0.26	N/S
Interpersonal	0.3	0.2	0.3	0.2	*	−0.15	N/S
Communication	1.7	0.6	1.5	0.6	***	0.37	N/S
Family	0.4	0.4	0.2	0.3	***	0.40	0.06
Friends	0.2	0.2	0.2	0.2	***	0.25	N/S
Female references	0.8	0.7	0.5	0.6	***	0.45	0.13
Culture	0.7	0.5	0.9	0.9	***	−0.35	N/S
Lifestyle	3.0	1.0	3.3	1.2	***	−0.27	N/S
Illness	0.2	0.2	0.1	0.1	***	0.54	N/S
Wellness	0.1	0.1	0.1	0.1	***	0.26	N/S
Mental health	0.1	0.1	0.02	0.07	***	0.50	N/S
Substances	0.1	0.2	0.1	0.1	***	0.21	N/S
Want	0.4	0.2	0.3	0.2	***	0.29	N/S
Fatigue	0.04	0.05	0.03	0.07	*	0.15	N/S
Reward	0.1	0.1	0.2	0.2	***	−0.28	N/S
Risk	0.3	0.1	0.2	0.1	***	0.25	N/S
Perception	8.3	1.2	8.5	1.6	**	−0.16	N/S
Feeling	0.5	0.3	0.4	0.3	***	0.42	N/S
Past focus	3.6	1.0	3.2	1.0	***	0.41	0.08
Present focus	3.9	0.8	4.1	1.0	***	−0.27	0.09
Conversational	1.0	0.6	1.3	0.9	***	−0.38	N/S

Table 4: Difference between psycholinguistic LIWC attributes between depression and control groups. Only statistically significant attributes are shown. Adjusted p-value thresholds for multiple comparisons is 0.00079 and denoted as * < 0.00079, ** < 0.00015, *** < 0.000015, equivalent to * < 0.05, ** < 0.01, *** < 0.001. The table includes effect sizes (Cohen's d) for both this study and Cohan et al. (2018), with 'N/S' indicating no statistical difference found in the latter study.

lence and sentiment patterns emerged between the depression and control groups. While the detailed examination of topic modeling outputs is not the central focus of our research, as its purpose is to segment the data into contexts for further analyses, the differences in topic distributions and sentiment distributions in topics between groups might be interesting in their own right. We provide some further analyses in Appendix C.

3.4 Topic-Specific LIWC Analysis

Using a linear regression model with 63 LIWC attributes as predictors, we assessed the sentiment differences outlined in Section 2.4.2. The model accounted for 26.6% of the variability in sentiment difference, $R^2 = 0.266, F(63, 4110) = 23.63, p < 0.001$, with 25 attributes being statistically significant ($p = 0.05$). Figure 2 illustrates the impact of the significant attributes on the sentiment difference outcome variable. Recall that positive attribute values refer to the depression group expressing more of that attribute in a topic. Thus, positive model weights indicate these features are associated with higher positive sentiment in the depression group posts across topics. Complete results are in Appendix E.
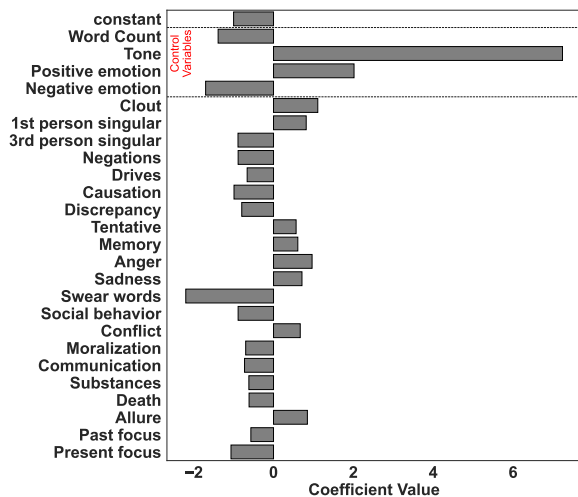
Figure 2: The linear regression model coefficients of the statistically significant features and their impact on the sentiment difference.

The bias term was significant, and its negative value shows that overall, the sentiment of the depression group tends to be more negative than that of the control group. Control variables such as word count, tone, and positive and negative emotions showed expected correlations with sentiment values. The average post length in our data was longer for the depression group compared to the control group (see Table 1). The negative coefficient shows that higher positive sentiment for the depression group is associated with shorter posts than the control group. The positive and negative emotion features are expected to be closely correlated to the positive and negative sentiment values. The coefficients of these features are in the expected direction. Finally, the tone feature, which includes both positive and negative tone, has the largest absolute coefficient and a positive weight, having the largest impact on the outcome variable.

A higher proportion of first-person singular pronouns is related to a higher positive sentiment, and a higher proportion of third-person singular pronouns is related to a higher negative sentiment. Time orientation features, such as past focus and present focus, have negative coefficients, meaning that a higher proportion of words in those categories are related to a higher negative sentiment. There are some features that, with their negative weight, are perhaps quite expectedly related to a higher negative sentiment: swear words, moralization, substances, and death. In contrast, more conflict-related words are associated with higher positive sentiment.

Finally, we highlight the two remaining emotion-related features: anger and sadness. These features have positive weights, correlating with more positive sentiment. In contrast to other findings mentioned above, these results are surprising, as anger and sadness as instances of negative emotions are expected to be more correlated with overall negative sentiment. In the next section, we will attempt to understand these findings.

4 Anger and Sadness

In the previous section, we found that the use of anger- and sadness-related words, typically seen as negative emotion words, correlates with positive sentiment. In order to get some idea of the observed phenomenon, we reviewed some of the posts with positive sentiments that contained anger- and sadness-related words. Examples 1–3 shown in Table 5 suggest that these posts express what might be called *mixed emotions*, i.e., containing a mixed usage of positive and negative emotion words. Building on this observation, we designed two analyses to study 1) if posts with mixed emotions could be responsible for overall higher positive sentiments observed for the depression group, and 2) if posts with mixed emotions might significantly contribute to the observed positive relationship between the anger- and sadness-related feature and positive sentiment.

Mixed Emotions: First, we need to operationalize what it means for a post to display mixed emotions. We define posts with mixed emotions as those that contain both positive and negative emotion words, i.e., the LIWC attribute of both positive and negative emotions is greater than zero. We aimed to examine the role of mixed emotion posts in our findings by comparing subsets of data both with and without mixed emotions. Given the possible combination of emotion words, we further categorized the data into four segments rather than with or without mixed emotions, which are: Mixed Emotions (both positive and negative emotion words, 3.6% of the total posts), Positive Emotions (only positive emotion words, 15.9%), Negative Emotions (only negative words, 8.7%), and Neutral Emotions (neither positive nor negative words, 71.8%).

Sentiment Distribution: First, we looked at the differences in sentiment proportions between depression and control groups in each of those data

No	Topic	Post text	Sentiment	Attribute
1	Art	I truly admire and appreciate the art; it's impressive. Yet, it's also causing me a great deal of frustration. It's amazing, though.	Positive	Anger
2	Empathy	In times of sadness, we seek understanding and compassion. It's music that has the power to uplift our spirits.	Positive	Sadness
3	Albums	My all-time favorite musical work is the second symphony. It deeply saddens me.	Positive	Sadness
4	Animals	Whenever I'm feeling low, a walk with my dog always helps. He invariably does something silly or amusing during our walk, which never fails to lift my spirits.	Positive	1st person pronoun
5	Gym	Previously, I relied on gym buddies, but their absence meant I stopped too. Now, I've taken control—working out alone, focusing on my diet, and tracking my progress. Sometimes friends join, but mostly, it's just me. This self-reliance has led to sustained success for the first time. My motivation and achievements are my own, though I welcome occasional companionship and encouragement. This self-empowered approach has transformed into my lifestyle, leaving no room for excuses, for myself or others.	Positive	1st person pronoun
6	Family	My father left when I was a child, leaving me confused about his reasons. Over time, living with my mother helped me understand his choice, though being with her has been challenging. I wish he hadn't left on my birthday. Despite this, we've reconnected and improved our relationship.	Negative	Past focus
7	Emotions	Do you ever worry that just as life gets better, something bad will happen? This fear of sudden, negative changes when things are going well makes me hesitant to fully invest myself. How do you deal with this anxiety?	Negative	Present focus

Table 5: Posts of depressed users with sentiment label and relative language attributes and associated topics. All posts have been rephrased to maintain the privacy of users.
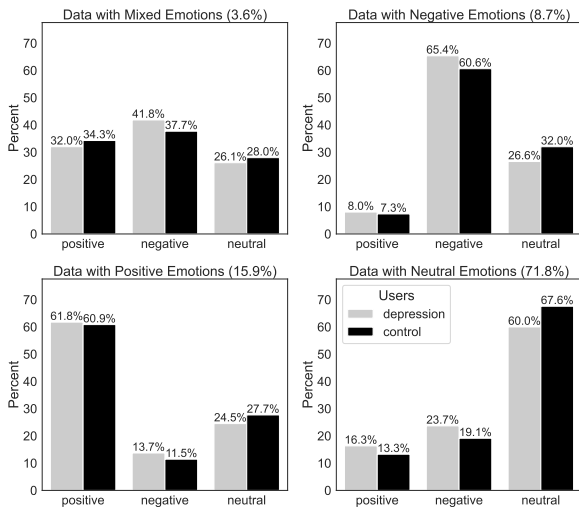

Figure 3: Sentiment distribution across data splits.

splits (shown in Figure 3). As expected, negative sentiments dominate in Negative Emotions split, positive sentiments in Positive Emotions split, and neutral sentiment in Neutral Emotions split. Mixed Emotions data shows an almost uniform distribution over sentiments. When looking at group differences in the positive sentiment, the proportion is similar in both groups in both the Mixed Emotions, Positive Emotions and Negative Emotions split, while in the Neutral Emotions group the depres-

sion group has more positive sentiment. Because the Neutral split is the biggest (72%), we conclude that this split, instead of the Mixed Emotions as we expected, drives the overall sentiment pattern observed in Figure 1. This analysis highlights an important limitation of lexicon-based systems, which struggle to grasp the overall context and sentiment in the absence of polarized emotional words. While the RoBERTa-based sentiment model is not perfect, it can capture emotional tone that is concealed from the lexicon-based LIWC system.

Anger and Sadness in Mixed Emotions: Next, we explored if mixed emotion posts might be related to positive correlations between anger and sadness features and the positive sentiment as found in Section 3.4. Figure 4 plots the median anger and sadness scores for the overall data, Mixed Emotions split and the Negative Emotions split.[2] Both anger and sadness scores were highest in the Negative Emotions split and lowest in the Mixed Emotions split across all sentiments. However, the overall pattern of median anger and sadness scores in Mixed Emotions split differs from other splits. In contrast to other slits, both the median anger and sadness scores are *highest* in

---

[2]Positive and Neutral Emotions splits are omitted as by definition they do not exhibit any negative emotion words.
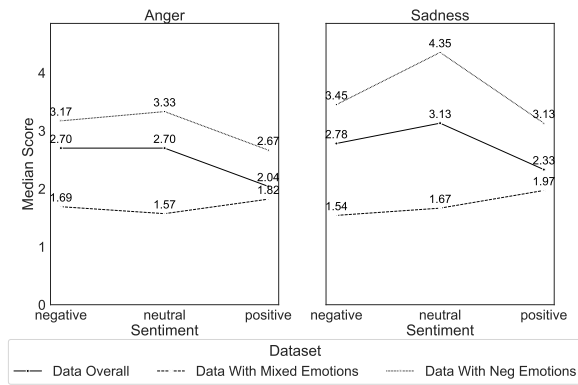
Figure 4: Median anger and sadness scores across data splits.

positive sentiment posts.

5 Discussion

This study aimed to investigate the relationship between (psycho)linguistic features and affective tone across contexts operationalized as topics in depressed and non-depressed reddit users. Overall sentiment analysis revealed that the depression group expressed more negative sentiment, aligning with prior studies (Liu et al., 2022), yet also more positive sentiment, which was unexpected. In further analyses that split the data into four subsections regarding the presence or absence of positive or negative emotion words, we found that the so-called neutral posts containing neither positive nor negative emotion words according to LIWC were driving this pattern. This finding highlighting the constraints of the lexicon-based systems, which fail to capture the full emotional context in absence of explicit emotional words.

A higher proportion of first-person singular pronouns was related to a higher positive sentiment, contrasting with previous research linking it to negative emotionality (Tackman et al., 2019; Bernard et al., 2016). However, another study Brockmeyer et al. (2015) found that for people with depression, the higher usage of first-person singular pronouns was related to texts elicited in the positive memory recall task but not in the negative memory recall task. Thus, although the majority of studies (Lyons et al., 2018; Rude et al., 2004; Stirman and Pennebaker, 2001; De Choudhury et al., 2013; Savekar et al., 2023; Trifu et al., 2017; Chung and Pennebaker, 2007) (including this one) have found the overall higher usage of first-person pronouns by individuals with depression, the interaction studies considering either affective, topical, or other con-

texts might show a different and more interesting picture, the analysis of which could be a topic of future studies. Examples 4 and 5 in Table 5 provide illustration; despite the self-referential nature of the language, the sentiments expressed in these posts are notably positive.

In the main effect LIWC analyses, the depression group displayed more past-focused language and less present-focused language, similar to previous works (Trifu et al., 2017; Smirnova et al., 2013; Imbault and Kuperman, 2018). At the same time, in linear regression analyses, both past-focused and present-focused features were negatively associated with the positive sentiment difference. As an illustration, in the Example 6 in Table 5, the user expresses negative affect in relation to past experiences. In contrast, in Example 7, the user conveys a current state of worry, leading to negative sentiments about present circumstances, demonstrating that in certain contexts, present focus might be an indication for depressive language.

Finally, we observed a significantly higher use of negative emotional words in the depression group, including anxiety-, anger- and sadness-related words, which is in line with previous studies (Trifu et al., 2017; Savekar et al., 2023; Rude et al., 2004; Chung and Pennebaker, 2007). In the linear regression analysis, although the association between the negative emotion words and the sentiment difference were, similarly to the main effect, negative, the correlations between the sentiment difference and the anger and sadness features were positive, i.e., the higher rate of anger and sadness words were related to more positive sentiment. Qualitative analysis of positive sentiment posts with non-zero anger or sadness scores revealed a pattern of mixed emotions, i.e., texts containing features of both positivity and negativity, such as starting by describing something negative, but ending in a positive note. When exploring the potential role of mixed emotions in this relation, we found that in contrast to other types of posts, the posts with mixed emotions have the highest anger and sadness scores in posts with positive sentiment. While this result does not provide definitive evidence for the role of mixed emotions in the observed positive correlations between anger and sadness features and positive sentiment, it shows that posts with mixed emotions behave differently from other posts containing negative emotion words and thus can play a different role in the depressive language.

6 Conclusion

In conclusion, our research highlights the important role of discussion context in shaping emotional expressions among individuals with (but also without) depression. Contrary to prior studies, we observed not only more negative sentiments but also more positive sentiments within the depression group—a pattern that was not captured by the LIWC, illustrating the limitations accurately interpreting emotions in the absence of explicit emotional words. Analyses also revealed notable interactions between linguistic markers—such as anger and sadness—and positive sentiments, suggesting a potentially important role of posts with mixed emotions. In summary, our findings support the notion that the research in linguistic markers of depression requires going beyond studying main effects and necessitates a contextual and multifaceted approach.

Limitations

There are several limitations to consider in our study. Firstly, the validity of the sentiment analysis model cannot be ensured because although the model is trained on social media data (Twitter), our data comprises Reddit posts, which, even with the length restriction imposed on our subsample, are, on average, considerably longer than tweets. We assessed the model's performance by manually annotating a random subset of 200 posts and found that the disagreements stemmed mostly from the model's tendency to categorize positive and negative posts as neutral. Thus, it is likely that the amount of posts with positive and negative sentiments is somewhat underestimated.

Two key limitations concern the SMHD dataset utilized. Firstly, the dataset spans from 2006 to 2017. It is important to acknowledge that the presence and severity of depression may vary over time for individuals (Harrigian and Dredze, 2022). There is a possibility that some users labeled as depressed did not have depression during the entire timeline, and such temporal uncertainty may impact the interpretation of results. Second, the control group is auxiliary (Ernala et al., 2019), i.e., although the control group was selected from non-mental health-related sub-Reddits, there is no way to be sure if controls are actually controls or if there are users in the control group who might be on the spectrum with any mental health disorders.

The LIWC tool, despite its widespread use, faces limitations due to its lexicon-based approach par-

ticularly with contextual nuances. An example includes incorrectly assigning a high anger emotion score to a statement like "Rita Madder is freaking great" because of the word "Madder," despite no anger being expressed. Our findings further validate this limitation, as neutral emotion settings revealed a significant difference in positive sentiment distributions among groups, highlighting the inadequacy of lexicon-based systems in capturing the true sentiment context.

Additionally, our regression analysis explains only 26% of the variance of the sentiment difference, indicating that unaccounted factors might influence the observed patterns in language and sentiments. This limitation could partly stem from LIWC's challenges in capturing context. Moreover, our findings are derived from a dataset specifically concerning depression, limiting the generalizability of our conclusions to broader populations and contexts. Further research with diverse datasets is necessary to apply these results more universally.

Ethical Considerations

In our study, we analyzed the language of the social media posts of both depressed and non-depressed users. We used the existing SMHD dataset (Cohan et al., 2018) that we obtained from its creators by signing a user agreement; we have adhered to the terms and conditions outlined in this agreement when conducting this study. In our work, we search for general patterns and do not make predictions or draw conclusions about any particular user in the dataset. Also, we believe that our findings are interesting for the social science sphere, however, we believe that they will not be directly useful for drawing conclusions about users posting in social media.

Data and Code Availability Statement

In the interest of fostering transparency and reproducibility, the source code supporting the findings of this study is publicly available. The code repository, which includes the scripts and any additional documentation necessary for replicating the analyses and results presented in this paper, can be accessed at the following GitHub link.[3] For access to the data itself, please contact the authors of Cohan et al. (2018).

---

[3]https://github.com/nehasharma666/Depression

Acknowledgement

References

Francesco Barbieri, Jose Camacho-Collados, Luis Espinosa Anke, and Leonardo Neves. 2020. TweetEval: Unified benchmark and comparative evaluation for tweet classification. In *Findings of the Association for Computational Linguistics: EMNLP 2020*, pages 1644–1650, Online. Association for Computational Linguistics.

Aaron T. Beck. 1979. *Cognitive therapy of depression.* Guilford press.

Aaron T. Beck and Brad A. Alford. 2009. *Depression: Causes and treatment*. University of Pennsylvania Press.

Jared D. Bernard, Jenna L. Baddeley, Benjamin F. Rodriguez, and Philip A. Burke. 2016. Depression, language, and affect: an examination of the influence of baseline depression and affect induction on language. *Journal of Language and Social Psychology*, 35(3):317–326.

Timo Brockmeyer, Johannes Zimmermann, Dominika Kulessa, Martin Hautzinger, Hinrich Bents, Hans-Christoph Friederich, Wolfgang Herzog, and Matthias Backenstrass. 2015. Me, myself, and I: self-referent word use as an indicator of self-focused attention in relation to depression and anxiety. *Frontiers in psychology*, 6:1564.

Stevie Chancellor and Munmun De Choudhury. 2020. Methods in predictive techniques for mental health status on social media: a critical review. *NPJ digital medicine*, 3(1):43.

Cindy Chung and James Pennebaker. 2007. The psychological functions of function words. In *Social communication*, pages 343–359. Psychology Press.

Arman Cohan, Bart Desmet, Andrew Yates, Luca Soldaini, Sean MacAvaney, and Nazli Goharian. 2018. SMHD: a Large-Scale Resource for Exploring Online Language Usage for Multiple Mental Health Conditions. In *Proceedings of the 27th International Conference on Computational Linguistics*, pages 1485–1497.

Jacob Cohen. 1987. Statistical power analysis for the behavioral sciences (revised edition). *Laurence Erlbaum Associates: Hillsdale, NJ, USA*.

Glen Coppersmith, Ryan Leary, Patrick Crutchley, and Alex Fine. 2018. Natural language processing of social media as screening for suicide risk. *Biomedical informatics insights*, 10.

Munmun De Choudhury, Michael Gamon, Scott Counts, and Eric Horvitz. 2013. Predicting depression via social media. In *Proceedings of the international AAAI conference on web and social media*, 1, pages 128–137.

Sindhu Kiranmai Ernala, Michael L. Birnbaum, Kristin A. Candan, Asra F. Rizvi, William A. Sterling, John M. Kane, and Munmun De Choudhury. 2019. Methodological gaps in predicting mental health states from social media: Triangulating diagnostic signals. In *Proceedings of the 2019 CHI conference on human factors in computing systems*, pages 1–16.

Maarten R. Grootendorst. 2022. BERTopic: Neural topic modeling with a class-based TF-IDF procedure. *ArXiv*, abs/2203.05794.

Sharath Chandra Guntuku, David B. Yaden, Margaret L. Kern, Lyle H. Ungar, and Johannes C. Eichstaedt. 2017. Detecting depression and mental illness on social media: an integrative review. *Current Opinion in Behavioral Sciences*, 18:43–49.

Keith Harrigian and Mark Dredze. 2022. Then and now: Quantifying the longitudinal validity of self-disclosed depression diagnoses. In *Proceedings of the Eighth Workshop on Computational Linguistics and Clinical Psychology*, pages 59–75.

Clayton Hutto and Eric Gilbert. 2014. Vader: A parsimonious rule-based model for sentiment analysis of social media text. In *Proceedings of the international AAAI conference on web and social media*, 1, pages 216–225.

Constance Imbault and Victor Kuperman. 2018. Emotional reactivity and perspective-taking in individuals with and without severe depressive symptoms. *Scientific reports*, 8(1):7634.

Molly Ireland and Micah Iserman. 2018. Within and between-person differences in language used across anxiety support and neutral reddit communities. In *Proceedings of the Fifth Workshop on Computational Linguistics and Clinical Psychology: From Keyboard to Clinic*, pages 182–193.

Tony Liu, Jonah Meyerhoff, Johannes C. Eichstaedt, Chris J. Karr, Susan M. Kaiser, Konrad P. Kording, David C. Mohr, and Lyle H. Ungar. 2022. The relationship between text message sentiment and self-reported depression. *Journal of affective disorders*, 302:7–14.

Minna Lyons, Nazli Deniz Aksayli, and Gayle Brewer. 2018. Mental distress and language use: Linguistic analysis of discussion forum posts. *Computers in Human Behavior*, 87:207–211.

Mattias R. Mehl, Megan L. Robbins, and Shannon E. Holleran. 2012. How taking a word for a word can be problematic: Context-dependent linguistic markers of extraversion and neuroticism. *Journal of Methods and Measurement in the Social Sciences*, 3(2):30–50.

Susan Nolen-Hoeksema, Blair E Wisco, and Sonja Lyubomirsky. 2008. Rethinking rumination. *Perspectives on psychological science*, 3(5):400–424.

James W. Pennebaker, Ryan L. Boyd, Kayla Jordan, and Kate Blackburn. 2015. The development and psychometric properties of LIWC2015. Technical report, the University of Texas in Austin.

Aubrey J. Rodriguez, Shannon E. Holleran, and Matthias R. Mehl. 2010. Reading between the lines: The lay assessment of subclinical depression from written self-descriptions. *Journal of personality*, 78(2):575–598.

Stephanie Rude, Eva-Maria Gortner, and James Pennebaker. 2004. Language use of depressed and depression-vulnerable college students. *Cognition & Emotion*, 18(8):1121–1133.

Anbu Savekar, Shashikanta Tarai, and Moksha Singh. 2023. Structural and functional markers of language signify the symptomatic effect of depression: A systematic literature review. *European Journal of Applied Linguistics*, 11(1):190–224.

Elizabeth M. Seabrook, Margaret L. Kern, Ben D. Fulcher, and Nikki S. Rickard. 2018. Predicting depression from language-based emotion dynamics: longitudinal analysis of facebook and twitter status updates. *Journal of medical Internet research*, 20(5):e168.

D. Smirnova, E. Sloeva, N. Kuvshinova, A. Krasnov, D. Romanov, and G. Nosachev. 2013. Language changes as an important psychopathological phenomenon of mild depression. *European Psychiatry*, 28:1.

Shannon Wiltsey Stirman and James W. Pennebaker. 2001. Word use in the poetry of suicidal and non-suicidal poets. *Psychosomatic medicine*, 63(4):517–522.

Allison M. Tackman, David A. Sbarra, Angela L. Carey, M. Brent Donnellan, Andrea B. Horn, Nicholas S. Holtzman, To'Meisha S. Edwards, James W. Pennebaker, and Matthias R. Mehl. 2019. Depression, negative emotionality, and self-referential language: A multi-lab, multi-measure, and multi-language-task research synthesis. *Journal of personality and social psychology*, 116(5):817.

Daniela Teodorescu, Tiffany Cheng, Alona Fyshe, and Saif Mohammad. 2023. Language and mental health: Measures of emotion dynamics from text as linguistic biosocial markers. In *Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing*, pages 3117–3133.

Katrine Bønneland Tølbøll. 2019. Linguistic features in depression: a meta-analysis. *Journal of Language Works-Sprogvidenskabeligt Studentertidsskrift*, 4(2):39–59.

Raluca Nicoleta Trifu, Bogdan Nemeş, Carolina Bodea-Hategan, and Doina Cozman. 2017. Linguistic indicators of language in major depressive disorder (MDD). An evidence based research. *Journal of Evidence-Based Psychotherapies*, 17(1).

Bernard L. Welch. 1947. The generalization of Student's problem when several different population variances are involved. *Biometrika*, 34(1-2):28–35.

Andrew Yates, Arman Cohan, and Nazli Goharian. 2017. Depression and self-harm risk assessment in online forums. In *Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing*, pages 2968–2978.

Appendices

A   Dataset

SMHD is a collection of self-reported mental health diagnoses from Reddit, designed for academic and research purposes (Cohan et al., 2018). SMHD contains posts of Reddit users with nine mental health disorders along with matched control users from a period spanning from January 2006 to December 2017 including depression, ADHD, anxiety, bipolar, PTSD, autism, OCD, schizophrenia, and eating disorder.

The SMHD dataset includes posts from 20,406 clinical users who have claimed to have been diagnosed with a mental health condition and 335,952 control users who are unlikely to have one of the mental health conditions studied. The clinical users were identified based on the textual patterns of self-reported diagnosis (e.g., I was diagnosed with depression) and keywords related to diagnoses (language related to mental health such as the name of a condition, and general terms like diagnosis, mental illness, or suffering from, etc.). Control users were selected based on the criteria that they had not posted in any mental health related subreddits. Control users were selected from a group of potential candidates based on their similarity to clinical users, determined by their subreddit activity and number of posts. The criteria for selecting control users were rigorous: candidates were excluded if they did not meet the required subreddit overlap or minimum post count criteria or used any mental health-related terms in their posts.

After collecting all related user posts, the authors of Cohan et al. (2018) removed all mental health-related posts for clinical users to make the data of both user groups similar. In the dataset, each user is represented by a unique identification number, and their data includes the text of the posts made

by that user. Table 6 shows the number of posts and tokens per diagnosis.

Diagnosis	Posts	Tokens
depression	1,272K	57.4M
adhd	872K	40.5M
anxiety	795K	36.9M
bipolar	575K	26.2M
ptsd	258K	13.7M
autism	248K	11.6M
ocd	203K	9.4M
schizophrenia	123K	6.1M
eating	53K	2.5M
control	115,669K	3,031.6M

Table 6: SMHD dataset statistics.

B Sentiment Analysis

We evaluated two sentiment analysis models for applicability to our dataset: a transformer-based model, RoBERTa, trained on Twitter (Barbieri et al., 2020), and the VADER sentiment analysis model (Hutto and Gilbert, 2014). After manually annotating 200 posts for ground truth by the first author of this paper, we compared both models' predictions against these annotations.

The VADER model exhibited a slightly higher overall accuracy of 69% compared to the 66% achieved by the transformer-based model (see Table 7). At the same time, compared to the RoBERTa model, VADER shows much lower precision and recall for the positive class, considerably lower precision for the neutral class, and considerably lower recall for the negative class. When looking at the errors made by both models (see the confusion matrices in Figure 5), we saw that while the RoBERTa model tended to confuse the positive or negative posts as neutral, i.e., it tended to overpredict the neutral label, VADER more often confused positive and negative posts. The latter is the limitation of its lexicon-based approach due to its reliance on a static list of sentiment-laden words without considering the broader context in which they appear. Because for our study, we considered confusing negative posts as positive and vice versa as a more severe error than predicting neutral instead of either valence, we chose the RoBERTa model despite its somewhat lower accuracy.

C Topic modeling

We explored the topics by looking 1) at the frequency of posts by user groups and 2) the sentiment distribution of each topic across user groups.

Topic frequencies: The frequency distribution of the top 50 most frequent topics is shown in Figure 6, which shows the proportion of posts for each topic, normalized by the total posts per group. These top 50 topics represent 11.3% of the depression posts and 11.2% of the control posts. See Figure 7 for more details about the topics and their word representations, as extracted from the BERTopic.

As shown in Figure 6, in our dataset, the depression group discusses topics that are related to, for example, animals, relationships, dieting, books, music albums, movies, grooming, LGBTQ themes, and emotions more compared to controls. Whereas the control group discusses topics related to Karambit/Gaming items, US elections, colors, ammunition, Reddit voting, cars, Online Gaming medals, football, rocket league/Online gaming items trade more compared to depression.

Topic and Sentiment Interactions: In addition to topics having different frequency distributions by groups, they also show different sentiment patterns. Figure 8 shows the sentiment distribution of both groups for the top 50 topics. For instance, regarding discussions related to animals, the depression group expressed more positive sentiments than the control group. On the other hand, for the relationship-related topic, the depression group expressed more negative sentiments. On the topic of US elections, although it is more frequent in the control group, the depression group expresses slightly more negative sentiments. In topics related to relationships, family, and LGBTQ, the depression and control users express a similar proportion of positive sentiment, while the depression users also express considerably more negative sentiment. In terms of the adult content-related topic, while the most prevalent sentiment is positive for both groups, the control group expresses considerably more positive sentiment and less negative sentiment than the depression group. At the same time, although most sentiments towards the topic related to animals are neutral, the depression group expresses more positive sentiments than the control group. There are several other topics where the depression group expresses the more positive sentiment; the ones with the most visible difference between the

	Precision		Recall		F1	
	Vader	RoBERTa	VADER	RoBERTa	VADER	RoBERTa
Positive	0.64	0.71	0.68	0.75	0.66	0.73
Negative	0.81	0.53	0.50	0.79	0.62	0.64
Neutral	0.66	0.88	0.85	0.48	0.74	0.62
Accuracy					0.69	0.66
Macro	0.70	0.71	0.68	0.68	0.67	0.67
Weighted	0.71	0.72	0.69	0.66	0.68	0.66

Table 7: Classification Report for the VADER and RoBERTa sentiment models.
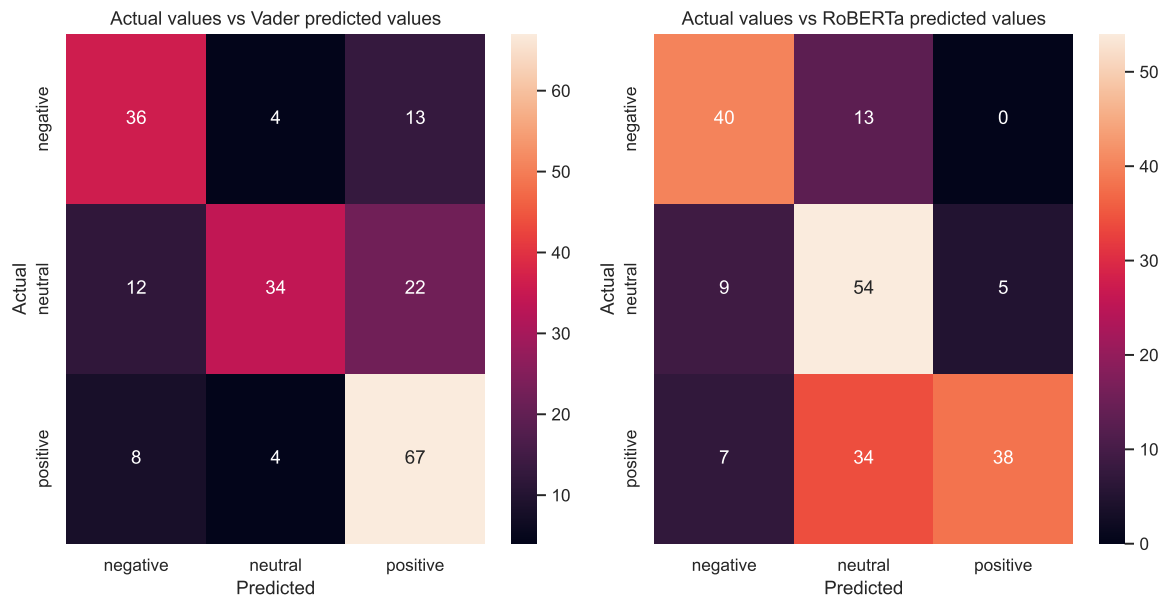

Figure 5: Confusion matrices of the VADER and RoBERTa predicted values compared to the manually annotated subset of 200 posts.
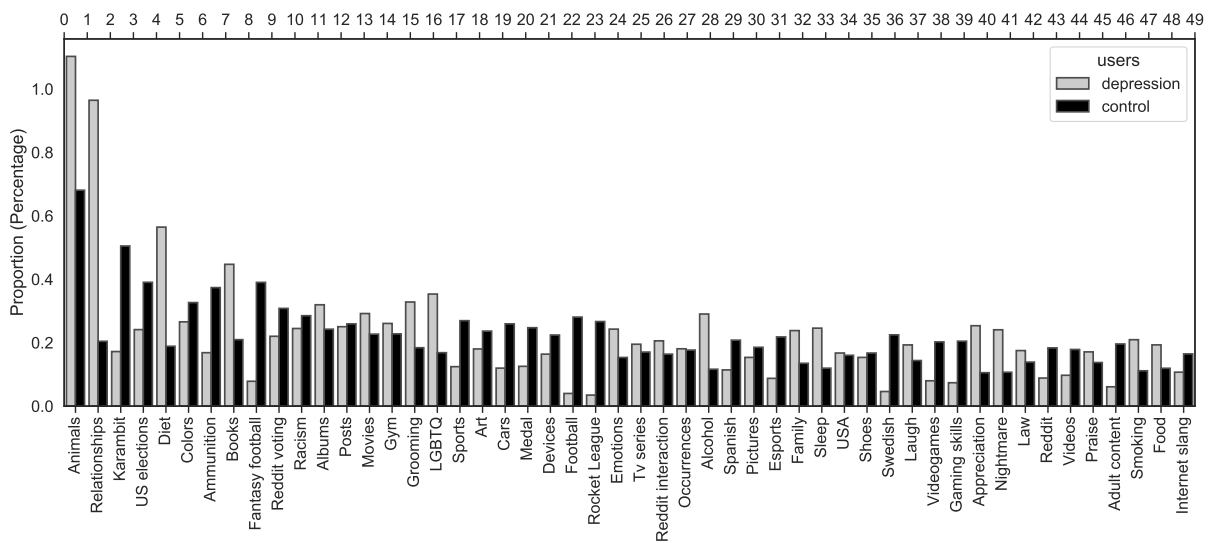

Figure 6: Frequency distribution of top 50 topics.

groups are related to grooming (topic 15), art (topic 18), TV series (topic 25), Reddit interactions (topic 26) and gaming skills (topic 39).

D LIWC

Linguistic Inquiry and Word Count (LIWC-22) software, developed by Pennebaker et al. (2015) was used in our research.[4]

LIWC contains four summary variables: Analytic Thinking, Clout, Authenticity, and Emotional Tone. The Analytic Thinking variable indicates how people use words that suggest formal, logical, and hierarchical thinking patterns. Clout is an indicator that refers to the relative social status, confidence, or leadership-related language. Authenticity shows the degree to which a person is self-monitoring, i.e., spontaneous language use with no self-regulation and filters. Emotional Tone is an indicator of positive and negative emotional tone dimensions. The linguistic dimensions contain attributes representing the percentage of words in a given text containing pronouns, articles, verbs, etc. Psychological processes contain the attributes related to cognitive processes, affect (emotional state, emotional tone), and social processes (social behavior, social references). The extended dictionary contains attributes related to culture, lifestyle, physical health, time orientations, and conversational aspects. A full description of these attributes can be found in (Pennebaker et al., 2015). The selected 63 attributes and their description are in Table 8.

E Linear regression results

Table 9 represents the impact of various linguistic and psychological attributes on the sentiment difference between the depression and control groups. Attributes associated with positive coefficients indicate a positive influence on sentiment in the depression group. In contrast, the negative coefficients suggest a negative influence on the depression group sentiment.

---

[4]We acknowledge that the use of LIWC is subject to a license, and we have obtained the necessary license for research purposes from the official LIWC website `https://www.liwc.app/`

Id	Topic	Word Representation
0	Animals	cat, dog, dogs, cats, kitten, puppy, name, pet, names, kitty
1	Relationships	relationship, ex, her, she, feelings, relationships, cheating, dating, cheated, breakup
2	Karambit	keys, steam, karambit, float, mw, doppler, bayonet, ft, ch, fv
3	US elections	trump, bernie, hillary, vote, obama, election, donald, clinton, president, marchagainsttrump
4	Diet	weight, fat, calories, calorie, pounds, lbs, overweight, muscle, diet, skinny
5	Colors	color, colors, green, blue, purple, red, yellow, colour, black, colorblind
6	Ammunition	rifle, ammo, gun, barrel, rifles, recoil, mag, shotgun, scope, bolt
7	Books	books, book, kindle, read, reading, author, novel, series, novels, textbooks
8	Fantasy football	wr, rb, ppr, rbs, scoring, wrs, te, gordon, fournette, jordy
9	Reddit voting	upvote, downvote, downvoted, upvotes, downvotes, upvoted, downvoting, voted, vote, votes
10	Racism	asian, racist, white, racism, privilege, black, race, blacks, asians, whites
11	Albums	album, song, albums, songs, music, band, lyrics, listen, listened, listening
12	Posts	posted, comment, posting, history, comments, post, threads, thread, posts, deleted
13	Movies	movie, film, movies, filming, scene, films, watched, filmed, watch, critics
14	Gym	gym, squat, squats, deadlift, crossfit, lifting, workout, bench, weights, lift
15	Grooming	hair, beard, shave, shaving, haircut, razor, shaved, mustache, hairy, hairs
16	LGBTQ	gay, marriage, wife, divorce, marry, married, divorced, husband, straight, homo
17	Sports	lebron, football, nba, rugby, basketball, kobe, jordan, soccer, sport, cavs
18	Art	art, paint, artwork, illustration, deadline, www, painting, commissions, storyboarding, infographics
19	Cars	ninja, truck, wrx, bmw, car, subaru, cars, mustang, suv, vw
20	Medal	mmr, bronze, silver, rank, ranked, plat, diamond, challenger, ranking, gold
21	Devices	watches, watch, phones, damasko, htc, galaxy, samsung, phone, iphone, nokia
22	Football	liverpool, arsenal, rooney, chelsea, striker, manchester, goal, messi, united, suarez
23	Rocket League	crates, crimson, heatwave, loopers, offers, lime, cert, keys, painted, tw
24	Emotions	hear, glad, welcome, sorry, heard, problem, help, bro, happy, hope
25	Tv series	episode, episodes, show, season, seasons, watched, eps, watching, shows, aired
26	Reddit interaction	reply, read, response, reading, responding, answer, replied, thanks, detailed, comment
27	Occurrences	happen, happened, happening, happens, exact, happend, gonna, prolly, often, going
28	Alcohol	drinking, drunk, sober, drink, alcohol, hangover, alcoholic, drinks, drank, alcoholism
29	Spanish	que, de, el, la, en, por, para, lo, je, los
30	Pictures	imgur, pics, pic, photo, pictures, photos, picture, image, upload, images
31	Esports	tsm, clg, fnatic, na, lcs, teams, eu, skt, worlds, fnc
32	Family	dad, mom, grandpa, mother, daddy, father, son, mum, uncle, naugthy
33	Sleep	sleep, wake, asleep, awake, waking, bed, nap, woke, sleeping, insomnia
34	USA	nc, ohio, michigan, carolina, virginia, ucf, nj, county, lsu, kentucky
35	Shoes	shoes, shoe, boots, heels, feet, sneakers, toe, pair, toes, wear
36	Swedish	det, att, som, och, inte, jag, har, om, en, med
37	Laugh	laughed, laugh, laughing, chuckle, laughter, loud, giggles, amused, hard, made
38	Videogames	xbox, console, consoles, pc, sony, gaming, platform, exclusives, xbone, playstation
39	Gaming skills	levelled, invention, skill, levelling, level, achieved, mining, skills, inventioni, thieving
40	Appreciation	advice, appreciate, thank, thanks, appreciated, excellent, heeney, oversee, solid, jtr
41	Nightmare	dream, dreams, nightmares, lucid, dreaming, nightmare, dreamt, vivid, woke, dreamed
42	Law	lawyer, jury, justice, judge, guilty, court, sue, lawsuit, innocent, lawyers
43	Reddit	added, ready, adding, online, available, trade, rn, add, whenever, spritzee
44	Videos	video, videos, youtube, vid, vids, link, uploaded, channel, cutinhalf, upload
45	Praise	awesome, amazing, cool, thank, incredible, dope, thanks, wonderful, dude, brilliant
46	Adult content	kik, sexy, hot, body, ass, ahm, meeting, hmu, chat, alexandraoweny
47	Smoking	smoking, smoke, cigarettes, cigarette, smokers, smoker, nicotine, smoked, smokes, tobacco
48	Food	breakfast, dinner, hungry, eat, meal, food, eating, lunch, meals, eats
49	Internet slang	gt, lt, ftfy, gtfo, gts, lolzdudezorz, div, zoop, drumjunkie, checkerboard

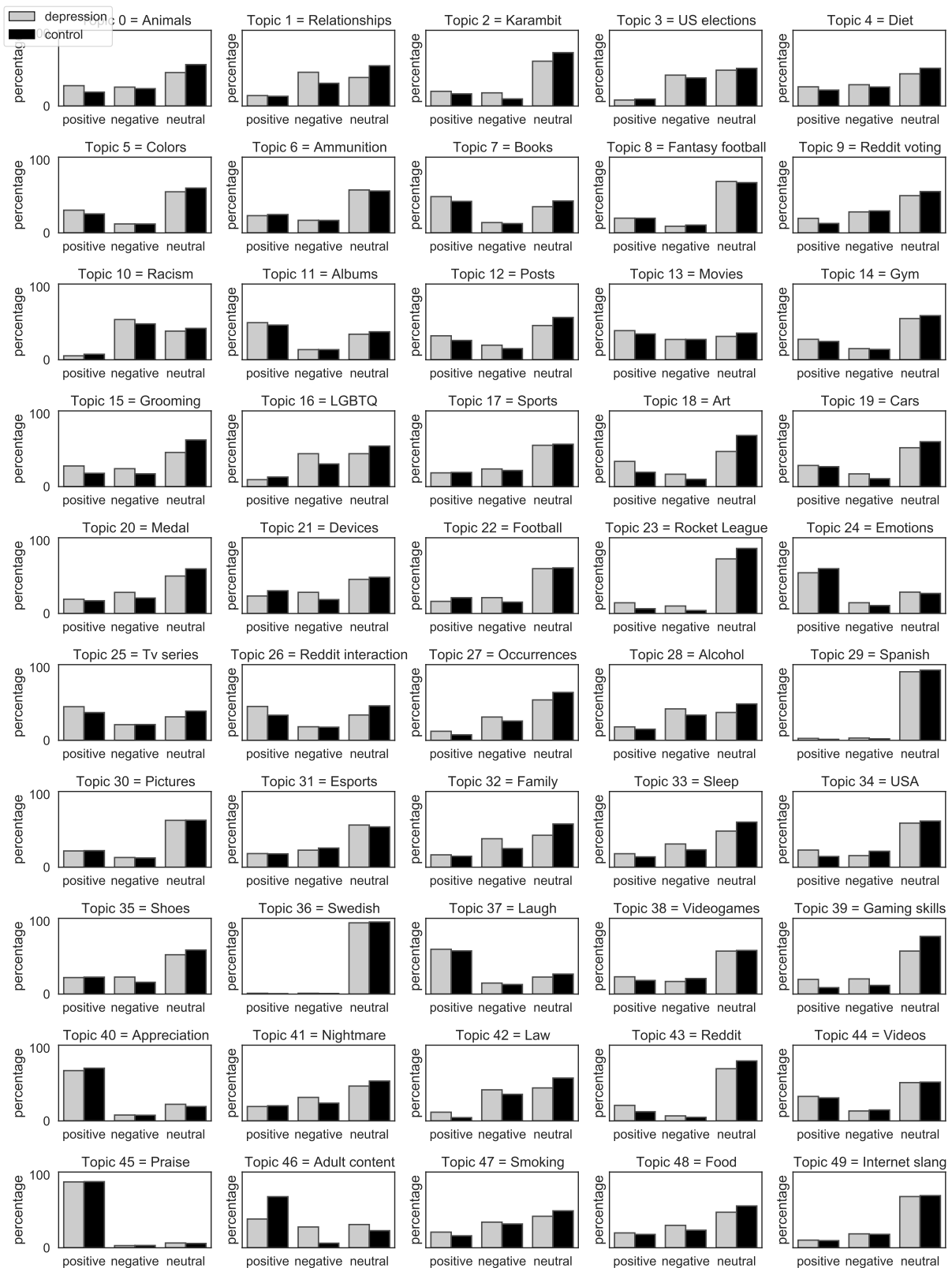Figure 7: Top 50 topics and their word representations.

Figure 8: Top 50 topics and their sentiment distributions.

Category	Description	Category	Description
Summary Variables		*Expanded Dictionary*	
Word Count	Total word count	Culture	car, united states, govern*, phone
Analytical thinking	Metric of logical, formal thinking	Lifestyle	work, home, school, working
Clout	Language of leadership, status	Religion	god, hell, christmas*, church
Authentic	Perceived honesty, genuineness	*Physical*	
Emotional tone	Degree of positive (negative) tone	*Health*	
Linguistic Dimensions		Illness	hospital*, cancer*, sick, pain
pronouns		Wellness	healthy, gym*, supported, diet
1st person singular	I, me, my, myself	Mental health	mental health, depressed, suicid*, trauma*
1st person plural	we, our, us, lets	Substances	beer*, wine, drunk, cigar*
2nd person	you, your, u, yourself	Sexual	sex, gay, pregnan*, dick
3rd person singular	he, she, her, his	Food	food*, drink*, eat, dinner*
3rd person plural	they, their, them, themsel*	Death	death*, dead, die, kill
Impersonal pronouns	that, it, this, what	*States*	
Auxiliary verbs	is, was, be, have	Need	have to, need, had to, must
Negations	not, no, never, nothing	Want	want, hope, wanted, wish
Psychological Processes		Acquire	get, got, take, getting
Drives	we, our, work, us	Lack	don't have, didn't have, *less, hungry
Cognition		Fulfilled	enough, full, complete, extra
All-or-none	all, no, never, always	Fatigue	tired, bored, don't care, boring
cognitive processes		*Motives*	
Insight	know, how, think, feel	Reward	opportun*, win, gain*, benefit*
Causation	how, because, make, why	Risk	secur*, protect*, pain, risk*
Discrepancy	would, can, want, could	Curiosity	scien*, look* for, research*, wonder
Tentative	if, or, any, something	Allure	have, like, out, know
Certitude	really, actually, of course, real	*Perception*	in, out, up, there
Differentiation	but, not, if, or	quad Feeling	feel, hard, cool, felt
Memory	remember, forget, remind, forgot	*Time orientation*	
Affect		Past focus	was, had, were, been
Emotion		Present focus	is, are, I'm, can
Positive emotion	good, love, happy, hope	Future focus	will, going to, have to, may
Negative emotion	bad, hate, hurt, tired	Conversational	yeah, oh, yes, okay
Anxiety	worry, fear, afraid, nervous		
Anger	hate, mad, angry, frustr*		
Sadness	:(, sad, disappoint*, cry		
Swear words	shit, fuckin*, fuck, damn		
Social processes			
Social behavior	said, love, say, care		
Politeness	thank, please, thanks, good morning		
Interpersonal	conflict fight, kill, killed, attack		
Moralization	wrong, honor*, deserv*, judge		
Communication	said, say, tell, thank*		
Social referents			
Family	parent*, mother*, father*, baby		
Friends	friend*, boyfriend*, girlfriend*, dude		
Female references	she, her, girl, woman		
Male references	he, his, him, man		

Table 8: Selected LIWC attributes and their descriptions.

Attribute	Coefficient	SE	t-value	p-value	95% CI
constant	−0.9998	0.248	−4.037	<0.001	−1.485 to −0.514
Summary Variables					
Word Count	−1.3876	0.263	−5.274	<0.001	−1.903 to −0.872
Clout	1.1090	0.485	2.285	0.022	0.158 to 2.060
Emotional tone	7.2384	0.287	25.198	<0.001	6.675 to 7.802
Linguistic Dimensions					
pronouns					
1st person singular	0.8189	0.377	2.175	0.030	0.081 to 1.557
3rd person singular	−0.8896	0.387	−2.298	0.022	−1.648 to −0.131
Negations	−0.8863	0.304	−2.916	0.004	−1.482 to −0.290
Psychological Processes					
Drives	−0.6603	0.298	−2.216	0.027	−1.245 to −0.076
Cognition					
cognitive processes					
Causation	−0.9897	0.264	−3.747	<0.001	−1.508 to −0.472
Discrepancy	−0.7955	0.320	−2.485	0.013	−1.423 to −0.168
Tentative	0.5659	0.271	2.086	0.037	0.034 to 1.098
Memory	0.6089	0.251	2.422	0.015	0.116 to 1.102
Affect					
Emotion					
Positive emotion	2.0170	0.280	7.195	<0.001	1.467 to 2.567
Negative emotion	−1.7006	0.420	−4.045	<0.001	−2.525 to −0.876
Anger	0.9660	0.342	2.828	0.005	0.296 to 1.636
Sadness	0.7130	0.296	2.408	0.016	0.132 to 1.294
Swear words	−2.1978	0.272	−8.086	<0.001	−2.731 to −1.665
Social processes					
Social behavior	−0.8867	0.416	−2.134	0.033	−1.701 to −0.072
Interpersonal	0.6695	0.314	2.131	0.033	0.053 to 1.286
Moralization	−0.6971	0.271	−2.575	0.010	−1.228 to −0.166
Communication	−0.7260	0.348	−2.086	0.037	−1.408 to −0.044
Expanded Dictionary					
Physical					
Health					
Substances	−0.6174	0.255	−2.419	0.016	−1.118 to −0.117
Death	−0.6125	0.284	−2.160	0.031	−1.168 to −0.057
Motives					
Allure	0.8486	0.278	3.049	0.002	0.303 to 1.394
Time orientation					
Past focus	−0.5674	0.286	−1.981	0.048	−1.129 to −0.006
Present focus	−1.0653	0.320	−3.334	0.001	−1.692 to −0.439

Table 9: Linear regression analysis summary.