



Article

A Corpus-Based Study of Adjectives and Collocates in Reddit Posts on Anxiety and Depression

Yang Yuhan¹, Lau Su Kia², Handoko³, Yap Leng Lee⁴, Sheena Kaur⁵

¹Tianjin Juemin Middle School, Tianjin, China

²Department of Malaysian Languages and Applied Linguistics, Faculty of Languages and Linguistics, Universiti Malaya, Kuala Lumpur, Malaysia

³English Department, Universitas Andalas, Indonesia; Faculty of Languages and Linguistics, Universiti Malaya, Kuala Lumpur, Malaysia

⁴Department of Asian and European Languages, Faculty of Languages and Linguistics, Universiti Malaya, Kuala Lumpur, Malaysia

⁵Department of English Language, Faculty of Languages and Linguistics, Universiti Malaya, Kuala Lumpur, Malaysia

SUBMISSION TRACK

Received: March 1, 2026
Final Revision: July 13, 2026
Accepted: July 14, 2026
Available Online: July 16, 2026

KEYWORDS

Anxiety, Depression, Adjective, Collocate, Emotional Meaning.

CORRESPONDENCE

*E-mail: lausk@um.edu.my

ABSTRACT

Talking openly about anxiety and depression (A&D) remains difficult for many people because of the stigma surrounding mental illness. Anonymous online platforms such as Reddit provide a space where users can express their thoughts and emotions more freely. This study considers how individuals linguistically construct and intensify emotional distress by examining (1) the adjectives used to express A&D, (2) the content-word collocates that co-occur with these adjectives, (3) the lexical features of these collocates, and (4) the emotional meanings conveyed through these collocational patterns. The dataset consisted of 1,440 Reddit posts (approximately 300,000 tokens) systematically sampled from the r/Anxiety and r/Depression subreddits between 2023 and 2025. An observational mixed-methods corpus linguistic approach was used to examine the data. Quantitative corpus linguistic analyses were carried out using AntConc, including frequency profiling and Mutual Information (MI) analysis, and were enhanced by qualitative concordance and Keyword-in-Context (KWIC) analysis to examine collocational patterns in context. The analysis shows a predominance of negatively valenced adjectives (e.g., anxious, depressed, hopeless, and suicidal), whose meanings are systematically intensified through their collocational environments. The collocates show distinctive lexical features. These include clinical nouns, linking and change-of-state verbs, and degree and frequency adverbs. These features construct varying levels of affective intensity and psychological distress. Emotional meaning is encoded in recurrent collocational patterns. Individual lexical items reveal only part of this meaning. This shows the value of collocational analysis for digital mental health research. The findings also possess practical implications. They may help improve the diagnostic sensitivity of automated digital mental health tools and foster more empathetic clinical communication.

I. INTRODUCTION

Anxiety and depression (A&D) are currently recognized among the most pervasive and debilitating mental health disorders worldwide, representing a serious worldwide health crisis

(García-Méndez et al., 2026; Salekin et al., 2018; Xiao et al., 2023). The World Health Organization (2022) estimated that nearly 970 million people were living with mental health conditions in 2019. Most of these cases involved anxiety and depression.

Depression alone affected around 280 million people, representing about 3.8% of the world's population, and was more common among women. Recent data from the United States also show that anxiety and depression are especially common among adolescents and people from lower-income groups. These findings highlight persistent social, economic, and demographic disparities (Brody & Hughes, 2025). Although anxiety and depression affect millions of people, social stigma still prevents many from seeking professional support. Meta-analytic findings show that stigma has a moderate but highly significant negative effect on a person's willingness to seek support, causing millions to suffer in silence (Schnyder et al., 2017).

In response to these far-reaching clinical barriers and the fear of social judgment, a growing number of individuals are turning to anonymous digital platforms to candidly share their emotional distress (Baumel et al., 2020; Habicht et al., 2024; Morris et al., 2015). Platforms such as Reddit have become vital sanctuaries for those searching for mental health communities and awareness of both anxiety and depression (Grub, 2025). Subreddits like r/Anxiety and r/Depression have experienced pronounced growth, particularly following the mental impact of the COVID-19 pandemic, thereby turning into significant repositories of naturally occurring mental health discourse (De Choudhury et al., 2013; Low et al., 2020). The intensity and openness of the self-disclosures found on these platforms can be theoretically comprehended through the lens of the Online Disinhibition Effect, which claims that the lack of identification and reduction of traditional social cues inherent in online environments enable individuals to express intense, vulnerable emotions far more candidly than they would in face-to-face clinical contexts (Knaevelsrud & Maercker, 2007; Mayopoulos & Farber, 2023; Prescott et al., 2017). These platforms grant researchers an unprecedented, unfiltered window into the real experiences of individuals struggling with A&D.

It is vital to examine the cognitive systems that dictate how individuals experiencing A&D process and articulate their worldviews. Individuals suffering from depressive disorders frequently develop negative cognitive schemas, deeply rooted patterns of distorted thinking that foster feelings of hopelessness, worthlessness, and profound despair (Gotlib & Joormann, 2010;

Kovacs & Beck, 1978). These cognitive distortions are reflected in the way people describe their emotional experiences. They contribute to the use of negatively biased language patterns (Gonda et al., 2015). The Cognitive Theory of Anxiety further stresses that anxious individuals possess threat-focused cognition, constantly anticipating danger and viewing ambiguous situations as highly threatening (Everaert et al., 2020; MacLeod & Cohen, 1993). Users' psychological states shape the language they use. Their posts often contain words that reflect negative thinking patterns and fear (Quillivic et al., 2023; Zarate et al., 2023).

Language does much more than simply transmit psychological information. It actively evokes and structurally shapes emotional experiences (Collins & Baker, 2023). The concept of lexical collocation in corpus linguistics, which refers to the habitual co-occurrence of words, provides a valuable structure for understanding how emotional meaning is structurally constructed (Soriano, 2013). Adjectives, in particular, are central to the expression of emotion, but they rarely function in isolation; they acquire their full affective resonance through their collocates, such as the nouns they modify, the verbs that precede them, or the adverbs that intensify them (Bélignon, 2020). For example, the adjective "low" carries a distinctly different affective charge in the phrase "low mood" than in "low risk," illustrating that emotional meaning is dynamic, culturally situated, and arises specifically from collocational patterns. Research in applied linguistics has shown that collocational analysis is an effective way to examine evaluative and emotional meaning across different types of texts. This makes it particularly useful for studying contexts where emotional expression plays an important role (Absattar et al., 2022; Edo Marzá, 2011; Hunston, 2010; Soriano, 2013).

Reddit offers a large collection of naturally occurring language data. However, there is still limited research on the detailed linguistic patterns found in these online discussions. Recent scoping reviews have identified dozens of studies using Reddit data to explore A&D (Almeida et al., 2024; Boettcher, 2021; De Choudhury et al., 2013; Low et al., 2020; Robertson et al., 2019; Zhang et al., 2022). The vast majority of these studies focus strictly on computational classification, risk detection, and automated symptom surveillance. Previous research frequently uses Natural

Language Processing (NLP) or machine learning to predict depression or track spikes in anxiety, but it fundamentally overlooks the detailed linguistic interactions between adjectives and their specific collocates (Kerz et al., 2023; Nguyen et al., 2022; O'Dea et al., 2025). Existing studies often treat words as independent data points for classification algorithms. This approach overlooks how language structurally conveys emotional intensity, nuance, as well as the lived reality of anxiety and depression.

The need to address this research gap is of considerable practical relevance to the extensive fields of psychology, applied linguistics, and digital healthcare. Specific collocational patterns provide important insights into how individuals articulate distress and are essential for the exact interpretation of online narratives. These insights are key for the development of increasingly empathetic and sensitive digital mental health interventions. Moving beyond solely keyword detection improves the interpretation of online emotional distress. It also helps bridge the gap between theoretical computational linguistics and practical mental health care. Building on these psychological and linguistic insights, this study examines how adjectives and their collocates are used to express A&D in Reddit posts. It examines not only the most frequent adjectives but also the collocational environments that shape and intensify their meanings, revealing how emotional intensity is linguistically constructed in online mental health discourse. Drawing on the gaps identified in the existing literature, this study addresses the following research questions (RQs):

1. What adjectives are commonly used by Reddit users to express anxiety and depression?
2. Which content-word collocates co-occur with these adjectives in Reddit posts?
3. How are these collocates characterized in terms of their lexical features?
4. Do the collocates of these adjectives convey positive, negative, or neutral emotional meanings?

II. METHODS

Research Design

Reddit was selected as the data source for its anonymity, which facilitates open and candid user expression, and its active mental health communities. A mixed-methods design, integrating quantitative

corpus analysis with qualitative interpretation, was employed to examine the linguistic patterns of adjectives and their collocates in Reddit discussions of anxiety and depression. The research follows a corpus-based approach, utilizing a large-scale collection of naturally occurring texts to identify linguistic patterns. A corpus of 1,440 Reddit posts (720 from r/Anxiety and 720 from r/Depression) was selected for the period 2023 to 2025, utilizing the anonymity and post-pandemic surge in online mental health discourse. The quantitative research method employed corpus linguistic techniques to identify commonly used adjectives and their collocates. In contrast, the qualitative research method examined the emotional meanings of these collocates through concordance analysis. Employing this dual approach enabled the study to provide a comprehensive understanding of how users articulate emotional distress in online discourse.

Collection of Corpus Data

Public posts were collected from the subreddits r/Anxiety and r/Depression using the Reddit API between April 2023 and March 2025. For each subreddit, 720 posts were sampled to achieve a balanced dataset. All data used in this study were collected from publicly available, open-access sources in accordance with Reddit's API terms of service. The dataset comprised only the full text of original user posts. Titles, comments, and replies were excluded to maintain the focus on original user reports. Additionally, the collection was limited to publicly available data, and all usernames and personal identifiers were excluded from the outset to protect user privacy. The raw text underwent a three-stage data cleaning process to prepare it for analysis. All personal metadata, including usernames, user IDs, and other identifiable information, was systematically removed to anonymize the dataset and protect user privacy. Non-textual elements, such as URLs, HTML tags, markdown formatting, emojis, and images, were then removed or replaced with neutral markers to ensure that the analysis focused solely on the linguistic content. Punctuation and whitespace were subsequently standardized by correcting inconsistent spacing and irregular line breaks, resulting in a clean and continuous body of text. These preprocessing procedures ensured that the corpus contained no personally identifiable information and consisted entirely of users' written language.

Corpus Construction

The collected Reddit posts were compiled into a single corpus for analysis, comprising approximately 300,000 tokens (words). This corpus provides a substantial database of naturally occurring language related to anxiety and depression. The total number of tokens was verified through a preliminary word-frequency analysis conducted using *AntConc* (see below). The resulting corpus formed the basis for all subsequent analysis, enabling robust examination and comparison of linguistic patterns across the r/Anxiety and r/Depression communities.

AntConc

AntConc, a freely available corpus analysis toolkit, was used for quantitative analysis (Anthony, 2022). The *Word*, *KWIC*, and *Collocate* functions from the package were used. The *Word* feature was used to generate a list of the most frequent lexemes; the *KWIC* feature was employed to analyze the concordance lines, in which each target word was displayed in its context of use, and the *Collocate* feature identifies words that frequently occur together within a specified context, revealing patterns of word association in a corpus. Apparently, *AntConc*'s capabilities enabled efficient processing of large text datasets, making it suitable for this study's objectives. In this study, *AntConc* was used to analyze the language of Reddit posts from r/Anxiety and r/Depression to meet the research objectives. To identify the adjectives most used by users (RQ 1), the Word List tool in *AntConc* was used to generate a frequency list from the collected data. This list was then examined to isolate adjectives, using common adjective endings such as *-ed*, *-ing*, and *-ful* to guide selection. The Collocates function in *AntConc* was used to address RQ 2, which focused on identifying collocates and examining their lexical patterns. This allowed the examination of words that frequently appeared near each target adjective, helping uncover patterns in how users describe their emotional states. The study used the Concordance tool to examine each collocate in its original context, thereby RQ 4 of interpreting its emotional meaning.

Adjective Extraction and Emotional Meaning Analysis

A systematic process of adjective extraction and collocate analysis was conducted to examine how emotional meaning is conveyed through

adjectives and their surrounding words. The extracted words were verified as adjectives through consultation of authoritative English dictionaries. Initially, the Word List function in *AntConc* was used to generate a comprehensive list of word frequencies in the corpus. All words tagged as adjectives and appearing at least 15 times were identified and extracted for further analysis. From this list, the top 15 most frequent adjectives were selected as node words (as shown in the shaded section of Table 1), based on their relevance to emotional expression in the context of anxiety and depression. For each selected adjective, a collocate analysis was performed to identify words that commonly appeared immediately to its left and right in the text. The Collocates function in *AntConc* was used for this step, with a span of one word to the left and one word to the right of the target adjective. Only collocated words that appeared at least twice in combination with the target adjective were included in the analysis to ensure that the identified patterns reflected consistent usage rather than isolated instances. MI scores were calculated using *AntConc*'s collocation function to measure the strength of association between the target adjectives and their collocates. MI is a statistical measure that compares the observed frequency of a word pair with the frequency expected if the two words occurred independently. Higher MI values indicate stronger and more meaningful associations. In this study, MI helped distinguish true collocational patterns from incidental co-occurrence. The emotional meaning of each collocate was examined through qualitative review. This involved manually inspecting concordance lines (KWIC) for each adjective-collocate pair to understand how these combinations contributed to the expression of emotional states in the posts.

III. RESULTS

A total of 33 adjectives occurred at least 15 times in expressions of anxiety and depression, as shown in Table 1. The most frequent were "anxious" (306), "panic" (291), "scared" (267), and "depressed" (207), with "desperate" (16) and "disappointed" (15) being the least frequent. Overall, the adjectives predominantly conveyed negative emotional states, reflecting the psychological distress associated with anxiety and depression. For subsequent collocational analysis, the 15 most frequent adjectives, ranging from

“anxious” (306) to “miserable” (49) were selected as node words.

Table 1. Commonly used adjectives by Reddit users to express anxiety and depression

Adjectives	Frequencies
anxious	306
panic	291
scared	267
depressed	207
mental	198
sick	132
afraid	114
sad	111
worried	109
struggling	91
crazy	84
suicidal	81
stressed	54
worrying	54
miserable	49
exhausted	47
nervous	45
terrified	45
emotional	41
hopeless	32
depressive	30
useless	27
weak	27
concerned	26
awkward	23
annoying	23
embarrassed	23
nauseous	20
frustrated	19
painful	19
pathetic	18
desperate	16
disappointed	15

Anxiety-related adjectives such as “anxious” (306), “panic” (291), and “scared” (267) reveal intense fear, nervousness, and unease. These words indicate quick reactions to perceived threats, suggesting a constant undercurrent of anxiety. Terms like “afraid” (114), “worried” (109), and “nervous” (45) reflect heightened alertness, as if bracing for potential problems. Adjectives like “concerned” (26) reflect ongoing worry and

emotional preoccupation, while “struggling” (91) and “terrified” (45) evoke the overwhelming impact of unresolved anxiety. These terms demonstrate that anxiety involves a mix of emotions, behaviors, and thought patterns. In addition to anxiety, depression-related adjectives are prevalent. Words such as “depressed” (207) and “suicidal” (81) signal severe distress. “Sad” (111), “miserable” (49), “hopeless” (32), and “depressive” (30) imply emotional numbness and despair. These lexical choices suggest emotional states that extend beyond sadness, indicating a deeper sense of hopelessness and loss of purpose for many individuals. Self-critical adjectives like “useless” (27), “weak” (27), “pathetic” (18), and “desperate” (16) emphasize low self-worth and helplessness. “Disappointed” (15) and “frustrated” (19) reflect weariness and a loss of faith, often leading to the internalization of setbacks as proof of failure.

Several adjectives overlap between anxiety and depression, including “stressed” (54), “emotional” (41), “crazy” (84), and “exhausted” (47). These words highlight common symptoms such as overwhelm and mood swings, illustrating how anxiety and depression are often experienced together. On the other side, 10 adjectives end with the morpheme *-ed*: “scared” (267), “depressed” (207), “worried” (109), “stressed” (54), “exhausted” (47), “terrified” (45), “concerned” (26), “embarrassed” (23), “frustrated” (19), “disappointed” (15). Additionally, three adjectives end in the *-ing* morpheme: “struggling” (91), “worrying” (54), and “annoying” (23). These adjectives play distinct roles in describing the emotional and physical states of anxiety and depression, where *-ed* focuses on the inner emotional state of users and *-ing* highlights the source or the root cause that represents the things or subjects. Typically, to express a person’s internal condition resulting from an experience or action, adjectives ending with *-ed* are used frequently, such as “scared” (267) and “depressed” (207). They typically express a person’s internal condition, anxiety, and depression resulting from an experience or action. In contrast, adjectives ending with *-ing*, like “worrying” (54) or “annoying” (23), describe the qualities of something or someone that provokes a reaction of anxiety and depression.

Adjectives such as “worried” (109) and “concerned” (26) denote feelings of anxiety, often tied to anticipatory distress or uncertainty, which are

consistent with the high-arousal emotional anxiety. For example, one user wrote on Reddit: *“I’m so scared and worried, and I definitely won’t sleep anytime soon.”* In contrast, “scared” (267), “afraid” (114), and “terrified” (45) represent escalating degrees of fear, ranging from mild apprehension to intense dread, and reflect acute emotional responses prevalent in anxiety-related discourse. Take “scared” (267), for example; one user said: *“I’m extremely scared of death, and it makes my stomach hurt”*. Meanwhile, “embarrassed” (23) and “awkward” (23) reflect discomfort in social situations, potentially highlighting stigma-related experiences of anxiety in mental health discussions. “Embarrassed” is often tied to feelings of shame or self-consciousness, perhaps triggered by perceived judgment or vulnerability when sharing one’s struggles. As can be seen in the narrative, *I feel too embarrassed to tell my friends, some of them make fun of LGBT people*. “Hopeless” (32) and “desperate” (16) both suggest a perceived lack of options or solutions, emphasizing severe emotional distress and aligning with narratives of depression, such as *“One day I’ll die which makes me feel so hopeless for the future, like what’s the purpose of life?”* As shown in the sentence, such adjectives carry significant weight in mental health discourse, often marking moments of crisis or profound suffering, and may overlap with discussions of suicidal thoughts or existential struggles, which show a strong feeling of depression.

Collocates of Anxiety- and Depression-Related Adjective Nodes and Their Lexical Features and Emotional Meanings

This section presents the results of the collocation analysis conducted on a corpus of Reddit posts from the r/Anxiety and r/Depression subreddits. The analysis examines how different parts of speech (POS), namely nouns, verbs, adjectives, and adverbs, function as collocates of A&D-related node words and contribute to the construction of emotional meaning. The emotional functions of the collocates are interpreted not only on the basis of their semantic roles but also in relation to their MI scores. Collocates that co-occurred with the node words significantly more often than would be expected by chance were identified using MI values. MI scores below 2 suggest weak relationships, scores between 2 and 3 indicate moderate relationships, and scores of 3 or higher

indicate strong relationships. MI values of 5 or more reflect very strong collocations, often indicating fixed expressions or specialized terms.

Noun Collocates and Their Lexical Features and Emotional Meanings

Nouns in the corpus show a clear tendency to represent psychological states, clinical conditions, and situations of distress, as demonstrated by the noun collocations in Table 2.

Table 2. Nouns that collocate with various anxiety- and depression-related adjective nodes

Adjective Nodes	Nouns (Collocates)	Freq		MI Score
		Left	Right	
anxious	person	0	9	4.05
	thoughts	0	9	4.33
	attachment	0	3	7.55
	level	0	2	4.87
	mess	0	2	4.67
panic	attacks	0	106	8.62
	attack	0	97	8.41
	disorder	0	16	6.66
	anxiety	7	3	2.16
mental	health	0	106	8.54
	breakdown	0	11	8.88
	illness	0	9	7.90
	hospital	0	7	6.59
	state	0	5	5.89
	issues	0	4	4.50
	disorders	0	3	7.59
	disorder	0	2	4.21
	disciplines	0	2	9.59
	stability	0	2	8.00
	clarity	0	2	8.00
	energy	0	2	4.48
	sound	0	2	6.62
	thoughts	0	2	3.63
crazy	anxiety	1	2	2.22
suicidal	thoughts	0	20	7.41
	ideation	0	7	10.88
	ideations	0	2	10.88
	depression	0	2	3.18

The noun “thoughts” functions as a collocate that frequently appears alongside emotionally intense adjective node words, such as “anxious” (Freq: 9; MI: 4.33), “sad” (Freq: 2; MI: 3.63), and “suicidal” (Freq: 20; MI: 7.41). In fact, “thoughts” is an abstract plural count noun that refers to internal mental processes, ideas, reflections, or mental images. In A&D discourse, particularly

on platforms like Reddit, thoughts often represent the inner dialogue or mental noise experienced by users suffering from emotional distress. When collocating with “anxious”, thoughts typically refer to racing, repetitive, or intrusive mental activity, such as in sentences like “*Anxious thoughts keep me up at night*”. This collocation highlights the cognitive side of anxiety, showing how mental health conditions not only affect emotion but also disturb patterns of thinking. Most significantly, the collocation of “thoughts” with “suicidal” indicates severe emotional distress, as shown in Table 2. The extremely high MI score reveals a strong semantic bond between these two words. Suicidal thoughts are a well-recognized phrase in both clinical and everyday language, used to describe dangerous internal experiences where the individual contemplates self-harm or ending their life. The repeated use of this collocation indicates that Reddit users often resort to language to express profound distress, using “thoughts” to convey the invisible yet overwhelming burden they carry.

“Attack” (Freq: 97; MI: 8.41) appears to be highly related to “panic”, and its “attacks” (Freq: 106; MI: 8.62) also play a significant role in “panic”. Lexically, an “attack” is a countable noun, often used to describe a sudden and intense episode of physical or emotional disturbance. The singular form “attack” is often used in contexts such as “had a panic attack”, where it functions as a concrete representation of a highly distressing event. It denotes a sharp, sudden onset of intense fear, racing heartbeat, shortness of breath, or loss of control, common symptoms reported by individuals suffering from panic disorder or severe anxiety. Its high MI score suggests that it is a key term in users’ descriptions of their psychological state. The use of “attack” implies something that happens to a person, beyond their control, reflecting how overwhelming and involuntary the experience feels. On the other hand, the plural form “attacks” appears just as frequently, if not more, and is typically used when users refer to recurring episodes. Phrases like “I get panic attacks regularly” show how the condition is perceived not as a one-time crisis, but as a chronic and repeated disruption to daily functioning. The shift from singular to plural highlights the persistent and debilitating nature of the emotional state, suggesting a cycle of fear and helplessness that many users endure over time.

Notably, the noun “ideation” and its plural

form “ideations” both appear in strong collocation with the adjective “suicidal” (MI: 10.88), reflecting a severe and specific lexical relationship. From a lexical perspective, “ideation” is an abstract, uncountable noun derived from the root “idea” and used primarily in medical and psychological contexts to refer to the process of forming ideas, particularly persistent or intrusive thoughts. In A&D discourse, however, “ideation” is rarely used in general terms; it most often appears in the fixed phrase “suicidal ideation”, which refers to recurring thoughts or preoccupations with ending one’s life. The high MI score indicates that “suicidal ideation” forms a highly specialized collocation, not typically found outside of clinical or mental health contexts. On the A&D subreddit, this term is used by individuals who either report their own experiences or discuss others’ experiences. The use of “ideation” rather than simpler alternatives (such as “thoughts”) demonstrates an effort to provide precise language for a highly sensitive and complex experience. It allows users to refer to suicidal thinking in a way that is both specific and distanced, which may help reduce stigma or make the topic easier to address. Moreover, the plural form, “ideations”, is far less common in general use but still appears with the same “suicidal”. Its usage often suggests the presence of multiple or varying types of harmful thoughts.

A range of other nouns also show strong collocational relationships with the adjective node words, as indicated by their high MI scores. For instance, the noun “attachment” frequently co-occurs with “anxious” (Freq: 3; MI: 7.55), suggesting a psychological or emotional context often discussed in mental health discourse, such as “anxious” attachment. These collocations, though less frequent, demonstrate significant lexical associations that reflect specialized usage and deeper connections in the discourse surrounding anxiety and depression. Most importantly, words like “depression”, “attack”, “disorder”, “anxiety”, and “breakdown” are negative, reflecting emotional distress or instability. Although “stability” and “clarity” are positive, they indicate a state of balance and mental focus. The terms “person”, “level”, “health”, and “issues” are neutral, describing conditions that have no substantial emotional impact. Therefore, most of these nouns convey negative emotions, with a few, such as stability and clarity, having positive or neutral meanings.

Verb Collocates and Their Lexical Features and Emotional Meanings

Verbs in the corpus are particularly important for expressing emotional states and changes in emotional experience. The data in Table 3 shows widespread use of psychological verbs and state-of-being verbs.

Table 3. Verbs that collocate with various anxiety- and depression-related adjective nodes

Adjective Nodes	Verbs (Collocates)	Freq		MI Score
		Left	Right	
anxious	feeling	13	0	3.95
	get	10	0	2.03
	being	9	0	2.95
	gets	2	0	3.48
	thinking	2	0	2.27
panic	have	14	0	1.45
	having	11	0	3.93
	blown	4	0	7.57
mental	break	0	2	4.48
scared	being	6	0	2.56
	getting	2	0	1.71
depressed	being	9	0	3.51
	feeling	3	0	2.51
sick	feel	18	0	3.53
	getting	5	0	4.04
	feeling	4	0	3.48
	got	3	0	2.82
sad	feel	8	0	2.61
	being	6	0	3.83
	feeling	3	0	3.00
	sound	2	0	6.62
worried	been	5	0	2.43
	getting	4	0	4.00
	get	3	0	1.78
struggling	feel	3	0	1.22
	been	18	0	4.54
crazy	going	9	0	4.76
	feel	5	0	2.33
	go	3	0	3.07
	have	2	0	0.44
	been	2	0	1.49
suicidal	feel	5	0	2.39
	get	3	0	2.21
	feeling	3	0	3.45
stressed	been	2	0	1.54
	get	4	3	3.01
	feel	2	0	1.65

worrying	stop	4	0	5.47
	started	2	0	3.90
	stuck	2	0	5.97
	spent	2	0	7.11
miserable	feel	2	0	1.79
	being	2	0	3.42

Table 3 shows that the verb “feel” frequently co-occurs with four adjective nodes: “sick” (Freq: 18; MI: 3.53), “sad” (Freq: 8; MI: 2.61), “suicidal” (Freq: 5; MI: 2.39), and “crazy” (Freq: 5; MI: 2.33). From these data, it can be inferred that the verb “feel” functions primarily as a linking (copular) verb, connecting the subject to an adjective that describes an emotional or physical state. In the context of Reddit posts on anxiety and depression, “feel” frequently collocates with adjectives such as “sad” and “suicidal”, reflecting negative emotional states commonly associated with mental health struggles.

The verb with the suffix morpheme *-ing*, “feeling”, collocates with five different adjectives, which include “anxious” (Freq: 13; MI: 3.95), “sick” (Freq: 4; MI: 3.48), “sad” (Freq: 3; MI: 3.00), “suicidal” (Freq: 3; MI: 3.45), and “depressed” (Freq: 3; MI: 2.51). This may suggest that the word “feeling”, when used as a progressive verb, plays a crucial role in expressing emotional experiences that are not brief or momentary, but ongoing. In Reddit posts about anxiety and depression, users often write sentences like “I’ve been feeling sad for weeks”. The use of the progressive form indicates that the emotions are not fleeting thoughts but rather part of a more prolonged struggle. Another verb, “get”, collocates with “anxious” (Freq: 10; MI: 2.03), “stressed” (Freq: 4; MI: 3.01), and “suicidal” (Freq: 3; MI: 2.21). The verb “get” frequently collocates with anxiety- and depression-related adjectives to express emotional transitions or the development of negative mental states. It signals the onset or intensification of an emotional experience rather than a fixed emotional condition. For example, “get stressed” (Freq: 4, MI: 3.01) indicates a rising sense of pressure or overwhelm.

Many other verbs collocate with the adjective node words, indicating strong lexical associations, as reflected in their high MI scores. For example, the verb “stuck” appears with “worrying” (Freq: 2; MI: 5.97), often reflecting a feeling of being trapped in persistent or overwhelming thoughts. Likewise, the verb “having” frequently co-occurs

with “panic” (Freq: 11; MI: 3.93), as in expressions such as “having panic attacks”, which point to direct, experiential descriptions of emotional distress. Mostly, verbs such as “feeling”, “getting”, “thinking”, “being blown”, “breaking”, “stopping”, “starting”, and “being stuck” express various emotional states. “Feel” and “feeling” are neutral terms that describe emotions or sensations. “Get” and “getting” often indicate negative transitions, such as becoming upset. “Thinking” is a neutral term that focuses on mental activity. “Blown” and “break” carry negative connotations, suggesting intense or destructive emotions. “Stop” and “start” are neutral, marking the beginning or end of a state. “Stuck” implies negativity, indicating emotional or physical immobility. “Have”, “having”, “been”, and “going” are neutral, representing ongoing or completed actions. “Spent” suggests emotional exhaustion, typically a negative connotation. Most of these verbs imply negative emotional states or transitions, with a few being neutral.

Adjective Collocates and Their Lexical Features and Emotional Meanings

Adjectives in the corpus play an evident role in explicitly conveying emotional intensity and valence as indicated in Table 4.

Table 4. Adjectives that collocate with various anxiety- and depression-related adjective nodes

Adjective Nodes	Adjectives (Collocates)	Freq		MI Score
		Left	Right	
anxious	depressed	4	3	4.07
	sad	0	2	3.17
panic	bad	8	0	3.22
	severe	7	0	5.23
	awful	2	0	4.30
mental	severe	3	0	3.97
sick	worried	2	0	4.40
depressed	anxious	3	4	4.07
	alone	0	2	2.80
sad	anxious	2	0	3.17
worried	sick	2	0	4.40

As presented in Table 4, the adjective “severe” co-occurs with two different adjectives (node words): “panic” (Freq: 7; MI: 5.23) and “mental” (Freq: 3; MI: 3.97). The data indicate that the collocation involving “panic” has the highest frequency. Here, the word “severe” functions as an adjective expressing a high degree of seriousness or intensity. In collocations such as “severe panic”, the adjective “severe” emphasizes the

extreme intensity of the emotional experience. In subreddit posts about anxiety and depression, “severe” helps users convey that their emotional state has reached a critical or unbearable level. The collocation “severe mental” is less frequent, but when it appears, it typically occurs in phrases like severe mental health issues. In these cases, “severe” again signals seriousness, often with long-term consequences, and possibly requiring medical or therapeutic attention. Moreover, other adjective collocates include “depressed”, “sad”, “bad”, “awful”, “worried”, and “sick”, which are also used to modify adjective-node words. These adjectives function as evaluative and emotive modifiers that enrich the meaning of the emotional states of anxiety and depression described.

“Depressed” (Freq: 7; MI: 4.07) as a collocate of “anxious” is a psychologically marked term that carries a clinical connotation, often associated with diagnosed mental illness. As an adjective, it directly conveys a persistent, deep emotional low, and when used in user-generated content, it reflects not only emotional exhaustion but also a sense of hopelessness or emptiness. The adjective “sad” (Freq: 2; MI: 3.17) expresses a basic emotional state characterized by unhappiness or sorrow. It is less clinical than “depressed” but no less impactful, often used to capture emotional vulnerability in a more accessible and universal way. Both words allow users to express pain clearly, but “depressed” frequently signals a more long-term and serious emotional state. The adjectives “bad” (Freq: 8; MI: 3.22) and “awful” (Freq: 2; MI: 4.30), as collocates of “panic”, contribute to the intensification of negative feeling, though they do so in different registers. “Bad” is a general, often vague descriptor, frequently used in casual speech to suggest discomfort, guilt, or distress without specifying the cause. In contrast, “awful” is stronger and more emotionally charged, indicating a deeper, often overwhelming experience of shame, regret, or emotional turmoil. Both are commonly used in personal narratives to convey emotional states, and their simplicity allows for quick and relatable expressions of suffering.

“Worried” (Freq: 2; MI: 4.40) and “sick” (Freq: 2; MI: 4.40) bring additional nuance. “Worried” describes a mental preoccupation with fear or uncertainty and is often used in a continuous or repetitive sense, emphasizing the ongoing nature of anxiety. It highlights the cognitive aspect of

emotional strain, focusing on thoughts rather than feelings alone. Meanwhile, “sick” straddles both physical and emotional realms. While it traditionally refers to physical illness, in mental health contexts, it is often used metaphorically to reflect a visceral reaction to emotional stress. Significantly, the words “depressed”, “sad”, “bad”, “severe”, “awful”, “worried”, “anxious”, “alone”, and “sick” all carry negative emotional meanings. “Depressed” and “sad” reflect emotional pain, with “depressed” being more intense, while “bad” is a broad term for something undesirable, and “severe” indicates extreme negativity. Other adjective collocates also suggest strong dislike or discomfort, referring to both physical and emotional distress. All these words describe situations or feelings that are overwhelmingly negative.

Adverb Collocates and Their Lexical Features and Emotional Meanings

Adverbs collocate with different adjective-node words, as shown in Table 5; these are among the most frequent collocations. In these collocations, adverbs modify adjectives, indicating their intensity or degree. They either strengthen or soften the meaning of the adjective, adding subtlety and precision to the emotional or descriptive expression.

Table 5. Adverbs that collocate with various anxiety- and depression-related adjective nodes

Adjective Nodes	Adverbs (Collocates)	Freq		MI Score
		Left	Right	
anxious	very	14	0	4.15
	more	12	0	3.42
	extremely	8	0	5.71
	incredibly	6	0	6.55
	super	4	0	4.77
	socially	3	0	6.55
	less	3	0	4.19
	constantly	3	0	3.44
	pretty	3	0	3.45
	highly	2	0	6.64
	whenever	0	2	3.81
	until	0	4	3.61
scared	so	43	0	3.35
	too	11	2	4.05
	extremely	3	0	4.49
	constantly	3	0	3.05
	mostly	2	0	4.43
	truly	2	0	4.30
	already	2	0	3.07
	pretty	2	0	3.06
depressed	since	0	6	3.47
	severely	5	0	7.68
	lately	0	4	5.20
	already	4	0	4.44
	actually	4	0	4.09
	before	0	4	3.10
	deeply	3	0	6.35
	completely	3	0	4.30
	incredibly	2	0	5.52
	extremely	2	0	4.28
sick	too	2	2	2.71
	most	2	0	2.73
	so	12	0	2.52
	physically	2	0	5.45
	literally	2	0	4.15
afraid	so	14	0	2.96
	too	4	0	3.57
	constantly	4	0	4.86
	more	3	0	2.85
	also	2	0	2.55
sad	so	16	0	3.10
	really	6	0	3.26
	more	3	0	2.98
	just	3	0	0.81
	still	2	0	2.48
worried	so	6	0	1.80
	very	5	0	4.16
	constantly	4	0	5.34
	really	3	0	2.29
	always	2	0	2.65
	still	2	0	2.50
struggling	really	14	0	4.77
	so	10	0	2.80
	still	3	0	3.35
	over	0	2	2.99
crazy	really	14	0	4.77
	so	9	0	2.98
	like	3	2	2.00
suicidal stressed	so	6	0	2.01
	so	7	0	3.04
	really	4	0	3.71
	super	3	0	6.86
	more	2	0	3.34
worrying	constantly	6	0	6.94
	always	2	0	3.66
miserable	so	6	0	2.69
	more	2	0	3.48

Table 5 shows the collocate “so” is the most frequently used adverb that co-occurs with adjectives including “scared” (Freq: 43; MI: 3.35), “sad” (Freq: 16; MI: 3.10), “afraid” (Freq: 14; MI: 2.96), “sick” (Freq: 12; MI: 2.52), “crazy” (Freq: 9; MI: 2.98), “suicidal” (Freq: 6; MI: 2.01),

“stressed” (Freq: 7; MI: 3.04), “miserable” (Freq: 6; MI: 2.69). Functioning as a degree adverb, “so” directly modifies adjectives. In A&D discourse, it conveys overwhelming feelings, particularly in informal or spontaneous writing, as commonly observed in A&D subreddit posts. For instance, “so scared” or “so miserable” expresses a sense of being deeply affected, rather than mildly or moderately. It can also be easy to find that “so” is used to amplify the emotional force of the adjective that follows, expressing a strong emotional state or reaction. From the narratives of Reddit users’ posts of anxiety and depression, it is frequently used to heighten the emotional impact of personal experiences. The high frequency and notable MI scores for collocations such as “so scared” (Freq: 43; MI: 3.35), “so sad” (Freq: 16; MI: 3.10), and “so stressed” (Freq: 7; MI: 3.04) show that users often choose this structure to communicate intense psychological distress. Similarly, other adverbs such as “really”, “too”, and “extremely” serve the same function as “so”, used not just to describe emotions but to underscore the seriousness and intensity of the A&D challenges being shared. In this way, they play a key role in personal narratives, helping users communicate in depth about suffering, often when words alone may feel insufficient.

The adverb “constantly”, formed with the suffix “-ly”, functions as a frequency adverb, conveying the persistence or repetition of a particular emotional state over time in this context. “Constantly” is used to modify adjectives that describe negative mental states, such as “worrying” (Freq: 6; MI: 6.94), “afraid” (Freq: 4; MI: 4.86), “worried” (Freq: 4; MI: 5.34), “anxious” (Freq: 3; MI: 3.44), and “scared” (Freq: 3; MI: 3.05). It can be inferred that this collocational pattern reveals that users often experience these emotions not as isolated feelings but as recurring or uninterrupted states. For instance, when someone writes “*I’m constantly worried*”, the adverb highlights that the sense of worry is not temporary, but ongoing, perhaps even invasive. Notably, the adverb “more” co-occurs with five adjectival node words: “anxious” (Freq: 12; MI: 3.42), “afraid” (Freq: 3; MI: 2.85), “sad” (Freq: 3; MI: 2.98), “miserable” (Freq: 2; MI: 3.48), and “stressed” (Freq: 2; MI: 3.34). These collocations commonly occur in phrases such as “more anxious than before” and “getting more stressed,” which indicate a progression or worsening of an individual’s emotional state over

time. In these expressions, the adverb “more” signals an increase in emotional distress, often reflecting a gradual intensification or a response to increasingly challenging circumstances. For example, “I feel *more anxious* every day”, which implies a building tension or deterioration in my mental well-being. Unlike intensifying adverbs such as “so” or “extremely”, which denote the degree of emotion at a particular moment, “more” expresses temporal progression or comparison. Its recurrent use suggests that users perceive their emotional state as worsening over time, reflecting an escalating experience of anxiety and depression.

Many other adverbs co-occur with adjective-node words, with some showing notably strong lexical connections, as indicated by high MI values. For example, the adverb “super” is used alongside “stressed” (Freq: 3; MI: 6.86), emphasizing heightened emotional pressure often described by users. Likewise, “very” is frequently combined with “worried” (Freq: 5; MI: 4.16), emphasizing the intensity of concern expressed in posts. Such adverb-adjective collocations demonstrate how individuals in online discussions about mental health tend to strengthen their emotional descriptions, conveying a more profound sense of distress and urgency in their language. More importantly, adverbs “very”, “extremely”, and “incredibly” are strong intensifiers, often conveying heightened emotions, with “extremely” and “incredibly” indicating extreme intensity. “Deeply” typically conveys profound intensity or emotional depth, whereas “super” and “really” function as informal high-degree intensifiers that can modify both positive and negative adjectives. In contrast, “pretty” generally expresses a milder or more moderate degree of intensity. “Severely” indicates a negative, intense situation, while “lately” suggests a recent change, often with a slightly negative tone. “More” and “less” are neutral, as are “before”, “whenever”, and “until”, which provide time or contextual information. Most of these adverbs are neutral, focusing on intensity, frequency, or scale, with a few, such as “too” and “severely” leaning negatively.

IV. DISCUSSION

This study explored how Reddit users in the r/Anxiety and r/Depression subreddits described their emotional experiences through adjectives and their lexical collocates. Negatively valenced adjectives

were among the most common words in the data. Common examples include “anxious”, “panic”, “scared”, and “depressed”. The significance of these data is not only in the frequency of these terms, but in their morphological features. Adjectives ending in *-ed* (e.g., “scared”, “depressed”, and “exhausted”) were overwhelmingly used to describe deep-seated internal states resulting from an experience, while *-ing* adjectives (e.g., “worrying” and “struggling”) indicated ongoing, active sources of distress. Several expressions, including “stressed”, “emotional”, “crazy”, and “exhausted”, are shared across both subreddits. This pattern indicates that anxiety and depression often overlap in users’ experiences. The frequent overlap of these expressions suggests that anxiety and depression often co-occur. This may increase users’ emotional burden and reflect the interconnected nature of their psychological experiences.

The MI analysis shows that users express emotions through recurring word combinations. Adjectives are not used in isolation. Their emotional meaning depends on the words that surround them. Noun collocates predominantly represent clinical conditions and psychological states (e.g., “thoughts”, “attacks”, and “ideation”). The extraordinarily high MI score (10.88) for the collocation “suicidal ideation” demonstrates a specialized, clinical lexical relationship, stressing the severe mental distress and specific phrasing users adopt when discussing self-harm. The verb often appears with words such as “get,” “feel,” and “stuck.” These combinations describe negative emotional changes and lasting emotional states. They suggest that users experience emotions over time rather than as isolated events. Adverbial collocates are usually neutral in meaning, but they increase emotional intensity and emphasize frequency. Adverbs such as “so”, “extremely”, and “constantly” (e.g., “constantly worrying”, with an MI score of 6.94) point to the chronic, persistent, and overwhelming severity of users’ mental health struggles.

These linguistic-oriented findings firmly correspond to established psychological theories and recent literature (Kerz et al., 2023; Nguyen et al., 2025; Pavalanathan & De Choudhury, 2015). The frequent use of words expressing fear, anticipation of danger, and panic supports Beck and Clark’s Cognitive Theory of Anxiety. According to the theory, anxious individuals tend to focus

on potential threats and often interpret ambiguous situations as dangerous (Clark & Beck, 2009). The frequent use of words such as “depressed”, “hopeless”, and “miserable” also aligns with Beck’s Cognitive Theory of Depression. Beck suggests that depression is shaped by negative schemas, or deeply rooted patterns of distorted thinking, that contribute to hopelessness (Clark & Beck, 2009). The language used by Reddit users reflects these patterns in the way they describe their experiences. These findings align with those of Boettcher (2021) and Robertson et al. (2023), who noted the high prevalence of negative affective language in online mental health platforms. The raw intensity and candid self-disclosure observed in this corpus can be contextualized through the online disinhibition effect (Coppersmith et al., 2014; Kim et al., 2023; Nguyen et al., 2025; Suler, 2004). Because sites like Reddit provide anonymity and reduce traditional social cues, users feel enabled to express acute emotional vulnerability, such as detailing “suicidal thoughts” or “severe panic”, that they might otherwise hide in face-to-face clinical settings due to stigma.

A key contribution of this study consists in its distinction from recent computationally-based research. Although studies by De Choudhury et al. (2013) and Low et al. (2020) successfully utilize Reddit data for risk detection and emotion classification, they often treat words as discrete data points. The combination of MI scores and KWIC analysis enables the current study to move beyond simple classification by demonstrating that emotional nuance is constructed through the interaction between adjectives and their collocates. The study makes a theoretical contribution to corpus linguistics by showing how collocational analysis can uncover emotional meaning in naturally occurring mental health discourse. It shows that users actively construct a mutual emotional lexicon that blends clinical terminology (e.g., “disorder” and “ideation”) with visceral, personal articulations (e.g., “awful”, “mess”, and “stuck”). These outcomes also have practical implications for the development of digital mental health interventions and automated support tools. Current NLP tools that monitor digital media for mental health crises commonly rely on isolated keyword spotting. Incorporating collocational patterns (e.g., distinguishing between “feeling anxious”, “constantly anxious”, and “severe panic”)

might improve the sensitivity and contextual interpretation of automated digital mental health tools. The findings also have consequences for clinical practice. A better understanding of the language frameworks through which individuals express distress could help clinical practitioners interpret patients' online and offline narratives more accurately while fostering greater empathy in clinical interactions.

Despite its important findings, this study has several limitations. The analysis is based on data collected over a two-year period (2023–2025) from a single platform, Reddit. As a result, the findings may reflect the demographic and cultural characteristics of Reddit users and may not be generalizable to the general population. The data's anonymity also makes it impossible to verify users' clinical diagnoses, as the analysis relies entirely on self-reported symptoms. Another limitation is that the study centres exclusively on English-language data. The collocational patterns identified may therefore not extend to the linguistic construction of anxiety and depression in other languages or cultural spheres. These limitations emphasize the need for future research that draws on a more extensive range of data sources and research methods. Extending collocational analyses to other social media platforms, such as X (formerly Twitter) and TikTok, would help determine how different digital environments and platform affordances, including character limits and communication styles, shape the linguistic expression of emotional distress. Subsequent research should also include cross-linguistic corpus studies to provide a larger cross-cultural understanding of how emotional intensity is linguistically constructed through different languages. Longitudinal research would further explain how users' collocational patterns evolve over time and whether these language changes correspond to worsening symptoms or recovery. Future research could also examine other linguistic features, such as metaphors, syntactic structures, and discourse markers. This would provide a deeper understanding of how people communicate their emotions in online mental health spaces.

V. CONCLUSION

This study found that Reddit users in the *r/Anxiety* and *r/Depression* subreddits describe and intensify their emotional experiences using

recurring word combinations rather than single words. Emotional meaning in online mental health discourse emerges through recurring collocational patterns. These patterns give insights into how people express psychological experiences online. They also point out the importance of examining lexical relationships in digital mental health research. A predominance of negatively valenced adjectives, such as “anxious”, “scared”, “depressed”, and “suicidal”, was identified, with their emotional meanings consistently intensified by their collocational environments. Noun, verb, and adverb collocates contribute in different ways to the expression of emotional meaning. Nouns such as “attacks” and “ideation” place users' experiences in clinically significant contexts. Verbs such as “feel” and “get” reflect the developing and perpetual nature of emotional distress. Adverbs such as “constantly” and “so” strengthen the emotional intensity of users' self-disclosures. Adjectives such as “severe” reinforce the sense of persistence and exigency. The study illustrates the value of collocational analysis in digital mental health research. It also shows how analyzing collocational patterns in context can improve automated mental health detection and support more informed and empathetic clinical communication. The framework developed in this study can be used in future research. It brings together ideas from corpus linguistics, digital communication, and mental health.

ETHICS STATEMENT

We have read and complied with the ethical requirements for publication in *Jurnal Arbitrer*, and the current work does not involve human subjects or animal experiments.

CREDIT AUTHOR STATEMENT

Yang Yuhan: Data collection, data analysis, manuscript drafting, and revision.

Lau Su Kia: Supervision, research design, manuscript review, and revision.

Handoko: Methodology, data analysis, proofreading, and revision.

Yap Leng Lee: Proofreading, manuscript formatting, and revision.

Sheena Kaur: Editing, proofreading, and revision.

ACKNOWLEDGEMENT

This research did not receive any grant from the government or sponsors. AI was used for language editing and grammar checking.

DECLARATION OF COMPETING INTERESTS

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

AI USE STATEMENT

The authors used ChatGPT and Grammarly to assist with language editing, including improving grammar, clarity, readability, and sentence structure. All scientific content, analyses, interpretations, and conclusions were developed by the authors. The authors carefully reviewed, revised, and verified all AI-assisted suggestions and take full responsibility for the accuracy and integrity of the final manuscript.

REFERENCES

- Absattar, A., Mambetova, M., & Zhubay, O. (2022). The potential of emotive language to influence the understanding of textual information in media coverage. *Humanities and Social Sciences Communications*, 9(1), Article 232. <https://doi.org/10.1057/s41599-022-01232-2>
- Almeida, A., Patton, T., Conway, M., Gupta, A., Strathdee, S. A., & Bórquez, A. (2024). The use of natural language processing methods in Reddit to investigate opioid use: Scoping review. *JMIR Infodemiology*, 4, e51156. <https://doi.org/10.2196/51156>
- Anthony, L. (2022). What can corpus software do? In *The Routledge handbook of corpus linguistics* (pp. 103–125). Routledge. <https://doi.org/10.4324/9780367076399-9>
- Baumel, A., Fleming, T., & Schueller, S. M. (2020). Digital micro interventions for behavioral and mental health gains: Core components and conceptualization of digital micro intervention care. *Journal of Medical Internet Research*, 22(10), e20631. <https://doi.org/10.2196/20631>
- Bélignon, S. (2020). Feeling, emotion and the company they keep: What adjectives reveal about the substantives *feeling* and *emotion*. *Lexis*, 15. <https://doi.org/10.4000/lexis.4322>
- Boettcher, N. (2021). Studies of depression and anxiety using Reddit as a data source: Scoping review. *JMIR Mental Health*, 8(11), e29487. <https://doi.org/10.2196/29487>
- Brody, D. J., & Hughes, J. P. (2025). *Depression prevalence in adolescents and adults: United States, August 2021–August 2023* (NCHS Data Brief No. 527). National Center for Health Statistics. <https://doi.org/10.15620/cdc/174579>
- Clark, D. A., & Beck, A. T. (2009). *Cognitive therapy of anxiety disorders: Science and practice*. Guilford Press.
- Collins, L., & Baker, P. (2023). *Language, discourse and anxiety*. Cambridge University Press. <https://doi.org/10.1017/9781009250139>
- Coppersmith, G., Dredze, M., & Harman, C. (2014). Quantifying mental health signals in Twitter. In *Proceedings of the workshop on computational linguistics and clinical psychology: From linguistic signal to clinical reality* (pp. 51–60). Association for Computational Linguistics. <https://doi.org/10.3115/v1/w14-3207>
- De Choudhury, M., Gamon, M., Counts, S., & Horvitz, E. (2013). Predicting depression via social media. *Proceedings of the International AAAI Conference on Web and Social Media*, 7(1), 128–137. <https://doi.org/10.1609/icwsm.v7i1.14432>
- Edo Marzá, N. (2011). A comprehensive corpus-based study of the use of evaluative adjectives in promotional hotel websites. *ODISEA: Revista de Estudios Ingleses*, 12, 97–123. <https://doi.org/10.25115/odisea.v0i12.222>
- Everaert, J., Bronstein, M. V., Castro, A. A., Cannon, T. D., & Joormann, J. (2020). When negative interpretations persist, positive emotions don't! Inflexible negative interpretations encourage depression and social anxiety by dampening positive emotions. *Behaviour Research and Therapy*,

124, Article 103510. <https://doi.org/10.1016/j.brat.2019.103510>

- García-Méndez, S., de Arriba-Pérez, F., Beiro-Suso, J., & González-Castaño, F. J. (2026). Real-time anxiety and depression detection by combining large language models and machine learning with explainability capabilities on a user-centric, engaging conversational assistant. *Expert Systems*, 43(4), e70229. <https://doi.org/10.1111/exsy.70229>
- Gonda, X., Pompili, M., Serafini, G., Carvalho, A. F., Rihmer, Z., & Döme, P. (2015). The role of cognitive dysfunction in the symptoms and remission from depression. *Annals of General Psychiatry*, 14(1), Article 27. <https://doi.org/10.1186/s12991-015-0068-9>
- Gotlib, I. H., & Joormann, J. (2010). Cognition and depression: Current status and future directions. *Annual Review of Clinical Psychology*, 6(1), 285–312. <https://doi.org/10.1146/annurev.clinpsy.121208.131305>
- Grub, M. F. (2025). Reddit as a “safe space”: Topic modeling of online mental health communities for depression and anxiety. *Weizenbaum Journal of the Digital Society*, 5(3). <https://doi.org/10.34669/wi.wjds/5.3.3>
- Habicht, J., Viswanathan, S., Carrington, B., Hauser, T. U., Harper, R., & Rollwage, M. (2024). Closing the accessibility gap to mental health treatment with a personalized self-referral chatbot. *Nature Medicine*, 30(2), 595–602. <https://doi.org/10.1038/s41591-023-02766-x>
- Hunston, S. (2010). *Corpus approaches to evaluation*. Routledge. <https://doi.org/10.4324/9780203841686>
- Kerz, E., Zanwar, S., Qiao, Y., & Wiechmann, D. (2023). Toward explainable AI (XAI) for mental health detection based on language behavior. *Frontiers in Psychiatry*, 14, Article 1219479. <https://doi.org/10.3389/fpsy.2023.1219479>
- Kim, S., Cha, J., Kim, D., & Park, E. (2023). Understanding mental health issues in different subdomains of social networking services: Computational analysis of text-based Reddit posts. *Journal of Medical Internet Research*, 25, e49074. <https://doi.org/10.2196/49074>
- Knaevelsrud, C., & Maercker, A. (2007). Internet-based treatment for PTSD reduces distress and facilitates the development of a strong therapeutic alliance: A randomized controlled clinical trial. *BMC Psychiatry*, 7(1), Article 13. <https://doi.org/10.1186/1471-244X-7-13>
- Kovacs, M., & Beck, A. T. (1978). Maladaptive cognitive structures in depression. *American Journal of Psychiatry*, 135(5), 525–533. <https://doi.org/10.1176/ajp.135.5.525>
- Low, D. M., Rumker, L., Talkar, T., Torous, J., Cecchi, G., & Ghosh, S. (2020). Natural language processing reveals vulnerable mental health support groups and heightened health anxiety on Reddit during COVID-19: Observational study. *Journal of Medical Internet Research*, 22(10), e22635. <https://doi.org/10.2196/22635>
- MacLeod, C., & Cohen, I. L. (1993). Anxiety and the interpretation of ambiguity: A text comprehension study. *Journal of Abnormal Psychology*, 102(2), 238–247. <https://doi.org/10.1037/0021-843x.102.2.238>
- Mayopoulos, G., & Farber, B. A. (2023). Disclosure in psychotherapy versus in anonymous and non-anonymous online spaces. *Psychotherapy Research*, 34(5), 638–647. <https://doi.org/10.1080/10503307.2023.2256954>
- Morris, R. R., Schueller, S. M., & Picard, R. W. (2015). Efficacy of a web-based, crowdsourced peer-to-peer cognitive reappraisal platform for depression: Randomized controlled trial. *Journal of Medical Internet Research*, 17(3), e72. <https://doi.org/10.2196/jmir.4167>
- Nguyen, K. T. P., Vu, B., Chandna, S., Schultz, J., & Mayer, G. (2025). Between the lines: Investigating health beliefs and emotional expressions in online mental health communities. *Frontiers in Psychology*, 16, Article 1521623. <https://doi.org/10.3389/fpsyg.2025.1521623>
- Nguyen, T. Q., Yates, A., Zirikly, A., Desmet, B., & Cohan, A. (2022). Improving the generalizability of depression detection by leveraging clinical questionnaires. In *Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)* (pp. 8446–

- 8459). Association for Computational Linguistics. <https://doi.org/10.18653/v1/2022.acl-long.578>
- O'Dea, B., Batterham, P. J., Braund, T. A., Chakouch, C., Larsen, M., Berk, M., Torok, M., Christensen, H., & Glozier, N. (2025). A randomised cross over trial examining the linguistic markers of depression and anxiety in symptomatic adults. *npj Mental Health Research*, 4(1), Article 30. <https://doi.org/10.1038/s44184-025-00140-y>
- Pavalanathan, U., & De Choudhury, M. (2015). Identity management and mental health discourse in social media. In *Proceedings of the 24th International Conference on World Wide Web* (pp. 315–321). Association for Computing Machinery. <https://doi.org/10.1145/2740908.2743049>
- Prescott, J., Hanley, T., & Ujhelyi, K. (2017). Peer communication in online mental health forums for young people: Directional and nondirectional support. *JMIR Mental Health*, 4(3), e29. <https://doi.org/10.2196/mental.6921>
- Quillivic, R., Gayraud, F., Auxéméry, Y., Vanni, L., Peschanski, D., Eustache, F., Dayan, J., & Mesmoudi, S. (2023). *Towards unlocking the linguistic code of post-traumatic stress disorder: A comprehensive analysis and diagnostic approach* [Preprint]. Research Square. <https://doi.org/10.21203/rs.3.rs-3747882/v1>
- Robertson, A., Aiello, L. M., & Quercia, D. (2019). The language of dialogue is complex. *Proceedings of the International AAAI Conference on Web and Social Media*, 13(1), 428–439. <https://doi.org/10.1609/icwsm.v13i01.3241>
- Robertson, C., Carney, J., & Trudell, S. (2023). Language about the future on social media as a novel marker of anxiety and depression: A big-data and experimental analysis. *Current Research in Behavioral Sciences*, 4, Article 100104. <https://doi.org/10.1016/j.crbeha.2023.100104>
- Salekin, A., Eberle, J. W., Glenn, J. J., Teachman, B. A., & Stankovic, J. A. (2018). A weakly supervised learning framework for detecting social anxiety and depression. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, 2(2), 1–26. <https://doi.org/10.1145/3214284>
- Schnyder, N., Panczak, R., Groth, N., & Schultze-Lutter, F. (2017). Association between mental health-related stigma and active help-seeking: Systematic review and meta-analysis. *The British Journal of Psychiatry*, 210(4), 261–268. <https://doi.org/10.1192/bjp.bp.116.189464>
- Soriano, C. (2013). Linguistic theories of lexical meaning. In J. F. R. Fontaine, K. R. Scherer, & C. Soriano (Eds.), *Components of emotional meaning: A sourcebook* (pp. 63–80). Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780199592746.003.0005>
- Suler, J. R. (2004). The online disinhibition effect. *CyberPsychology & Behavior*, 7(3), 321–326. <https://doi.org/10.1089/1094931041291295>
- World Health Organization. (2022). *World mental health report: Transforming mental health for all*. <https://www.who.int/publications/i/item/9789240049338>
- Xiao, Y., Baes, N., Vylomova, E., & Haslam, N. (2023). Have the concepts of ‘anxiety’ and ‘depression’ been normalized or pathologized? A corpus study of historical semantic change. *PLOS ONE*, 18(6), e0288027. <https://doi.org/10.1371/journal.pone.0288027>
- Zarate, D., Ball, M., Prokofieva, M., Kostakos, V., & Stavropoulos, V. (2023). Identifying self-disclosed anxiety on Twitter: A natural language processing approach. *Psychiatry Research*, 330, Article 115579. <https://doi.org/10.1016/j.psychres.2023.115579>
- Zhang, T., Schoene, A. M., Ji, S., & Ananiadou, S. (2022). Natural language processing applied to mental illness detection: A narrative review. *npj Digital Medicine*, 5(1), Article 89. <https://doi.org/10.1038/s41746-022-00589-7>