Distribution Agreement

In presenting this thesis or dissertation as a partial fulfillment of the requirements for an advanced degree from Emory University, I hereby grant to Emory University and its agents the non-exclusive license to archive, make accessible, and display my thesis or dissertation in whole or in part in all forms of media, now or hereafter known, including display on the world wide web. I understand that I may select some access restrictions as part of the online submission of this thesis or dissertation. I retain all ownership rights to the copyright of the thesis or dissertation. I also retain the right to use in future works (such as articles or books) all or part of this thesis or dissertation.

Signature:

_____ _____

Victoria Lawlor Date

The phenomenology of psychopathology in online communities.

By

Victoria Lawlor

Master of Arts

Psychology

Michael Treadway

Advisor

Edward Craighead

Committee Member

Phillip Wolff

Committee Member

Accepted:

_____

Lisa A. Tedesco, Ph.D.

Dean of the James T. Laney School of Graduate Studies

_____

Date

The phenomenology of psychopathology in online communities.

By

Victoria Lawlor

B.S., Northeastern University, 2017

Advisor: Michael T. Treadway, Ph.D.

An abstract of

A thesis submitted to the Faculty of the

James T. Laney School of Graduate Studies of Emory University

in partial fulfillment of the requirements for the degree of

Master of Arts

in Psychology

2021

Abstract

The phenomenology of psychopathology in online communities.

By Victoria Lawlor

Current DSM-based criteria for mental disorders boast high reliability, however, widespread comorbidity, symptom overlap between disorders, and heterogeneity within disorders point to the limited validity of this classification system. Advancement in this domain is, in part, hindered by reification of these diagnostic criteria. A returned focus on clinical phenomenology may offset this and guide identification of alternative classification criteria. The popularity of online mental health communities and emergence of natural language processing methods provide opportunities to characterize clinical phenomena in large samples of text written by individuals endorsing psychopathology. In the present study, we examined free-form, unprompted text posts from six online mental health communities ("subreddits") from Reddit, a popular discussion-based website. To characterize discussions within these communities we quantified semantic similarity between their content, created topic models to describe their thematic elements, and examined membership patterns between communities. We found high overall semantic similarity among mental health subreddits. Interpretable, face-valid themes also emerged from the topic modeling analyses. These themes helped contextualize variability in semantic similarity and were moderately related to known clinical phenomenology. Membership patterns observed in these communities showed little correspondence to established nosology. Taken together, these findings help identify the symptoms and concerns most salient to members of online mental health communities and demonstrate the utility and feasibility of using these samples to characterize psychopathology.

The phenomenology of psychopathology in online communities.

By

Victoria Lawlor

B.S., Northeastern University, 2017

Advisor: Michael T. Treadway, Ph.D.

A thesis submitted to the Faculty of the

James T. Laney School of Graduate Studies of Emory University

in partial fulfillment of the requirements for the degree of

Master of Arts

in Psychology

2021

Table of Contents

Introduction

The introduction of the third edition of the Diagnostic and Statistical Manual of Mental Disorders (DSM-III; APA, 1980) marked a turning point in the field of psychiatry that gave rise to our current approach to diagnosis of psychopathology. Influenced by mounting skepticism toward the legitimacy of psychiatry, the DSM-III formalized diagnostic criteria for each disorder—medicalizing American psychiatry and establishing mental disorders as disease categories (Fischer, 2012; Kawa & Giordano, 2012). The benefits of this shift were far reaching. In particular, it provided clinicians and researchers a shared language to describe clinical phenomena, helping to pave the way for new epidemiological and psychopharmacological research (Kawa & Giordano, 2012). Additionally, it increased diagnostic reliability by establishing precise symptom criteria for each disorder (APA, 1980). In comparison to its predecessors, DSM-III categories were grounded to empirical data rather than clinical judgment, which helped to legitimize the field of psychiatry.

Despite these acknowledged benefits, this change also resulted in a significant unintended consequence: the reification of DSM-based diagnostic criteria and the assumption that diagnostic categories could accurately differentiate distinct "natural kinds" of psychiatric disease (Hyman, 2010). This assumption has been widely criticized in light of widespread diagnostic comorbidity, significant overlap among symptom criteria, and substantial heterogeneity within diagnostic categories—all of which point to limited validity of DSM-based classification (Plana-Ripoll et al., 2019, Zbozinek et al., 2012, Wardenaar & de Jong, 2013). The failure to acknowledge the distinction between clinical features of a disorder and its diagnostic criteria poses a number of problems. In clinical contexts, an overemphasis on diagnostic criteria may result in oversimplistic conceptualizations devoid of information that may not neatly map onto a DSM-based disorder. On a broader level, reification may impede advancements in classification by

asserting the incorrect notion that the "true" disorders have already been identified (Hyman, 2010).

Dissatisfaction with the current conceptualization of mental disorders has prompted research initiatives like the Research Domain Criteria (RDoC) project aimed at identifying components of psychopathology, and motivated the development of an alternative classification system, the hierarchical taxonomy of psychopathology (HiTOP; Kotov et al., 2017). However, breaking away from DSM criteria and categories remains a challenge in light of how entrenched they have become in conceptualizations of psychopathology. Clinical measures used in research are largely based on symptoms put forth in the DSM—limiting their utility for forming new classification systems. Similarly, research in clinical samples predominantly uses the DSM to determine eligibility, which imparts the diagnostic criteria onto any data by constructing a sample only of patients whose symptoms already conform to DSM disorders (Guloksuz, Os, & Pries, 2017).

Returning to clinical phenomenology may help to offset the effects of reification on advancing psychopathology classification. Specifically, meaningful symptoms and constructs that were initially overlooked or excluded from diagnostic criteria may emerge from focusing directly on experienced psychiatric phenomena (Kendler & Parnas, 2015). Supporting this idea, Kendler (2016a) reviewed textbooks with the aim of characterizing historical descriptions of depression and identified five signs and symptoms that are not captured by DSM criteria. A similar approach using textbooks and review articles that described schizophrenia revealed eleven signs and symptoms not found in any diagnostic system (Kendler, 2016b). These examples illustrate how a greater understanding of clinical phenomenology may spur the identification of promising new diagnostic criteria.

Online communities show excellent potential as sources of rich, first-hand descriptions of clinical phenomena—particularly the popular discussion-based website Reddit, which draws over 430 million users each month (Reddit, 2019). Reddit is comprised of subreddits, or forums for particular topics, in which users can create posts or contribute to ongoing discussions. Notably, there are a number of subreddits devoted to mental health, many of which correspond to DSM-based disorders. These subreddits are promising sources of data for several reasons. First, they provide large amounts of data going back several years; this public data can be readily collected using established tools. The freeform, unprompted nature of subreddit posts may also help to reveal the thoughts and experiences that are most salient to patients themselves. Additionally, subreddits are anonymous, which may encourage users to post their unfiltered thoughts with minimal impression management. Lastly, users self-select into these communities—so, while the creation of subreddits is likely influenced by DSM diagnostic categories, users may gravitate toward communities that best reflect their lived experiences.

Natural language processing, a subfield of computer science, is well suited for extracting clinical phenomena from these large, text-based datasets. Recent work using methods from natural language processing to analyze social media data—including Reddit communities—demonstrates the utility of this approach. Classifiers successfully trained to predict MBTI personality types, the expression of anxiety in individual posts, symptoms of depression, and annotated categories from Reddit posts illustrate the wealth of information contained in these naturalistic datasets (Gjurkovic & Snajder, 2018; Shen & Rudzics, 2017; Tadesse, Lin, Xu, and Yang, 2019; Gaur et al., 2019). Additionally, topic modeling has successfully been used to summarize themes in both mental health and non-clinical communities (Yoo, Lee, & Ha, 2019; Curiskis, Drake, Osborn, Kennedy, 2020; Park, Conway, Chen 2018). Unsupervised methods

like topic modeling may be particularly well suited for identifying clinical phenomenology, as they can detect patterns in large, complex datasets in a bottom-up manner. Interestingly, a majority of studies that have looked at mental health topics in social media are found in the computer science literature. In many of these studies, the mental health data appear to be used primarily as a means of proposing and testing new modeling or visualization techniques, as opposed to extracting clinical insights. Accordingly, Chancellor, Baumer, and De Choudhury (2019) note the need for integration between these two fields, asserting that incorporating domain knowledge into these analyses will result in more meaningful conclusions.

In the present study, we examined posts from mental health subreddits to characterize the experiences of Reddit users who identify with a form of psychopathology. We first quantified similarities in discussion content among mental health subreddits and between mental health and control subreddits. Given symptom overlap and comorbidity rates between disorders, we expected to see higher similarity among mental health subreddits than similarity between mental health and control subreddits (Kessler et al., 2012; Zbozinek 2013). We next used topic modeling to identified themes within each mental health subreddit to a) summarize discussions in a digestible manner, b) contextualize similarity findings by locating thematic overlap between subreddits, and c) compare these user-generated topics to DSM criteria. Lastly, we explored user membership patterns within these data—specifically, the number of members in each subreddit and cross-subreddit overlap—to compare with prevalence and comorbidity rates found in the general population.

Methods and Materials

Data Collection

Data were sourced from subreddits on the popular discussion-based website Reddit; individual subreddits are denoted "r/subreddit-name." There are over 138,000 active subreddits (Marotti, 2018) dedicated to a vast array of topics, including specific hobbies (e.g. r/baking), locations (e.g. r/Atlanta), skill development (e.g. r/coding), humor (e.g. r/memes), sports (e.g. r/tennis), and mental health (e.g. r/depression). Sample posts from three mental health subreddits are provide in Table 1.

Table 1. Sample Reddit Posts

Subreddit	Representative Post
r/depression	**This "happy" mask I'm forced to wear is getting awfully heavy…** I've got too much to lose to let it slip off and for people to see how seriously depressive I've been and suicidal I'm becoming, but it's getting really heavy to keep up all day.
r/anxiety	**I'm tired of replaying the same thoughts in my head.** Whenever someone says something to me that I didn't like or when a conversation doesn't go as planned I replay it in my head for hours or days and it gets so exhausting. How do you just forget things and shut your brain down
r/ptsd	**Unnoticed adrenaline?** Does anyone else randomly feel their entire body go weak from exhaustion and not even realize they were pumping adrenaline?

Representative posts from selected sample mental health subreddits.

The dataset from the present study was generated by collecting posts from seven subreddits that were based on topics corresponded to DSM-based mood and anxiety disorders: r/anxiety, r/depression, r/ocd, r/panicattack, r/panicdisorder, r/ptsd, r/socialanxiety. Posts from r/panicattack and r/panicdisorder were aggregated to form a single "Panic Disorder" community. We collected posts that were written during the one-year period between March 1, 2019 to March 1, 2020 using the Pushshift Reddit API (Baumgartner et al., 2020) and custom Python scripts.

To serve as a baseline for select analyses, we also collected posts from six subreddits not associated with mental health that were written during the same one-year period: r/frugal, r/youshouldknow, r/books, r/fitness, r/teaching, and r/personalfinance. These subreddits, selected from among the control subreddits used by Hanwen-Shen & Rudzics, 2017, are structurally similar to the clinical subreddits (both primarily involve first-person narratives) despite their different topic focus.

As a Reddit post gets more "upvotes" from users, it increases in visibility on r/all, the homepage of Reddit made up of the most popular posts from all subreddits; this can cause the comment sections of some posts to be filled with users who are not members of the subreddit to which the post was initially submitted to. We therefore collected only text from primary posts (i.e. did not include text from the comment section) for clinical and control subreddits.

Data Preparation

Raw posts were cleaned using several standard preprocessing steps prior to analysis. Reddit users can format posts using a Reddit-specific Markdown syntax (e.g. using '__text__' would display 'text' written in bold), resulting in additional characters in the raw data; we removed these characters used for formatting using the redditcleaner Python package (Version 1.1.2) and custom Python scripts. To help standardize the text, we converted it to lowercase and removed punctuation and digits.

To reduce noise in the data and improve the runtime of analyses, we eliminated words expected to provide little to no semantic information meaningful to our analyses—we removed highly common terms referred to as "stop words" (e.g. "the," "that," and "a") and very short words (≤ 3 letters). Posts that contained fewer than ten words after this cleaning process as well

as non-text posts (e.g. an uploaded media file) were removed. Table 2 shows the number of posts per subreddit before and after this cleaning process. For the topic modeling analyses, we additionally converted words to their canonical form (e.g. "shopping" was converted to "shop") using the *lemma* function from the WordNet module of NLTK (Natural Language Toolkit; Bird, Klein, Loper, 2009, Miller, 1995).

Table 2. Summary of Clinical Subreddit Data

Subreddit	Proxy dx	Subscribers[a]	Total posts[b]	Used posts
r/anxiety	GAD	333,162	81,345	74,926
r/depression	MDD	608,903	256,843	232,281
r/ocd	OCD	69,945	27,793	23,569
r/panicattack	Panic Disorder	8,430	2,814	2,496
r/panicdisorder	Panic Disorder	2,046	802	705
r/ptsd	PTSD	41,174	10,101	9,763
r/socialanxiety	Social Anxiety	237,155	28,303	21,605

GAD = Generalized Anxiety Disorder, MDD = Major Depressive Disorder, OCD = Obsessive Compulsive Disorder, PTSD = Post Traumatic Stress Disorder
[a] N subscribers as of 3/1/2020
[b] Posts from 3/1/2019–3/1/2020

Similarity Analysis

While each subreddit is centered around a particular disorder, users in these different communities might be discussing comparable concerns due to diagnostic comorbidity, common symptoms within diagnostic categories, and shared impact of disorders on quality of life. Natural language processing provides a number of methods for determining the semantic similarity—a metric representing overlapping meaning—between text documents. To examine the relationships between clinical communities and determine which communities share the greatest overlap in subject matter, we quantified the semantic similarity between each pair of clinical

communities and compared this to similarity between clinical and non-mental health communities.

Text representation. In order to quantify similarity between communities, text must first be transformed to allow for manipulation and computation. Embeddings are representations of text as numeric vectors that are semantically meaningful. A number of algorithms have been created to map words into vector space while preserving semantic information—vectors of words similar in meaning are closer, and even complex relationships between words are maintained (e.g. the embedding for 'queen' is a combination of the embeddings for 'royalty' and 'woman'). By using these rich representations of text as our foundation, we can compare the overlap in the *meaning* of subreddit content rather than simply the overlap of identical word use. We used 300-dimension GloVe embeddings (Global Vectors for Word Representation; Pennington, Socher, & Manning, 2014) pretrained on a large news corpus, and averaged embeddings for words in a post, resulting in vectors that represented each post. The subreddit-level embeddings were formed by taking the average of each subreddit's constituent post vectors, as in Khan (2019).

Computing similarity. We then used these embeddings to compute semantic similarity, measured by the cosine similarity between all pairs of clinical communities. As such, similarity between vector representations of two subreddits A and B was computed as: $\cos(\theta) = \frac{A \bullet B}{||A||||B||}$

Semantic similarity between subreddits was high, likely stemming from the similar function of posts (e.g. ask for advice, share a story, reflect on the day), as well as from the similar structure of posts (i.e. first-person narrative form). Consequently, for greater ease of interpretation we computed a baseline similarity score using the average similarity between every disorder and the

six control subreddits. Similarity scores for pairs of subreddits are reported as the percent difference from this baseline.

It is plausible that overlap in content may in part be driven by overlap in community members, who may be submitting similar posts to multiple subreddits. Semantic similarity could, in turn, be overestimated for subreddits that boast a high number of intersecting users. To control for this potential confound, we repeated these analyses after excluding posts from users who have posted in more than one community.

Topic Modeling

We next sought to characterize the themes that emerged in each clinical community in a quantitative manner using topic modeling. Topic modeling is an approach from natural language processing that extracts latent semantic structures (i.e. topics) from text. The resulting topics can serve as digestible summaries of large text data and provide insight into their thematic content. To identify topics discussed by each clinical community, we used Latent Dirichlet Allocation (Blei et al., 2003), a widely used probabilistic topic-modeling method. With this method, each post is viewed as a mixture of topics and topics are probability distributions over words. The topics that emerge are collections of words that frequently co-occur in similar contexts.

To perform the LDA, we used the MALLET (Machine Learning for Language Toolkit; McCallum, 2002) model via a Gensim Python wrapper. As LDA requires the number of topics (k) to be explicitly provided, we tested multiple models (with topics ks ranging from 5 to 15; hyperparameters were determined automatically using hyperparameter optimization function in MALLET) for each clinical community. The final models were determined through examination of coherence scores (computed using Gensim) for each cluster n and visual inspection of cluster

interpretability using the LDAvis Python package. The word sets for the identified topics from the final models were manually assigned a cluster "label" based on their perceived themes (e.g. "Social," "Workplace," "Treatment").

LDA models can be used to estimate the proportions of each topic in new, unseen text. This enables us to examine shared themes among communities, which will also help to contextualize the semantic similarity results. Using the final LDA model for a given community, we estimated the proportion of its topics in posts from the remaining communities. We then averaged across posts in each community to compute the average proportion of each topic by community. Given the large number of total topics from all communities combined, we manually selected topics from each community to serve as our primary focus. We selected a mixture of topics related to clinical symptoms (e.g. intrusive thoughts, physical manifestations of anxiety) and daily life (e.g. pastimes, employment, school). Among clinical symptoms, we selected some topics that reflect symptoms that show specificity to a given disorder and others that we would expect to find transdiagnostically.

User analysis

We were interested in exploring whether high comorbidity rates between DSM-based disorders were reflected in the overlap of community members. As Reddit users can post to any number of subreddits, we can see where these communities intersect most frequently based on posting history. To this end, we created sets of members for each clinical community. We considered a user a "member" of a community if they posted in it two or more times within the one-year span of collected data. To quantify overlapping membership between pairs of communities, we used the Jaccard similarity coefficient, measured as the number of common members over total unique members (i.e. the set intersection divided by the union). We

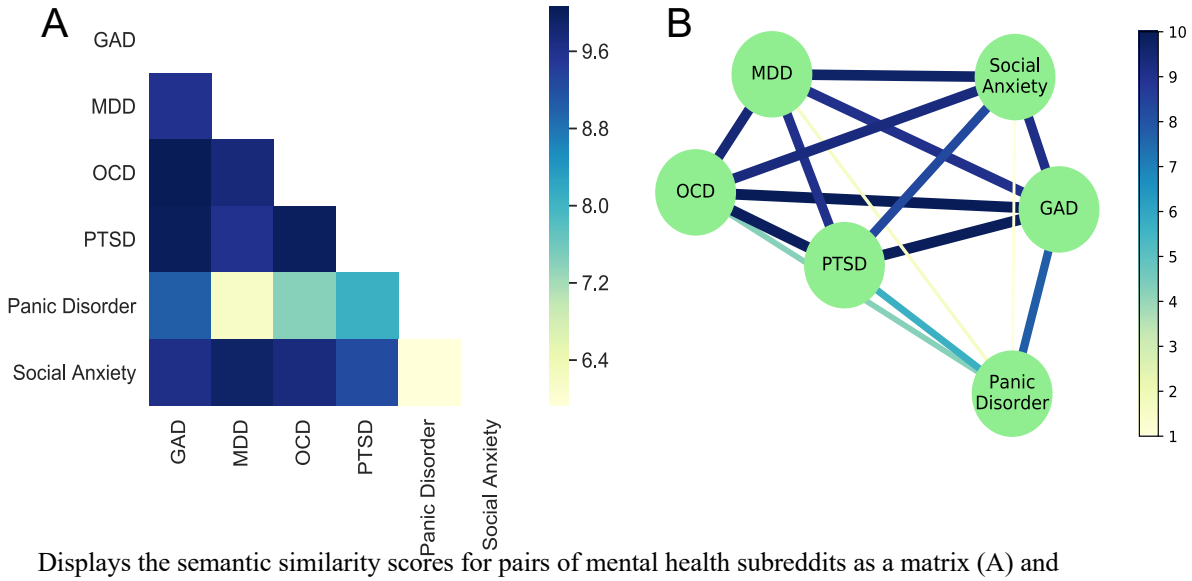
computed scores for each pair of clinical communities, as well as all clinical communities compared to the six control subreddits to serve as a baseline. We also visually inspected the overlap in communities to identify more complex patterns (i.e. user overlap between three or more communities).

Results

Semantic Similarity

Semantic similarity scores between all pairs of clinical communities are presented in Table 3 and Figure 1. As expected, semantic similarity for all pairs of clinical communities was higher compared to the control similarity scores—suggesting that clinical communities discussed similar content relative to non-mental health communities. Specifically, scores ranged from 5.93% (Panic Disorder & Social Anxiety) to 10.06% (GAD & OCD) higher than the average control similarity. The GAD community had relatively high similarity to the other clinical communities, averaging 9.68% higher than baseline. This was followed by PTSD and OCD (both 9.40%), MDD (9.02%), Social Anxiety (8.90%), and Panic Disorder (7.33%) communities.

Figure 1. Semantic Similarity



Displays the semantic similarity scores for pairs of mental health subreddits as a matrix (A) and network (B). Color represents the % change from baseline (i.e. control similarity scores).

Table 3. Semantic similarity

Community	GAD	MDD	OCD	Panic Disorder	PTSD	Social Anxiety
All posts						
GAD	—					
MDD	9.61	—				
OCD	10.06	9.79	—			
Panic Disorder	9.03	6.16	7.43	—		
PTSD	10.04	9.61	10.01	8.08	—	
Social Anxiety	9.65	9.92	9.73	5.93	9.27	—
Restricted to unique posters[a]						
GAD	—					
MDD	9.41	—				
OCD	9.90	9.69	—			
Panic Disorder	9.10	6.23	7.39	—		
PTSD	9.95	9.55	9.95	8.12	—	
Social Anxiety	9.49	9.80	9.80	6.02	9.22	—

[a]Excludes posts from users who have posted in more than one clinical community
Scores are displayed as % change from baseline (i.e. control) similarity scores

In our data, 23.7% of posts were written by users who had posted to multiple clinical communities. As shown in Table 3, similarity patterns between subreddits remained even after removing these posts—suggesting the variance in semantic similarity is not solely not a confound of the users posting similar content in multiple communities.

Topic modeling

The LDA was able to characterize themes in each community as cohesive, face valid topics (*n* topics PTSD: 8, Panic Disorder: 9, GAD: 11, MDD & Social Anxiety: 12, and OCD: 15). Table 4 depicts notable topics that emerged and their manually assigned label; the full set of topics and frequent words associated with them are included in the Appendix.

Table 4. Selected topics & words

Community	Label	Frequent words
GAD		
Topic 3	Sleep	attack, panic, sleep, night, wake, hour, felt, morning
Topic 7	Physical symptoms	anxiety, feel, pain, heart, symptom, chest, blood, doctor
OCD		
Topic 5	Intrusive thoughts	thought, intrusive, harm, head, image, violent, disturb
Topic 9	Contamination	hand, wash, clean, touch, shower, food, dirty
MDD		
Topic 2	Life & death	life, love, pain, hurt, live, kill, suicide, anymore, alive
Topic 6	Vitriol/Humiliation	fuck, hate, ugly, people, stupid, life, bullshit, pathetic
Social Anxiety		
Topic 4	Physical symptoms	speak, voice, nervous, shake, heart, panic, attack, stutter
Topic 7	Social settings	party, drink, invite, night, birthday, house, home, plan
Panic Disorder		
Topic 2	Panic onset	feel, start, breathe, head, feeling, minute, body, numb
Topic 5	Panic symptoms	heart, pain, chest, blood, rate, doctor, normal, check
PTSD		
Topic 3	Sleep	sleep, nightmare, night, dream, wake, attack, flashback
Topic 7	Traumatic event	accident, blood, pain, drive, ambulance, hospital, scream

Selected topics from the LDA analysis for each subreddit, shown as the top words in the cluster and the manually assigned label.

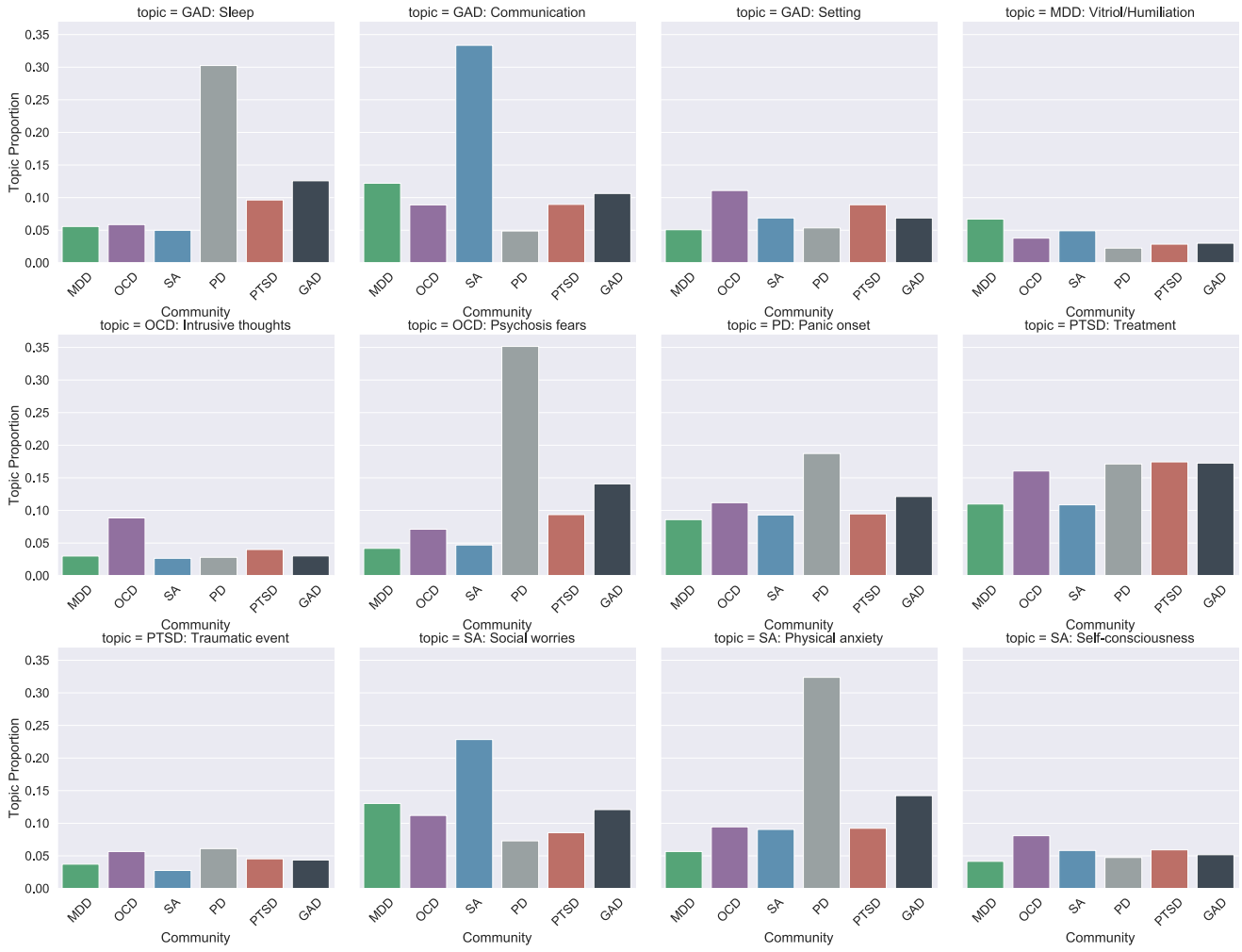
Communities had a number of topics in common that are related to clinical symptoms. The GAD, Social Anxiety, and Panic Disorder subreddits contained topics that could be broadly categorized as somatic concerns/physiological manifestations of anxiety (e.g. “feel,” “pain,” “symptom”). However, we observed variability in the concerns that were most salient to each subreddit. In the Social Anxiety subreddit, discussion primarily focused on observable symptoms of anxiety (e.g. “shake,” “voice,” “stutter,” “panic”), while somatic concerns in the Panic Disorder subreddit centered on symptoms that may be interpreted as dangerous (e.g. “heart,” “chest,” “palpitation”). Additionally, Sleep topic emerged for GAD, Panic Disorder, and PTSD communities. All three appeared to broadly indicate difficulties with sleep. In Panic Disorder, the Sleep topic included additional terms that may reflect efforts to induce sleep (e.g. “smoke,” “drink,” “weed”), while PTSD users additionally discussed “nightmares,” “dreams,” and “flashbacks.”

Topics relating to treatment were identified in all communities. In the GAD and OCD subreddits, Pharmacological treatment and Psychotherapy emerged as two separate topics. The Treatment topics that emerged for Social Anxiety and Panic Disorder were largely comprised of terms related to pharmacological interventions (e.g. “medication,” “prescribe,” “dose,” “ssri”), while the Treatment topics for PTSD and MDD were mixed between pharmacological and psychotherapy terms.

We additionally observed symptom-related topics that showed specificity to a given community. Many of the topics in the OCD community mapped on to common categories of fears and obsessions seen clinically in OCD samples (i.e. Contamination, Religious obsessions, Scrupulosity, Violent intrusive thoughts, and Checking behaviors). Interestingly, an Excoriation and Trichotillomania cluster also emerged from the OCD community—in line with the DSM-5 reclassification of Trichotillomania from Impulse-Control Disorders Not Classified Elsewhere to Obsessive-Compulsive and Related disorders. Discussion in the MDD subreddit was expectedly characterized the most negatively-valenced words—with themes discussion Isolation (e.g. “people,” “lack,” “emptiness,” “despair,” “understand”) and Vitriol (e.g. “fuck,” “hate,” “ugly,” “stupid,” “pathetic”). In the PTSD subreddit, we additionally noted themes associated with specific traumatic events (e.g. sexual trauma, military trauma).

In addition to these symptom-related themes, a number of co-occurring themes unrelated to clinical phenomena emerged. A School topic emerged in GAD, OCD, MDD, and Social Anxiety; Employment in GAD, MDD, and Social Anxiety; and Pastimes in GAD, OCD, and Social Anxiety communities.

Figure 2. Topic Overlap



We selected twelve topics for additional exploration: PTSD: Traumatic event, PTSD: Treatment, OCD: Intrusive thoughts, OCD: Psychosis fears, SA: Social worries, SA: Physical anxiety, SA: Self-consciousness, PD: Panic onset, GAD: Sleep, GAD: Communication, GAD: Setting, MDD: Vitriol/Humiliation. Figure 2 displays the proportion of these selected topics in the remaining communities; there are notable relationships between GAD: Communication & Social Anxiety, OCD: Psychosis fears & Panic Disorder, SA: Physical anxiety & Panic Disorder, and GAD: Sleep and Panic Disorder.

Community membership

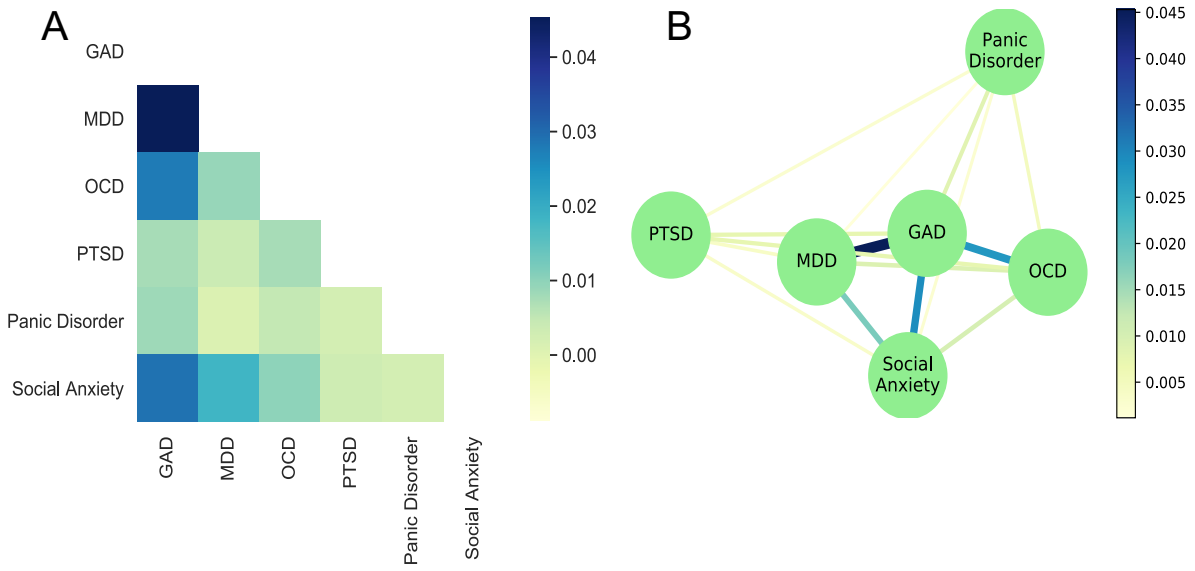
Table 5. Community Membership

Proxy dx/ subreddit	Total posters	Total members [a]	% Members
GAD	49,000	12,305	25.11%
MDD	134,148	38,690	28.84%
OCD	12,414	4,074	32.81%
Panic Disorder	2,207	516	23.38%
PTSD	6,139	1,462	23.81%
Social Anxiety	16,282	4,356	26.75%
r/frugal	9,613	1,559	16.22%
r/youshouldknow	7,443	1,996	26.82%
r/personalfinance	123,679	25,665	20.75%
r/fitness	64,774	13,351	20.61%
r/books	34,438	6,535	18.98%
r/teaching	3,728	799	21.43%

[a] Members are defined as users who have authored two or more posts

As shown in Table 4, 25-32% of posters in the clinical communities had written two or more posts and were considered "members" for this analysis. Figure 3 displays membership similarity between subreddits, which was highest for GAD + MDD, followed by GAD + Social Anxiety, GAD + OCD, and MDD + Social Anxiety. All scores, including membership overlap with control communities, are available in the Appendix, as well as a figure displaying raw user overlap..

Figure 3. Community membership



Discussion

The overarching goal of the present study was to characterize discussion in online communities associated with psychopathology. Using methods from natural language processing, we found high semantic similarity between communities for depression, anxiety, panic disorder, social anxiety, and OCD compared to non-clinical baselines. In addition, we show that LDA topic modeling successfully distilled these subreddits into digestible, face-valid themes—which contextualized the variability we observed from the semantic similarity analysis and were moderately related to known clinical phenomenology. Lastly, we characterized membership rates and membership overlap between communities, finding modest agreement with established nosology.

Given the high comorbidity rates of DSM-based disorders and notable overlap in their symptom criteria (Kessler et al., 2012; Zbozinek 2013), we would expect there to be overlap in the content of discussion between communities of different disorders. Supporting this, we show that all pairs of clinical communities exhibit higher semantic similarity compared to the control baselines. Importantly, this finding raises the possibility that patterns of epidemiology and diagnosis of mental disorders may be captured from individuals' verbal behaviors in online communities. Notably, this finding remained even after removing posts from multi-community users—suggesting that similarity in the content of discussion between communities of different disorders was not solely driven by rates of comorbidity/comembership and may reflect common features among the disorders themselves.

Overall, we found that discussions in all clinical communities contained a mix of 1) symptoms fairly specific to a given disorder (e.g. Contamination fears), 2) non-specific clinical concerns (e.g. Somatic symptoms) and 3) non-clinical activities (e.g. Pastimes). Themes from the

topic modeling analysis enable us to interpret differences in semantic similarity between clinical communities. The GAD subreddit had the highest average semantic similarity with the other clinical communities, which may be related to GAD's fairly general symptom criteria that are often seen transdiagnostically (i.e. restlessness, fatigue, difficulty concentrating, irritability, muscle tension, and sleep disturbances, APA, 2013). An alternative, less clinically meaningful possibility for this finding is simply that the name for the GAD subreddit is relatively general (r/anxiety) and may draw users with other anxiety disorders. This also illustrates a challenge of interpreting findings from this type of data and highlights the value of including analyses that are more descriptive in nature, such as topic modeling. Based on the topic modeling, high similarity between GAD and other communities appears to primarily stem from discussions about living with a disorder, sleep, treatment, and physiological anxiety—suggesting similarity partially stems from discussing overlapping symptoms as well as more general conversations about having and treating a mental disorder.

The identification of both specific and general clinical symptoms concords nicely with hierarchical approaches of describing psychopathology—general themes could feasibly correspond to higher-order levels while themes with a high degree of specificity could correspond to lower-order levels. Supporting this notion, we observe a moderate degree of overlap between the topics identified from the LDA and internalizing components of HiTOP (Kotov et al., 2017). Specifically, we found corresponding topics for the components: Insomnia (GAD: Sleep, PD: Sleep, PTSD: Sleep), Suicidality (MDD: Life & death), Dissociation (PD: Derealization), Interactive Anxiety (SA: Social worries, SA: Self-consciousness), Public Places (PTSD: Settings), Physiological Panic (GAD: Physical symptoms, PD: Panic symptoms, SA: Physical anxiety), Cleaning (OCD: Contamination), and Checking (OCD: Checking behaviors).

This speaks to the potential utility of natural language processing tools for studying the structure of psychopathology. Future work could examine whether these text-based topics correspond to extant measures of these components.

In comparing the membership size of these communities to lifetime rates of their corresponding diagnoses, we found that MDD was ranked highest in both (Kessler et al., 2012), but these metrics were otherwise unrelated (prevalence rank: MDD, social anxiety, PTSD, GAD, panic disorder, OCD; membership rank: MDD, GAD, social anxiety, OCD, PTSD, panic disorder). These differences are interesting, though difficult to interpret. Taken together with the semantic similarity findings, the GAD community may be relatively large in our sample based on its fairly broad content and non-specific name, appealing to a larger number of Reddit users. It's also possible that these differences stem from characteristics of the disorders themselves. For example, users with high degrees of social anxiety may find it anxiety provoking to engage with an online community. In fact, the social anxiety subreddit has the lowest ratio of posts to subscribers, supporting this possibility.

Membership similarity among clinical communities and lifetime comorbidity rates were both highest for GAD + MDD. These metrics were also both high for MDD + social anxiety. There was otherwise minimal overlap between comorbidity rates and membership similarity. Prior work has shown that comorbidity between GAD and MDD is strongly influenced by the overlap in symptom criteria (Zbozinek et al., 2012). However, in turning to the topic modeling results we observed the greatest thematic overlap between GAD and MDD on the more neutral topics (e.g. Living with anxiety/MDD, Relationships, Communication, and Treatment). The high similarity we observed may stem from a shared focus day-to-day concerns rather than specific symptoms.

While the communities we examined are all associated with a DSM-based disorder, we are cautious in extrapolating conclusions to real-world patients diagnosed with these disorders. We are unable to determine how representative Reddit users are of the general population or verify whether a user meets criteria for a particular disorder. Given that diagnostic reliability is an issue even among trained clinicians (Regier et al., 2013; Chmielewski et al., 2015), we expect many posts are written by users who would not meet diagnostic criteria for these disorders. In some ways, this is advantageous for studying psychopathology structure, however, it does limit our ability to make one-to-one comparisons between these online communities and their corresponding diagnosis. Additionally, similar to real-world communities, each subreddit has its own particular culture that may influence discussion in ways that are not yet fully known. Further work disentangling these complexities will be needed in order to make full use of these rich naturalistic datasets. The variation in membership size between subreddits makes it difficult to characterize and interpret membership similarity. We decided not to down-sample our data for the membership analyses, as we felt it was a conceptually important constraint on the similarity index. More advanced network representations would better represent the structure of the membership data and likely allow us to draw more precise conclusions.

We focused our analyses on communities associated with internalizing disorders; future work could examine the multitude of additional subreddits dedicated to other disorder and mental health more broadly (e.g. r/addiction, r/anger, r/aspergers, r/mentalhealth, r/opiatesrecovery, r/selfharm, etc.). With a larger number of subreddits, it would additionally be interesting to perform topic modeling on all posts aggregated together. Furthermore, we only looked at text from posts in order to best achieve our aims, however, comments and post meta-

data (e.g. number of likes, time of posting, length) likely contain meaningful information, particularly related to the social aspects of these communities.

The present study applies methods from natural language processing to study large online communities for users with DSM-based disorders. These findings help refine our understanding of common and specific symptoms of psychopathology as described directly by individuals with these disorders, point to promising avenues for future work in this area, and demonstrate the feasibility and utility of studying psychopathology from discussions in online communities.

References

Association, A. P. (2013). *Diagnostic and statistical manual of mental disorders (DSM-5®)*: American Psychiatric Pub.

Baumgartner, J., Zannettou, S., Keegan, B., Squire, M., & Blackburn, J. (2020). *The pushshift reddit dataset.* Paper presented at the Proceedings of the International AAAI Conference on Web and Social Media.

Bird, S., Klein, E., & Loper, E. (2009). *Natural language processing with Python: analyzing text with the natural language toolkit*: " O'Reilly Media, Inc.".

Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). Latent dirichlet allocation. *Journal of machine Learning research, 3*(Jan), 993-1022.

Chancellor, S., Baumer, E. P., & De Choudhury, M. (2019). Who is the" Human" in Human-Centered Machine Learning: The Case of Predicting Mental Health from Social Media. *Proceedings of the ACM on Human-Computer Interaction, 3*(CSCW), 1-32.

Chancellor, S., & De Choudhury, M. (2020). Methods in predictive techniques for mental health status on social media: a critical review. *NPJ digital medicine, 3*(1), 1-11.

Chmielewski, M., Clark, L. A., Bagby, R. M., & Watson, D. (2015). Method matters: Understanding diagnostic reliability in DSM-IV and DSM-5. *Journal of abnormal psychology, 124*(3), 764.

Curiskis, S. A., Drake, B., Osborn, T. R., & Kennedy, P. J. (2020). An evaluation of document clustering and topic modelling in two online social networks: Twitter and Reddit. *Information Processing & Management, 57*(2), 102034.

Feldhege, J., Moessner, M., & Bauer, S. (2020). Who says what? Content and participation characteristics in an online depression community. *Journal of Affective Disorders, 263*, 521-527.

Fischer, B. A. (2012). A review of American psychiatry through its diagnoses: the history and development of the Diagnostic and Statistical Manual of Mental Disorders. *The Journal of nervous and mental disease, 200*(12), 1022-1030.

Gjurković, M., & Šnajder, J. (2018). *Reddit: A gold mine for personality prediction.* Paper presented at the Proceedings of the Second Workshop on Computational Modeling of People's Opinions, Personality, and Emotions in Social Media.

Guloksuz, S., Pries, L., & Van Os, J. (2017). Application of network methods for understanding mental disorders: pitfalls and promise. *Psychological medicine, 47*(16), 2743.

Hyman, S. E. (2010). The diagnosis of mental disorders: the problem of reification. *Annual review of clinical psychology, 6*, 155-179.

Kawa, S., & Giordano, J. (2012). A brief historicity of the Diagnostic and Statistical Manual of Mental Disorders: issues and implications for the future of psychiatric canon and practice. In: BioMed Central.

Kendler, K. S. (2016). The phenomenology of major depression and the representativeness and nature of DSM criteria. *American Journal of Psychiatry, 173*(8), 771-780.

Kendler, K. S. (2016). Phenomenology of schizophrenia and the representativeness of modern diagnostic criteria. *JAMA psychiatry, 73*(10), 1082-1092.

Kendler, K. S., & Parnas, J. (2015). *Philosophical issues in psychiatry: Explanation, phenomenology, and nosology*: JHU Press.

Kessler, R. C., Petukhova, M., Sampson, N. A., Zaslavsky, A. M., & Wittchen, H. U. (2012). Twelve-month and lifetime prevalence and lifetime morbid risk of anxiety and mood disorders in the United States. *International journal of methods in psychiatric research, 21*(3), 169-184.

Kho, S. J., Padhee, S., Bajaj, G., Thirunarayan, K., & Sheth, A. (2019). Domain-specific Use Cases for Knowledge-enabled Social Media Analysis. In *Emerging Research Challenges and Opportunities in Computational Social Network Analysis and Mining* (pp. 233-246): Springer.

Kotov, R., Krueger, R. F., Watson, D., Achenbach, T. M., Althoff, R. R., Bagby, R. M., . . . Clark, L. A. (2017). The Hierarchical Taxonomy of Psychopathology (HiTOP): a dimensional alternative to traditional nosologies. *Journal of abnormal psychology, 126*(4), 454.

McCallum, A. K. (2002). Mallet: A machine learning for language toolkit. *http://mallet. cs. umass. edu*.

Miller, G. A. (1995). WordNet: a lexical database for English. *Communications of the ACM, 38*(11), 39-41.

Park, A., Conway, M., & Chen, A. T. (2018). Examining thematic similarity, difference, and membership in three online mental health communities from Reddit: a text mining and visualization approach. *Computers in human behavior, 78*, 98-112.

Pennington, J., Socher, R., & Manning, C. D. (2014). *Glove: Global vectors for word representation.* Paper presented at the Proceedings of the 2014 conference on empirical methods in natural language processing (EMNLP).

Plana-Ripoll, O., Pedersen, C. B., Holtz, Y., Benros, M. E., Dalsgaard, S., De Jonge, P., . . . Greve, A. N. (2019). Exploring comorbidity within mental disorders among a Danish national population. *JAMA psychiatry, 76*(3), 259-270.

Reddit. (2019). Reddit by the Numbers. Retrieved from https://www.redditinc.com/

Regier, D. A., Narrow, W. E., Clarke, D. E., Kraemer, H. C., Kuramoto, S. J., Kuhl, E. A., & Kupfer, D. J. (2013). DSM-5 field trials in the United States and Canada, Part II: test-retest reliability of selected categorical diagnoses. *American Journal of Psychiatry, 170*(1), 59-70.

Shen, J. H., & Rudzicz, F. (2017). *Detecting anxiety through reddit.* Paper presented at the Proceedings of the Fourth Workshop on Computational Linguistics and Clinical Psychology—From Linguistic Signal to Clinical Reality.

Spitzer, R. L., Md, K. K., & Williams, J. B. (1980). *Diagnostic and statistical manual of mental disorders.* Paper presented at the American psychiatric association.

Tadesse, M. M., Lin, H., Xu, B., & Yang, L. (2019). Detection of depression-related posts in reddit social media forum. *IEEE Access, 7*, 44883-44893.

Wardenaar, K. J., & de Jonge, P. (2013). Diagnostic heterogeneity in psychiatry: towards an empirical solution. *BMC medicine, 11*(1), 1-3.

Yoo, M., Lee, S., & Ha, T. (2019). Semantic network analysis for understanding user experiences of bipolar and depressive disorders on Reddit. *Information Processing & Management, 56*(4), 1565-1575.

Zbozinek, T. D., Rose, R. D., Wolitzky-Taylor, K. B., Sherbourne, C., Sullivan, G., Stein, M. B., . . . Craske, M. G. (2012). Diagnostic overlap of generalized anxiety disorder and major depressive disorder in a primary care sample. *Depression and anxiety, 29*(12), 1065-1071.

Appendix

LDA topics

Proxy dx	Label	Frequent words
GAD		
Topic 1	Living w/ anxiety	feel, thing, make, thought, life, worry, anxious, hate, time
Topic 2	Psychotherapy	anxiety, therapy, therapist, mental, health, disorder
Topic 3	Sleep	attack, panic, sleep, night, wake, hour, felt, morning
Topic 4	Communication	friend, talk, people, conversation, social, girl, person
Topic 5	Relationships	relationship, family, love, move, live, parent, friend
Topic 6	Setting	drive, walk, door, room, house, noise, wear, clean, road
Topic 7	Physical symptoms	anxiety, feel, pain, heart, symptom, chest, blood, doctor
Topic 8	Employment	work, interview, call, money, manager, company, office
Topic 9	Pastimes	video, game, watch, play, music, book, read, create, song
Topic 10	School	school, class, college, grade, study, exam, teacher
Topic 11	Pharmacological tx	effect, medication, prescribe, zoloft, dose, doctor
OCD		
Topic 1	Daily life	feel, make, thing, love, relationship, person, feeling
Topic 2	Living w/ OCD	obsession, compulsion, memory, fear, anxiety, life
Topic 3	Psychotherapy	therapist, diagnose, symptom, health, treatment
Topic 4	Settings	life, school, year, live, work, home, college, family
Topic 5	Intrusive thoughts	thought, intrusive, harm, head, image, violent, disturb
Topic 6	Psychosis fears	panic, schizophrenia, dream, fear, psychosis, develop
Topic 7	Checking behaviors	check, door, number, lock, time, walk, count, turn, ritual
Topic 8	Social	friend, people, talk, post, message, comment
Topic 9	Contamination	hand, wash, clean, touch, shower, food, dirty
Topic 10	Scrupulosity	girl, woman, sexual, porn, child, attract, pedophile
Topic 11	Excoriation & trich	hair, skin, mouth, nail, swallow, pick, teeth, body, face
Topic 12	Pharmacological tx	medication, effect, drug, zoloft, ssri, dose, prescribe
Topic 13	School	study, exam, class, write, student, information, book
Topic 14	Pastimes	game, watch, play, video, movie, song, music, listen
Topic 15	Religious obsessions	promise, pray, religious, christian, prayer, punishment
MDD		
Topic 1	Living w/ MDD	feel, depression, time, scar, thought, happy, felt, hard
Topic 2	Life & death	life, love, pain, hurt, live, kill, suicide, anymore, alive
Topic 3	Misc.	didnt, wont, isnt, feel, people, care, talk, whats, havent,
Topic 4	Social	friend, talk, girl, relationship, date, hang, group, text
Topic 5	Isolation	people, lack, emptiness, negative, social, despair
Topic 6	Vitriol/Humiliation	fuck, hate, ugly, people, stupid, life, bullshit, pathetic
Topic 7	Daily routine	sleep, drink, night, hour, wake, clean, shower, food
Topic 8	Treatment	depression, medication, doctor, therapy, anxiety
Topic 9	School	school, class, college, grade, study, high, game, play
Topic 10	Employment/finances	work, money, move, afford, bill, debt, rent, company
Topic 11	Family	mother, father, brother, sister, parent, family, child
Topic 12	Misc.	song, music, light, movie, world, soul, voice, dark, listen

Social Anxiety		
Topic 1	Social worries	people, talk, conversation, friend, person, group, quiet
Topic 2	Life w/ anxiety	anxiety, issue, overcome, fear, problem, struggle, suffer
Topic 3	Home life	life, year, parent, live, feel, family, move, depression
Topic 4	Physical anxiety	speak, voice, nervous, shake, heart, panic, attack, stutter
Topic 5	Coping	thought, negative, mind, fear, learn, life, people, accept
Topic 6	Relationships	date, girl, message, text, friend, crush, talk, post, respond
Topic 7	Social settings	party, drink, invite, night, birthday, house, home, plan
Topic 8	Employment	work, interview, call, customer, manager, apply, phone
Topic 9	Self-consciousness	walk, wear, smile, hair, face, laugh, stand, blush, stare
Topic 10	School	class, school, teacher, grade, college, presentation, group
Topic 11	Pastimes	play, game, video, chase, watch, boring, music, youtube
Topic 12	Pharmacological tx	medication, drug, doctor, effect, prescribe, ssri, dose
Panic Disorder		
Topic 1	Life w/ panic	panic, attack, anxiety, feel, experience, disorder, fear, life
Topic 2	Panic onset	feel, start, breathe, head, feeling, minute, body, numb
Topic 3	Panic setting	friend, leave, school, home, work, house, call, talk, drive
Topic 4	Sleep	sleep, night, wake, smoke, drink, asleep, weed, morning
Topic 5	Panic symptoms	heart, pain, chest, blood, rate, doctor, normal, check
Topic 6	Pharmacological tx	xanax, medication, prescribe, lexapro, dose, zoloft, ssri
Topic 7	Self-help	read, recovery, step, survive, learn, recover, approach
Topic 8	Derealization	watch, noise, scream, movie, reality, universe, exist
Topic 9	Coping	solution, conceptual, tangible, situation, concept
PTSD		
Topic 1	Life w/ PTSD	feel, make, thing, people, life, hard, time, hurt, understand
Topic 2	Treatment	ptsd, therapy, diagnose, symptom, trauma, medication
Topic 3	Sleep	sleep, nightmare, night, dream, wake, attack, flashback
Topic 4	Sexual trauma	sexual, relationship, abuse, assault, touch, rape, kiss
Topic 5	Settings	work, home, call, door, walk, leave, house, police
Topic 6	Family	parent, family, mother, brother, school, father, sister, child
Topic 7	Traumatic event	accident, blood, pain, drive, ambulance, hospital, scream
Topic 8	Military trauma	people, veteran, study, military, group, service, survivor

Membership overlap

Proxy dx/ subreddit	GAD	MDD	OCD	Panic Disorder	PTSD	Social Anxiety
GAD	—					
MDD	.045	—				
OCD	.028	.009	—			
Panic Disorder	.008	.001	.005	—		
PTSD	.008	.004	.008	.003	—	
Social Anxiety	.029	.018	.010	.002	.003	—
r/frugal	.002	.001	.001	0	.002	.002
r/youshouldknow	.001	.001	.001	0	.001	.002
r/personalfinance	.006	.006	.001	0	.001	.002
r/fitness	.008	.007	.003	0	.001	.005
r/books	.005	.003	.002	0	.001	.004
r/teaching	.001	0	.001	0	.000	.001

Membership measured as Jaccard similarity coefficient