Within and Between-Person Differences in Language Used Across Anxiety Support and Neutral Reddit Communities

Molly E. Ireland and **Micah Iserman**
Department of Psychological Sciences, Texas Tech University, Lubbock, Texas
{molly.ireland,micah.iserman}@ttu.edu

Abstract

Although many studies have distinguished between the social media language use of people who do and do not have a mental health condition, within-person context-sensitive comparisons (for example, analyzing individuals' language use when seeking support or discussing neutral topics) are less common. Two dictionary-based analyses of Reddit communities compared (1) anxious individuals' comments in anxiety support communities (e.g., /r/PanicParty) with the same users' comments in neutral communities (e.g., /r/todayilearned), and, (2) within popular neutral communities, comments by members of anxiety subreddits with comments by other users. Each comparison yielded theory-consistent effects as well as unexpected results that suggest novel hypotheses to be tested in the future. Results have relevance for improving researchers' and practitioners' ability to unobtrusively assess anxiety symptoms in conversations that are not explicitly about mental health.

1 Introduction

Approaches to automatically identifying general psychological distress or specific mental health conditions tend to focus on between-person comparisons, often including yoked controls that are matched on demographic characteristics (Coppersmith et al., 2016; Smith et al., 2017). Particularly in the area of computational linguistics, which historically has focused more on prediction or classification than psychological insight (cf. Schwartz et al., 2013), within-sample variance due to differences in communicative contexts is typically ignored. Such differences (for example, in how individuals who are distressed talk when they are seeking support versus having conversations that are irrelevant to mental health) may wash out in sufficiently large text samples; likewise, a common research aim is to classify a person's men-

tal health condition or distress level accurately in the absence of contextual information, given that such information is frequently unavailable (Coppersmith et al., 2015; Schwartz et al., 2016). When within-person analyses—comparing a person with themselves, versus matched controls—have been carried out in computational linguistics, the aim has typically been to identify change points over time or temporal patterns that precede important events, such as suicide attempts or panic attacks (Benton et al., 2017; Coppersmith et al., 2016; De Choudhury et al., 2016; Loveys et al., 2017).

It is clearly useful to be able to recognize distress or clinically relevant changes in situations where contextual data is absent or sparse. However, when details about the communicative context are available, understanding how individuals' goals and the social context influence language use may be valuable in interpreting linguistic signals more accurately. For example, using language to identify mental health conditions or classify symptom severity (i.e., triage) in support settings, such as crisis support forums, may be very different from attempting the same classification in everyday conversations about topics other than mental health (Friedenberg et al., 2016).

Research in psychology supports the premise that certain emotions, personality traits, and mental health symptoms manifest differently across various settings, with negative affective traits being virtually invisible in many situations (Ireland and Mehl, 2014; Mehl et al., 2012). For example, in transcripts of naturalistic recordings of students' everyday lives, neuroticism correlated with increased physical activity for men and decreased verbosity and laughter for women, with no other linguistic correlates for either sex (Mehl et al., 2006). Neuroticism—described by Jack Block as "an overinclusive, easy-to-invoke, societally evaluative wastebasket label" (Block, 2010, p. 9)—is

the Big Five trait that is typically the least legible, or most difficult to reliably and accurately detect in verbal or nonverbal behavior (Tskhay and Rule, 2014). Neuroticism is characterized by vulnerability to stress and negative affect, including depression, anxiety, and irritability (John and Srivastava, 1999).

There are two main reasons for the difficulty of detecting neuroticism in everyday social interactions. First, expressing negative affect publicly is often non-normative or socially undesirable. That is, people tend to dislike and avoid negativity—particularly sadness (Tiedens, 2001). Separately, neuroticism involves internalizing emotions such as anxiety and sadness (Zahn-Waxler et al., 2000), which are directed inward and do not require the involvement of other people (in contrast with other Big Five facets, such as gregariousness or conformity). As a result of these characteristics, even people ranking high in neuroticism will often avoid verbalizing their negative thoughts and feelings in public (e.g., conversations at work) and reveal those traits through negative emotional language only in private (e.g., diaries; Holleran and Mehl, 2008; Jarrold et al., 2011; Mehl et al., 2006, 2012).

Avoiding self-disclosures of sadness or anxiety may be particularly common among men (Nadeau et al., 2016), given that men are discouraged from expressing emotion in most cultures (Garside and Klimes-Dougan, 2002), and negative affect or neuroticism is more normative among women (Schmitt et al., 2008). For both sexes, strategically suppressing or masking negative affect in order to avoid social censure may present a barrier to coping with psychological distress, given that disclosing negative emotions is a critical step in seeking social support (Davison et al., 2000; Taylor et al., 2004).

Building on personality research on how neuroticism manifests across public and private contexts (Mehl et al., 2012), we are specifically interested in how individuals may suppress indicators of negative affect (anxiety, sadness, or irritation) as they move from talking in support-seeking settings—where presumably expressing negative affect is more normative—to neutral settings. As a test case, we analyzed users in subreddit communities for general anxiety, social anxiety, health anxiety, and panic disorder.

We focused on anxiety because it is enormously common, has severe consequences for individuals' well-being and health, and has been overlooked, relative to depression, in studies of language and clinical psychology. Several studies have investigated anxiety in concert with other disorders (Coppersmith et al., 2014, 2015; Gkotsis et al., 2017), but studies that focus on a single condition more commonly focus on depression (De Choudhury et al., 2013; for a review, see Conway and OConnor, 2016). Worldwide, anxiety is the second most prevalent mental health condition and, among all mental disorders, accounts for the second greatest variance in disability-adjusted life years (Whiteford et al., 2013). Anxiety is frequently comorbid with depression (Sartorius et al., 1996), the primary cause of suicidality, but contributes unique variance to the prediction of suicide attempts and deaths by suicide (Khan et al., 2002).

Past research on the linguistic indicators of anxiety on social media has shown that anxious individuals' language use resembles the more general distress pattern observed in other mental health conditions (particularly depression) and neuroticism (Resnik et al., 2013, 2015). This pattern includes more references to negative affect (particularly anxiety words for anxious individuals), greater self-focus, more tentativeness, more references to health, and, in some cases, more socially distant language, relative to average (Coppersmith et al., 2014, 2015; Resnik et al., 2013, 2015).

We selected Reddit for analysis because of its large base of daily active users and broad range of well-defined, active communities (or subreddits) on both mental health and other topics (Barthel et al., 2016). Subreddits are defined by clear descriptions and rules. For example, the sidebar of one anxiety support forum states, "Welcome to /r/PanicParty. This subreddit is intended to be a place of help and support for those suffering from anxiety and panic disorders." As a result, at least for the more narrowly defined mental health communities, subreddits comprise relatively coherent groups of people who all assert that they have the symptoms described in the group's rules. Although not all commenters will be suffering from the anxiety symptoms they are discussing at the time of posting, there is an expectation that community members have experienced anxiety themselves and are not participating solely in an expert (or voyeuristic) capacity. Because the same Reddit users often post in both mental health sup-

Subreddit	Anxiety	Comparison
AskReddit	5746	5901
relationships	580	571
politics	506	1096
Advice	382	173
funny	378	749
pics	361	763
aww	351	443
news	326	743
AskWomen	322	426
worldnews	312	623
todayilearned	273	473
gaming	253	548
Showerthoughts	239	464
CasualConversation	238	530
videos	194	509
gifs	177	232
Fitness	159	232
SkincareAddiction	123	331
teenagers	70	180
actuallesbians	69	115

Table 1: Number of posts in non-anxiety forums by each group.

port forums and general forums about neutral topics (such as /r/AskReddit or /r/IamA), Reddit allows for within-person same-site comparisons that would not be possible in most other online anxiety support communities (such as 7 Cups[1] or DailyStrength[2]).

Reddit is a popular news sharing and social media site used by 4-6% of adult internet users (Duggan and Smith, 2013). Its users are approximately 67% male, and 64% of all Reddit users are between 18 and 29, based on recent Pew research (Barthel et al., 2016). Given that concealing negative emotions may be a particular concern among men (Nadeau et al., 2016), and given the relatively low participation of men in most psychology convenience samples, the possibility of oversampling male users may be a benefit rather than a limitation of Reddit analyses. Furthermore, the site's use of upvotes and downvotes (or "karma") tends to discourage most everyday users—that is, people not using dedicated "trolling" accounts—from behaving more antisocially than they would in real life (Barthel et al., 2016; Chen and Wojcik, 2016).

The following study analyzes naturalistic lan-

---

[1] https://www.7cups.com
[2] https://www.dailystrength.org

guage use on Reddit to ask two simple, exploratory research questions: (1) In a within-person analysis, how do individuals use language differently in mental health support forums versus neutral contexts? (2) In a between-person analysis, do anxiety forum members and comparison users who do not belong to anxiety forums talk differently when posting in subreddits that are not explicitly about mental health? We explored both question across all available Linguistic Inquiry and Word Count (LIWC; Pennebaker et al., 2015) categories, with special attention to the categories that have previously served as indicators of anxiety, and commonly used individual words (Coppersmith et al., 2015). Our aim is to produce insights that will be useful in clinical practice, particularly for clinicians interested in monitoring clients between sessions or on an outpatient basis after a health crisis, such as a substance use relapse or suicide attempt.

2   Method

We collected three sets of text from two groups of users. For the anxiety group, we collected the recent activity of members of six anxiety-related subreddits (or forums; /r/Anxiety, /r/HealthAnxiety, /r/PanicAttack, /r/panicdisorder, /r/PanicParty, and /r/socialanxiety). The memberships of these forums vary, with /r/Anxiety and /r/socialanxiety having over 80,000 members, and the rest under 3,000 members. From this sample of anxiety-poster activity, we identified the 20 most common non-anxiety-related forums, then identified a sample of users who posted in those common forums but not in any anxiety-related forums (referred to as comparison users; Table 1).

We collected and processed texts in R (R Core Team, 2018), using the jsonlite (Ooms, 2014) and RedditExtractoR (Rivera, 2015) package. The initial scraping of the anxiety related subreddits resulted in 2,636 replies from 1,423 unique users. From each user's profile, we collected 100 of their most recent replies and dropped anyone with fewer than 50 words in anxiety-related forums. This left 28,154 replies from 1,409 unique users. We then combined the text from each user by context (anxiety versus non-anxiety-related forums) and processed the texts with LIWC (Pennebaker et al., 2015). We also translated texts into a document-term matrix for word-level analyses, which involved some cleaning to better identify

word boundaries, and case standardization. Finally we excluded users with fewer than 50 words in either their non-anxiety set, which resulted in 15,516 replies from 523 unique users.

For the comparison sample, we first collected up to 102 of the most popular threads from each of the 20 non-anxiety-related forums. After excluding content from the anxiety-posting users, this resulted in 139,680 replies from 73,976 unique users. From this potential set of users, we aimed to identify a sample similar to the anxiety-posting users, in terms of their non-anxiety-related forum activity. To do this, we drew random users from the potential set of users one by one; if the user had more than 50 words across their replies, we looked at which non-anxiety-related forums they posted in. If including the user would not increase the Canberra distance[3] between the anxiety posting sample and the comparison by .04[4] or more, they would be added to the comparison set. This was done until there were 523 users included in the comparison group. The resulting sample included 15,102 replies aggregated into 1,046 texts (two per user), which, in terms of percentages of subreddits, is .189 Canberra distance (.989 Pearson's r) from the anxiety-poster sample. The final dataset included 1,569 texts, with 523 from each source (anxiety, non-anxiety, and comparison).

3 Results and Discussion

3.1 Full Sample Analyses

To capture an initial picture of the data, we constructed word clouds based on logistic regressions (calculated for each word, with the word and the percent of each user's political posts predicting each source separately; Figure 1), and fit a decision tree (using the rpart package; Therneau and Atkinson, 2018) to the entire dataset (Figure 2). The decision tree's predictions matched the real sample 68% of the time; that is to say, knowing the values of the anx, shehe, and netspeak LIWC categories, what percentage of words in the text were captured by LIWC (Dic), and the frequency of *that* and *know*, you could use these rules to appropriately categorize the texts 68% of the time, within this sample. Both of these visualizations give

Figure 1: 100 words or marks with the largest log-odds when predicting the given source. Size and color are determined by scaled log-odds.
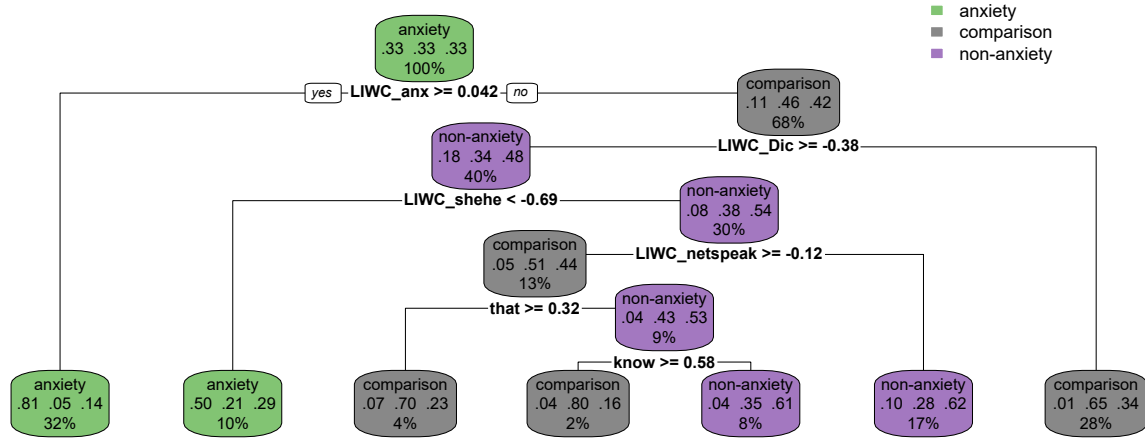
185

Figure 2: Decision tree fit to the full dataset, predicting source sample. Each colorful node is named for the dominant class, with the probability of each class and the percent of total sample underneath. The split values under each node are z-scores.

	anxiety	comparison	non-anxiety
anxiety	.922	.113	.224
comparison	.017	.662	.315
non-anxiety	.061	.226	.461

Table 2: Confusion matrix comparing actual text source to that predicted by the decision tree fit to the entire sample. Cells show probability of predicted source (rows) given actual source (columns), with columns summing to 1.

the similar impression that, at minimum, anxiety-related words characterize the texts from anxiety-related forums. Table 2 breaks down the decision tree's accuracy to make this point again; 92% of the anxiety posts were accurately classified, compared with 66% of the comparison texts, and only 46% of the non-anxiety texts (non-anxiety posts from anxiety posters).

The word clouds in Figure 1 are bound to be somewhat specific to the users that we sampled and may not generalize well to new data; nevertheless, they provide a vivid snapshot of the content of each sample, and some patterns in these word-level correlates fit with past research on anxiety disorders. Figure 1 shows that anxiety users' neutral posts are characterized by references to unpleasant aspects of relationships (*separating, doormat*) or other people (*immaturity, pestering*), counterbalanced to a degree by a few positive affective words (*wellbeing, masterpiece, hugged*). The same group of posts seemed to use more moral words than the comparison or anxiety forum posts, with terms that may reflect concerns about harm (*humane, wronged*), subversion or question-

ing authority (*denies, dissent*), and perhaps unfairness or injustice (*gays, inmates, interracial, greed*; Graham et al., 2009). In contrast, the comparison posts seemed to discuss social injustice in a less personal or more analytic way (*indictment, counsel, Vladimir*).

There were a few commonalities between words used in neutral and anxiety support forums by anxiety forum members. Echoing past findings concerning anxious individuals' greater use of LIWC's health category on Twitter (Coppersmith et al., 2015), health references were more common in anxiety users' posts in both neutral (*nurses, overdosing*) and anxiety forums (*meds, strokes*). References to specific symptoms (*palpitations, hyperventilating*), medications (*propranolol, mirtazapine*), and behavioral coping strategies (*mindfulness, meditation*) were more common in anxiety support forums. Although posts in anxiety forums do refer to anxiety more often and more specifically than the two comparison samples (*panic, nervousness, spiraling*), anxiety users' posts in neutral forums were also characterized by broader negative affective terms, such as *curse* and *bawling*. Finally, anxiety forums, relative to the two comparison samples, used higher rates of psychological terms that are not necessarily unique to the etiology or treatment of anxiety—including *stressor, subconsciously*, and *amygdala*— perhaps reflecting users' research on or knowledge about psychology more broadly.

Category/Word	log-odds	z	p
Intercept	.071	.881	.378
anxiety	.945	3.736	.000
notice	.501	3.094	.002
LIWC: Dic	.366	2.663	.008
parents	.234	2.386	.017
though	.204	2.429	.015
me	.191	2.299	.022
sometimes	.182	2.023	.043
LIWC: health	.156	1.834	.067
LIWC: negemo	.152	1.730	.084
LIWC: home	.113	1.415	.157
was	.107	1.019	.308
and	.103	1.025	.305
LIWC: anx	.064	.608	.543
LIWC: focuspast	.064	.592	.554
depression	.053	.421	.674
LIWC: conj	.019	.175	.861
set	-.064	-.838	.402
america	-.129	-1.505	.132
LIWC: netspeak	-.154	-1.819	.069
LIWC: funct	-.165	-1.211	.226
panic	-.183	-1.999	.046
LIWC: number	-.198	-2.412	.016
asian	-.228	-2.730	.006
russia	-.253	-2.399	.016
guns	-.258	-1.931	.053
punct: qmark	-.277	-3.410	.001
lucky	-.413	-2.423	.015
thank	-.440	-4.208	.000

Table 3: Logistic regression in the between-person analysis predicting user group (anxious or non-anxious) within non-anxiety forums, with positive log-odds associated with anxiety forum users.

Anxiety versus non-anxiety		
	anxiety	non-anxiety
anxiety	.960	.069
non-anxiety	.040	.931
Non-anxiety versus comparison		
	non-anxiety	comparison
non-anxiety	.672	.351
comparison	.328	.649

Table 4: Accuracies of within and between person regression models broken down

3.2 Out-of-Sample Predictions

Next, we explored how predictable the texts' source would be outside of the sample. To do this, we randomly selected 174 users (1/3 of the sample; keeping the number of texts from each source about even) each from the anxiety-posting sample and comparison sample to be held out for testing, then used the remaining sample of 699 users for training. We considered most LIWC categories (excluding percentile and punctuation variables, which were processed manually) and all unique words, making for 12,297 variables. From here, we separated the data into a set only including the anxiety and non-anxiety posts, and a set only including the non-anxiety and comparison posts. For each of these sets, we fit regularized (elastic net, using the glmnet package; Friedman et al., 2010) logistic regressions and decision trees, both predicting each text's source. We considered both of these methods for their potential to reduce the number of variables and thus make the results more interpretable.

Within-Person Comparison. The first sample we tested contained two sets of posts from each user, with the goal of predicting which set of forums the given post was coming from (anxiety-related or non-anxiety-related). To find the optimal penalty parameter (α; affecting the smoothness of weighting) for the model, we tested 5 values from 0 to 1 (considering L_1 and L_2 regularization, and in between). The optimal weighting parameter (λ; affecting the strength of weighting) was selected by cross-validation within the training set. For the reported model, $\alpha = .25$ and $\lambda = .083$.

Regularization left 216 variables with coefficients greater than 0. Among these, positive predictors of anxiety-related forums with the largest coefficients were the dictionary (Dic; % of dictionary words captured) and anxiety (anx) LIWC categories. The positive predictors of non-anxiety forums with the largest coefficients were the male, sexual, and female LIWC categories, and the word *the*. This model accurately classified 94.54% of the test sample texts (Table 4).

In other words, when posting in anxiety forums, people tended to use higher-frequency words and, unsurprisingly, used words related to anxiety (e.g., *scare, worried*) more often. When the same people moved to other non-anxiety-related forums, they discussed men, women, and sex. Whether

187

this pattern represents masking (intentionally imitating Reddit norms in order to appear typical), a type of disengagement coping (avoiding distress through distractions), the anxiety forum members' personalities when they are not feeling anxious, or even the source of users' anxiety itself is unclear based on these data alone (Carver and Connor-Smith, 2010). Fitting a decision tree to the within-person sample with the same outcome yielded similar results (Figure 3), accurately classifying 83.82% of the texts in the test sample.

Between-Person Comparison. The between-person analysis attempted to answer the potentially more challenging question of how to distinguish anxiety forum members' and others' comments in neutral forums. The same sort of regularized model was fit here in the same manner; $\alpha = .25$, $\lambda = .199$. This model accurately classified 66.09% of test sample texts (Table 4). Regularization left 28 variables with coefficients over 0. These are presented in Table 3, which also shows the results of an unregularized logistic regression, including only those variables, and fit to the entire dataset. The decision tree for this set had an out-of-sample accuracy of 60.01%.

Results showed that, relative to the comparison sample (people who were not members of popular anxiety forums), anxiety subreddit members posting in neutral forums used more common words and more conjunctions (Coppersmith et al., 2015), perhaps reflecting a simpler and more conversational (as opposed to analytical) writing style (Pennebaker et al., 2014). Notably, anxiety forum members used more anxious language than others even in neutral forums that were ostensibly irrelevant to mental health. Finally, anxiety forum members showed signs of being less social than others, asking fewer questions (fewer *what*s, fewer question marks) and thanking other posters less often—perhaps reflecting social withdrawal, which has been implicated in both the etiology and maintenance of anxiety disorders (especially social anxiety; Rubin et al., 2009).

Finally, consistent with past findings regarding neuroticism and anxiety, anxiety forum members were more self-focused (more *me*) than comparison users (Tackman et al., 2018). That *me* and not *I* predicted anxiety in this sample could indicate that anxiety users' self-focus specifically takes a more passive or less agentic form, discussing events or actions that that happened to them rather than
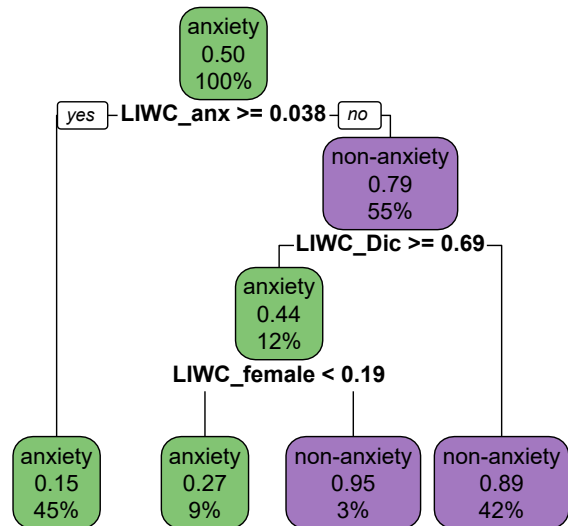

Figure 3: Decision tree fit to the within-person test set contrasting posts in anxiety and non-anxiety forums by the same users. The number at the center of each node is the probability of anxiety posters within that split.
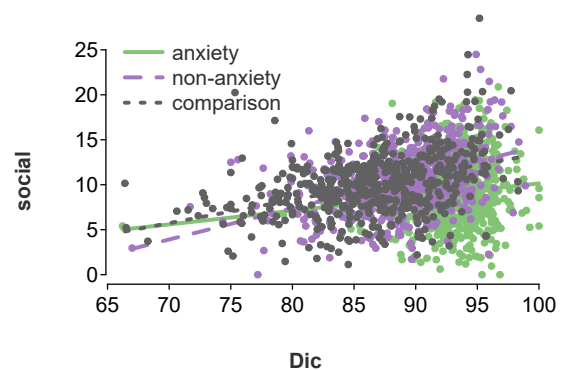

Figure 4: Interaction between the social LIWC category and the percent of words captured by LIWC.

their own actions or thoughts. Recent research has examined psychological differences in subjective and objective first-person singular pronouns (*I* versus *me*, respectively) in depression (Zimmermann et al., 2017), finding that the objective *me* is more indicative of depression than the subjective *I*. Our results suggest that it may be worthwhile to revisit the subjective vs. objective distinction in research on anxiety as well.

Next, we explored the data more visually, with a focus on LIWC variables of interest. For example, Figure 4 shows an interaction between the anxiety and non-anxiety posts. Posts that are particularly well captured by LIWC (Dic) but use very few social words seem to be the main cause of this interaction. Texts fitting this description seem to describe experiences with anxiety and treat-
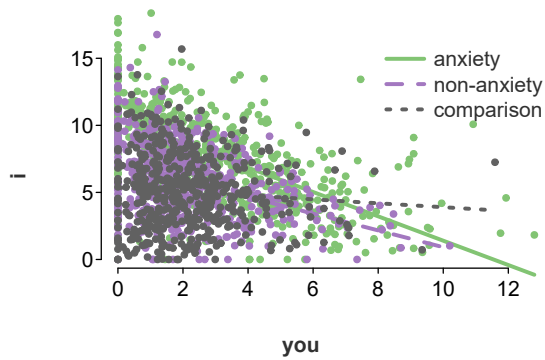
188

Figure 5: Interaction between the i and you LIWC categories.

ment, as in "I don't know why, and I just wake up like that sometimes. So far, breathing exercises just makes it worse, but maybe I'm not doing it right" (r/PanicAttack). Comments in non-anxiety-related forums did not tend to involve this sort of recounting, but it occurred occasionally; for example, "This happened to me when I was in college. I was trying to sleep because I had to be up early the next morning" (r/AskReddit). These are well-captured and low in social language because they are describing individuals' thoughts and experience in a particular moment or sequence.

Finally, past LIWC research has demonstrated the centrality of personal pronouns in understanding how focus on oneself versus others relates to personality and mental health (Tackman et al., 2018). Figure 5 shows the association between I and you within each sample. A negative correlation between first-person singular pronouns (i) and second-person singular pronouns (you) is prominent in the anxiety sample, which appears to be most driven by texts with high *I* use and low *you* use. Considering that approaching personal challenges from a first-person rather than second-person perspective tends to be associated with increased psychological distress, the pronoun usage of people posting in anxiety forums could represent a ruminative or otherwise suboptimal method of seeking and providing support (Dolcos and Albarracin, 2014; Kross and Ayduk, 2011).

4   Future Work and Limitations

The aims of this study were to observe how anxious individuals' language use changes from support-seeking to neutral settings, and investigate whether those same anxiety-subreddit users' language could be differentiated from others' lan-

guage use in neutral forums. As a preliminary proof-of-concept study, the present findings provide a foundation for future work on these topics; however, our approach had several limitations. First, after selecting only the 20 most common neutral subreddits that anxiety community members also posted in, and after excluding users who did not use at least 50 words in each context, the sample for the within-person comparison was relatively small ($N = 523$). Future analyses may hand code all 5,562 subreddits that the users in the original anxiety sample also posted in, providing a more nuanced portrait of how individuals with anxiety post across popular and niche communities.

Partly due to the relatively small sample, we primarily used a dictionary approach to analyzing these texts. Because of their transparency, theory-driven nature, and ease of use, they are more readily disseminated to researchers outside of computational linguistics (such as practicing clinicians) than more mathematically sophisticated or data-driven natural language processing methods (Tauszcik and Pennebaker, 2010). LIWC is also arguably more appropriate than open-vocabulary approaches in smaller samples ($N < 5,000$), where individual words or topics may not occur often enough to be useful predictors (Schwartz et al., 2016, 2017). However, dictionary approaches also have many acknowledged limitations. The LIWC affect categories in particular can be difficult to interpret without significant text cleaning that we did not carry out in this study (e.g., disambiguating uses of *like*) and may not be reliably related to self-reported positive emotions (Sun et al., under review). In dictionaries, it may also be unclear whether the effect of an entire category is being driven by one or a few relatively common words (see Ireland et al., 2015).

In terms of psychological insights offered by this study, a primary concern is whether the individuals in our sample are representative of other individuals with anxiety. People commonly have separate handles for different purposes in order to provide some privacy, or use "throwaway" reddit usernames when they wish to discuss personally identifiable or intimate information on Reddit. It may be relatively rare to share personal details relating to mental health conditions (e.g., describing recent panic attacks at work or childhood physical abuse) and then chat about less intimate topics

(e.g., video games or world news) in other subreddits under the same username. In our sample, only about one-third (37.12%) of the people who used at least 50 words in anxiety subreddits also used at least 50 words in popular neutral or non-anxiety subreddits under the same name. By definition then, the people with sufficient text to analyze in both contexts are atypical, even for members of Reddit anxiety forums. Speculatively, people who are willing to use consistent usernames in support-seeking and neutral contexts may be more extraverted (John and Srivastava, 1999), more verbally disinhibited (Swann Jr and Rentfrow, 2001), or lower in self-monitoring (the tendency to alter one's behavior to fit social expectations; Ickes et al., 1986), relative to an average person—all characteristics that may limit the generalizability of our results. More simply, they could have milder anxiety symptoms (particularly for social anxiety) or better overall mental health than those who post only in mental health forums.

Along the same lines, the six anxiety communities that we sampled from do not provide full coverage of all anxiety disorders; there are also notable differences among the conditions those subreddits represent. Panic disorder, social anxiety, and generalized anxiety are in the same broad category of the Diagnostic and Statistical Manual of Mental Disorders 5 (Anxiety Disorders), but those conditions have key differences in both etiology and treatment (APA, 2013). Future analyses should determine whether changes in language use from support-seeking to neutral contexts are similar across all mental health conditions that relate to the experience of chronic negative affect, including depression, PTSD, and bipolar disorder, among many others. Within anxiety disorders as well, it is unclear whether our results will generalize to communities focusing on more narrowly defined or less common conditions, such as specific phobias or agoraphobia.

Finally, by collapsing across posts, we sacrificed granularity for parsimony. That is, for the moment, we intentionally ignored a wealth of potentially useful information about specific subreddits, time, upvotes, and thread structure. There is clearly much more to be explored, particularly in terms of social and temporal dynamics (see Coppersmith et al., 2016). For example, due to social anxiety or simply the cognitive burden of inhibiting negative emotions, anxiety users may be less

socially engaged—and therefore receive fewer upvotes and replies—relative to controls when they post in neutral communities. They also may post more slowly, less often, or in atypical temporal patterns, relative to less anxious Reddit users (Loveys et al., 2017).

5 Conclusion

Two sets of analyses explored how individuals' language use changes from support-seeking to neutral settings, and further demonstrated that anxious individuals' language use can be differentiated from comparison posts even in neutral settings, when the topics of conversation rarely focus on mental health. Results revealed not only face-valid content differences (e.g., in references to anxiety, negative affect, and social language), but also subtler stylistic differences (e.g., in self-focus, conjunctions, word frequency, and questions). Findings were largely consistent with past research and existing theory (Coppersmith et al., 2015; Mehl et al., 2012; Tackman et al., 2018), while also suggesting novel data-driven hypotheses to be tested in future research.

We are particularly encouraged by some of the unexpected results (for example, regarding question marks and *thanks*) that, despite not being directly predicted by past work, are nevertheless consistent with research and theory on the nature of anxiety. In terms of informing future behavior change interventions, it may be especially valuable to identify behavior patterns in neutral settings that maintain or exacerbate anxiety—for example, being less interactive or positive even when ostensibly engaging in prosocial behavior, such as posting in discussion forums.

Information about the communication context is typically unavailable in large-scale social media classification tasks; however, clinicians or medical practitioners often operate at the level of individual clients. In cases with abundant information about the person and the context—for example, when reviewing chat messages from online outpatient therapy sessions (Wolf et al., 2010) or analyzing clients' social media messages between health center visits (Padrez et al., 2015)—appreciating how aspects of the situation influence the linguistic signal of psychological distress may prove to have near-future applied value.

References

APA. 2013. *Diagnostic and statistical manual of mental disorders-5*. American Psychiatric Association.

Michael Barthel, Galen Stocking, Jesse Holcomb, and Amy Mitchell. 2016. Nearly eight-in-ten reddit users get news on the site. *Pew Research Center*, 25.

Adrian Benton, Margaret Mitchell, and Dirk Hovy. 2017. Multitask learning for mental health conditions with limited social media data. In *Proceedings of the 15th Conference of the European Chapter of the Association for Computational Linguistics: Volume 1, Long Papers*, volume 1, pages 152–162.

Jack Block. 2010. The five-factor framing of personality and beyond: Some ruminations. *Psychological Inquiry*, 21(1):2–25.

Charles S Carver and Jennifer Connor-Smith. 2010. Personality and coping. *Annual review of psychology*, 61:679–704.

Eric Evan Chen and Sean P Wojcik. 2016. A practical guide to big data research in psychology. *Psychological methods*, 21(4):458.

Mike Conway and Daniel OConnor. 2016. Social media, big data, and mental health: current advances and ethical implications. *Current opinion in psychology*, 9:77–82.

Glen Coppersmith, Mark Dredze, and Craig Harman. 2014. Quantifying mental health signals in twitter. In *Proceedings of the Workshop on Computational Linguistics and Clinical Psychology: From Linguistic Signal to Clinical Reality*, pages 51–60.

Glen Coppersmith, Mark Dredze, Craig Harman, and Kristy Hollingshead. 2015. From adhd to sad: Analyzing the language of mental health on twitter through self-reported diagnoses. In *Proceedings of the 2nd Workshop on Computational Linguistics and Clinical Psychology: From Linguistic Signal to Clinical Reality*, pages 1–10.

Glen Coppersmith, Kim Ngo, Ryan Leary, and Anthony Wood. 2016. Exploratory analysis of social media prior to a suicide attempt. In *Proceedings of the Third Workshop on Computational Linguistics and Clinical Psychology*, pages 106–117.

Kathryn P Davison, James W Pennebaker, and Sally S Dickerson. 2000. Who talks? the social psychology of illness support groups. *American Psychologist*, 55(2):205.

Munmun De Choudhury, Michael Gamon, Scott Counts, and Eric Horvitz. 2013. Predicting depression via social media. *ICWSM*, 13:1–10.

Munmun De Choudhury, Emre Kiciman, Mark Dredze, Glen Coppersmith, and Mrinal Kumar. 2016. Discovering shifts to suicidal ideation from mental health content in social media. In *Proceedings of the 2016 CHI conference on human factors in computing systems*, pages 2098–2110. ACM.

Sanda Dolcos and Dolores Albarracin. 2014. The inner speech of behavioral regulation: Intentions and task performance strengthen when you talk to yourself as a you. *European Journal of Social Psychology*, 44(6):636–642.

Maeve Duggan and Aaron Smith. 2013. 6% of online adults are reddit users. *Pew Internet & American Life Project*, 3:1–10.

Meir Friedenberg, Hadi Amiri, Hal Daumé III, and Philip Resnik. 2016. The umd clpsych 2016 shared task system: Text representation for predicting triage of forum posts about mental health. In *Proceedings of the Third Workshop on Computational Lingusitics and Clinical Psychology*, pages 158–161.

Jerome Friedman, Trevor Hastie, and Robert Tibshirani. 2010. Regularization paths for generalized linear models via coordinate descent. *Journal of Statistical Software*, 33(1):1–22.

Rula Bayrakdar Garside and Bonnie Klimes-Dougan. 2002. Socialization of discrete negative emotions: Gender differences and links with psychological distress. *Sex roles*, 47(3-4):115–128.

George Gkotsis, Anika Oellrich, Sumithra Velupillai, Maria Liakata, Tim JP Hubbard, Richard JB Dobson, and Rina Dutta. 2017. Characterisation of mental health conditions in social media using informed deep learning. *Scientific reports*, 7:45141.

Jesse Graham, Jonathan Haidt, and Brian A Nosek. 2009. Liberals and conservatives rely on different sets of moral foundations. *Journal of personality and social psychology*, 96(5):1029.

Shannon E Holleran and Matthias R Mehl. 2008. Let me read your mind: Personality judgments based on a persons natural stream of thought. *Journal of Research in Personality*, 42(3):747–754.

William Ickes, Susan Reidhead, and Miles Patterson. 1986. Machiavellianism and self-monitoring: As different as" me" and" you". *Social Cognition*, 4(1):58.

Molly E Ireland and Matthias R Mehl. 2014. Natural language use as a marker. *The Oxford handbook of language and social psychology*, pages 201–237.

Molly E Ireland, H Andrew Schwartz, Qijia Chen, Lyle H Ungar, and Dolores Albarracín. 2015. Future-oriented tweets predict lower county-level hiv prevalence in the united states. *Health Psychology*, 34(S):1252.

William Jarrold, Harold S Javitz, Ruth Krasnow, Bart Peintner, Eric Yeh, Gary E Swan, and Matthias Mehl. 2011. Depression and self-focused language in structured interviews with older men. *Psychological reports*, 109(2):686–700.

Oliver P John and Sanjay Srivastava. 1999. The big five trait taxonomy: History, measurement, and theoretical perspectives. *Handbook of personality: Theory and research*, 2(1999):102–138.

Arif Khan, Robyn M Leventhal, Shirin Khan, and Walter A Brown. 2002. Suicide risk in patients with anxiety disorders: a meta-analysis of the fda database. *Journal of affective disorders*, 68(2):183–190.

Ethan Kross and Ozlem Ayduk. 2011. Making meaning out of negative experiences by self-distancing. *Current directions in psychological science*, 20(3):187–191.

Kate Loveys, Patrick Crutchley, Emily Wyatt, and Glen Coppersmith. 2017. Small but mighty: Affective micropatterns for quantifying mental health from social media language. In *Proceedings of the Fourth Workshop on Computational Linguistics and Clinical Psychology—From Linguistic Signal to Clinical Reality*, pages 85–95.

Matthias R Mehl, Samuel D Gosling, and James W Pennebaker. 2006. Personality in its natural habitat: Manifestations and implicit folk theories of personality in daily life. *Journal of personality and social psychology*, 90(5):862.

Matthias 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.

Miranda M Nadeau, Michael J Balsan, and Aaron B Rochlen. 2016. Mens depression: Endorsed experiences and expressions. *Psychology of Men & Masculinity*, 17(4):328.

Jeroen Ooms. 2014. The jsonlite package: A practical and consistent mapping between json data and r objects. *arXiv:1403.2805 [stat.CO]*.

Kevin A Padrez, Lyle Ungar, Hansen Andrew Schwartz, Robert J Smith, Shawndra Hill, Tadas Antanavicius, Dana M Brown, Patrick Crutchley, David A Asch, and Raina M Merchant. 2015. Linking social media and medical record data: a study of adults presenting to an academic, urban emergency department. *BMJ Qual Saf*, pages bmjqs–2015.

James W. Pennebaker, Ryan L. Boyd, Kayla Jordan, and Kate Blackburn. 2015. The development and psychometric properties of liwc2015. *UT Faculty/Researcher Works*.

James W Pennebaker, Cindy K Chung, Joey Frazee, Gary M Lavergne, and David I Beaver. 2014. When small words foretell academic success: The case of college admissions essays. *PloS one*, 9(12):e115844.

R Core Team. 2018. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. Version 3.4.4.

Philip Resnik, William Armstrong, Leonardo Claudino, Thang Nguyen, Viet-An Nguyen, and Jordan Boyd-Graber. 2015. Beyond lda: exploring supervised topic modeling for depression-related language in twitter. In *Proceedings of the 2nd Workshop on Computational Linguistics and Clinical Psychology: From Linguistic Signal to Clinical Reality*, pages 99–107.

Philip Resnik, Anderson Garron, and Rebecca Resnik. 2013. Using topic modeling to improve prediction of neuroticism and depression in college students. In *Proceedings of the 2013 conference on empirical methods in natural language processing*, pages 1348–1353.

Ivan Rivera. 2015. *RedditExtractoR: Reddit Data Extraction Toolkit*. R package version 2.0.2.

Kenneth H Rubin, Robert J Coplan, and Julie C Bowker. 2009. Social withdrawal in childhood. *Annual review of psychology*, 60:141–171.

Norman Sartorius, T Bedirhan Üstün, Yves Lecrubier, and Hans-Ulrich Wittchen. 1996. Depression comorbid with anxiety: Results from the who study on" psychological disorders in primary health care.". *The British journal of psychiatry*.

David P Schmitt, Anu Realo, Martin Voracek, and Jüri Allik. 2008. Why can't a man be more like a woman? sex differences in big five personality traits across 55 cultures. *Journal of personality and social psychology*, 94(1):168.

H Andrew Schwartz, Johannes C Eichstaedt, Margaret L Kern, Lukasz Dziurzynski, Stephanie M Ramones, Megha Agrawal, Achal Shah, Michal Kosinski, David Stillwell, Martin EP Seligman, et al. 2013. Personality, gender, and age in the language of social media: The open-vocabulary approach. *PloS one*, 8(9):e73791.

H Andrew Schwartz, Salvatore Giorgi, Maarten Sap, Patrick Crutchley, Lyle Ungar, and Johannes Eichstaedt. 2017. Dlatk: Differential language analysis toolkit. In *Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing: System Demonstrations*, pages 55–60.

H Andrew Schwartz, Maarten Sap, Margaret L Kern, Johannes C Eichstaedt, Adam Kapelner, Megha Agrawal, Eduardo Blanco, Lukasz Dziurzynski, Gregory Park, David Stillwell, et al. 2016. Predicting individual well-being through the language of social media. In *Biocomputing 2016: Proceedings of the Pacific Symposium*, pages 516–527. World Scientific.

Robert J Smith, Patrick Crutchley, H Andrew Schwartz, Lyle Ungar, Frances Shofer, Kevin A Padrez, and Raina M Merchant. 2017. Variations in facebook posting patterns across validated patient health conditions: A prospective cohort study. *Journal of medical Internet research*, 19(1).

Jessie Sun, H. Andrew Schwartz, Youngseo Son, Margaret L. Kern, and Simine Vazire. under review. The language of well-being: Tracking within-person emotion fluctuations through everyday speech.

William B Swann Jr and Peter J Rentfrow. 2001. Blirtatiousness: cognitive, behavioral, and physiological consequences of rapid responding. *Journal of Personality and Social Psychology*, 81(6):1160.

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. 2018. Depression, negative emotionality, and self-referential language: A multi-lab, multi-measure, and multi-language-task research synthesis. *Journal of personality and social psychology*.

Yla R Tausczik and James W Pennebaker. 2010. The psychological meaning of words: Liwc and computerized text analysis methods. *Journal of language and social psychology*, 29(1):24–54.

Shelley E Taylor, David K Sherman, Heejung S Kim, Johanna Jarcho, Kaori Takagi, and Melissa S Dunagan. 2004. Culture and social support: who seeks it and why? *Journal of personality and social psychology*, 87(3):354.

Terry Therneau and Beth Atkinson. 2018. *rpart: Recursive Partitioning and Regression Trees*. R package version 4.1-13.

Larissa Z Tiedens. 2001. Anger and advancement versus sadness and subjugation: the effect of negative emotion expressions on social status conferral. *Journal of personality and social psychology*, 80(1):86.

Konstantin O Tskhay and Nicholas O Rule. 2014. Perceptions of personality in text-based media and osn: A meta-analysis. *Journal of Research in Personality*, 49:25–30.

Harvey A Whiteford, Louisa Degenhardt, Jürgen Rehm, Amanda J Baxter, Alize J Ferrari, Holly E Erskine, Fiona J Charlson, Rosana E Norman, Abraham D Flaxman, Nicole Johns, et al. 2013. Global burden of disease attributable to mental and substance use disorders: findings from the global burden of disease study 2010. *The Lancet*, 382(9904):1575–1586.

Markus Wolf, Cindy K Chung, and Hans Kordy. 2010. Inpatient treatment to online aftercare: E-mailing themes as a function of therapeutic outcomes. *Psychotherapy Research*, 20(1):71–85.

Carolyn Zahn-Waxler, Bonnie Klimes-Dougan, and Marcia J Slattery. 2000. Internalizing problems of childhood and adolescence: Prospects, pitfalls, and progress in understanding the development of anxiety and depression. *Development and psychopathology*, 12(3):443–466.

Johannes Zimmermann, Timo Brockmeyer, Matthias Hunn, Henning Schauenburg, and Markus Wolf. 2017. First-person pronoun use in spoken language as a predictor of future depressive symptoms: Preliminary evidence from a clinical sample of depressed patients. *Clinical psychology & psychotherapy*, 24(2):384–391.