Computers in Human Behavior 145 (2023) 107786

Contents lists available at ScienceDirect

Computers in Human Behavior

journal homepage: www.elsevier.com/locate/comphumbeh

Sharing is in fact about caring: Care concerns feature prominently in subreddits devoted to self-injurious thoughts and behaviors☆

Emma G. Preston[a,*], Suhaib Abdurahman[a], Diana W. Woodward[a], Amy E. West[a,b]

[a] *Department of Psychology, University of Southern California, 3620 S. McClintock Ave, SGM 501, Los Angeles, CA, 90089, USA*
[b] *Department of Clinical Pediatrics, Psychology, and Psychiatry & Behavioral Sciences, Keck School of Medicine, Children's Hospital of Los Angeles, 3250 Wilshire Blvd, Los Angeles, CA, 90010, USA*

ARTICLE INFO

Handling Editor: Catalina L Toma

Keywords:
Self-injurious thoughts and behaviors (SITBs)
Social media
Treatment
Stigma
Values

ABSTRACT

Objective: Given poor treatment efficacy and utilization for self-injurious thoughts and behaviors (SITBs) and the pervasive SITB-stigma present in many treatment settings, research that identifies values and experiences that resonate with individuals engaged in SITBs is urgently needed to improve current treatment offerings. The present study uses Moral Foundations Theory, a leading framework for conceptualizing values and behavior, to identify responses to SITBs most likely to resonate with those who have lived SITB experience.
Methods: Natural language processing methods (topic modeling, neural network-based classifier) were used to extract latent conversation topics and moral concerns from 1.68 M messages on the two largest SITB forums on Reddit. Once conversation topics and moral concerns were extracted, a linear regression model was fit to describe the relationship between likes on Reddit, moral concerns, and latent conversation topics.
Results: Findings revealed several types of messages most likely to resonate with individuals engaged in SITBs: 1) Specific situational narratives compared to general messages of sadness ($p < .01$); 2) messages that expressed care ($p < .001$); and 3) specific messages that expressed care, fairness, loyalty, and purity.
Conclusions: Specific, care-focused content (kind, nurturing content that discusses avoidance of emotional/physical harm to others) resonates most with individuals engaged in SITBs, providing insight to the real-time needs and experiences of those engaged in SITBs and suggesting the importance of framing SITB interventions in care language rather than success/failure language. Study findings may inform one-on-one clinical interactions, the development of SITB-specific interventions, and SITB training for clinicians.

1. Public health significance statements

When working with individuals who engage in self-injurious thoughts and behaviors, assessing online community use may provide important information to inform treatment needs specific to that individual.

This study indicates that clinical interventions for self-injurious thoughts and behaviors may be more effective when language emphasizes support and caring versus success and failure.

Study findings suggest moral, values-based content (specifically content that is care-focused) resonates strongly with individuals engaged in self-injurious thoughts and behaviors.

Self-injurious thoughts and behaviors (SITBs) cover a range of behaviors including suicidal ideation, planning, attempts, and non-suicidal self-injury (Nock, Joiner, Gordon, Lloyd-Richardson, & Prinstein, 2006). Approximately 17% of adolescents and young adults will engage in NSSI (Brown & Plener, 2017), which is the single most robust predictor of future suicidality (National Institute for Health Care Excellence, 2013). Although suicide is the second leading cause of death among adolescents in the United States (Miron et al., 2019), SITB treatment utilization and efficacy remain low—between 30-50% of individuals who engage in NSSI report that they have never received any kind of treatment and current estimates suggest that less than half of individuals engaged in suicidal behavior will seek treatment (Hom et al., 2015; Rowe et al., 2014).

While there are many hypotheses for low rates of SITB treatment

☆ **Author Note:** This study was not pregreistered. Requests for the dataset can be made to the first author and analytic code is available at https://github.com/goytoom/eSITB. We have no known conflicts of interest of funding sources to disclose.

* Corresponding author.
E-mail address: epreston@usc.edu (E.G. Preston).

https://doi.org/10.1016/j.chb.2023.107786
Received 18 November 2022; Received in revised form 3 April 2023; Accepted 10 April 2023
Available online 14 April 2023
0747-5632/© 2023 Elsevier Ltd. All rights reserved.

utilization, SITB stigma as a treatment barrier is highly supported in empirical literature (Long, 2017; Meheli & Saha 2022; Staniland et al., 2021; MacDonald et al., 2020). To effectively diminish the negative impact of SITB stigma and increase treatment utilization and efficacy, it is essential to understand underlying processes that facilitate development and maintenance of stigma. Moral foundations theory (MFT), a leading framework for conceptualizing and quantifying moral values, offers insight towards better understanding how stigma associated with SITBs developed and functions. According to MFT, moral values can be divided into five cross-cultural domains: (1) care/harm; (2) fairness/cheating; (3) loyalty/betrayal, to one's ingroup; (4) authority/subversion; and (5) purity/degradation, which includes disgust for both biological and spiritual contaminants (Graham et al., 2013, see Table 5 for a description of each value). The MFT framework has been increasingly used in clinical research to understand stigma attribution in traditionally stigmatized behaviors and disorders (e.g., substance abuse; Henderson et al., 2019) and to help researchers and clinicians identify treatment targets that may map onto values-based care (care that is congruent with patient needs, experiences, and values, e.g., purity concerns in OCD; Kang et al., 2016). This growing body of clinical MFT research extends to SITBs, and two MFT dimensions are especially useful for understanding the field's historic, ineffective response to SITBs in treatment: purity and care.

Purity. Because the purity dimension of MFT includes disgust for both spiritual and biological contaminants, it is a particularly relevant moral dimension for SITBs. Theoretical, philosophical, legal, and clinical precedent indicate that SITBs are viewed as both spiritually and biologically disgusting, particularly in Western society. It was not until 1983 that the Catholic church allowed funeral rites and burial in church cemeteries for individuals who died by suicide (Gearing & Lizardi, 2009), and in 1992, a clinician sued by the family of someone who died by suicide argued that since the deceased individual had completed an "immoral and illegal act", their family had no right to sue, and the Supreme Court of Virginia dismissed the case (Wackwitz v. Roy, 1992). Thus, despite the view of many that the United States is a secular society, an implicit belief in suicide as a spiritually disgusting act continues to inform societal processes. To systematically study how these societal processes impact perception of suicide, Rottman and colleagues designed several rigorous experiments in which they presented participants with identical obituaries except for cause of death (suicide or homicide). They found that most people implicitly associate suicide with purity concerns, regardless of their espoused spirituality. Participants were more likely to condemn suicide if they believed suicide tainted a soul, endorsed greater purity concerns in an independent questionnaire, or had elevated levels of disgust, providing further evidence of implicit social belief that suicide is spiritually wrong (2014).

Moreover, from a biomedical perspective, humans have developed to avoid pathogens and contaminants (Sarabian et al., 2018). A robust literature of experimental and clinical research indicates that humans are biased to associate blood and injury with aversive outcomes (Pury & Mineka, 1997). Blood and wounds (hallmarks of SITBs) are potential sources of infection that ancestors likely would have encountered, and thus over time, humans have developed adaptive disgust reflexes to protect themselves against potential negative health outcomes (Curtis & Biran, 2001). Thus, clinicians working with individuals engaged in SITBs must also overcome both implicit societal norms of suicide as spiritually wrong *and* evolutionary disgust associated with wounds, blood, infections, etc. Research indicates that many individuals including clinicians have a particularly difficult time overcoming these evolutionary disgust barriers with people engaged in SITBs because the wounds are self-inflicted and defy "basic" human instinct (Gunderson & Choi-Kain, 2019). Thus, despite being a highly vulnerable population at large, individuals engaged in SITBs are not desirable patients (Schoppmann et al., 2007).

Care. The care dimension of MFT is rooted in mammalian attachment systems and relates to humans' ability to feel (and dislike) pain in others (Haidt, 2012; Graham et al., 2013). According to MFT, care is associated with underlying virtues such as kindness, gentleness, and fairness (e.g., psychological "care" and medical "care" in colloquial language). However, when considering SITB prevention, some researchers and clinicians have been concerned that expressing care might actually reinforce SITBs (Carr et al., 1977; Lovaas, 1965; Lovaas & Simmons, 1969; Nock and Prinstein, 2004; Bergen et al., 2023). Much of the early research on SITBs was conducted in the context of the DSM-IV's "Factitious" category, with researchers and clinicians consistently expressing worry about "malingering" in patients engaged in SITBs (Fliege et al., 2002; Cummings et al., 2008; Willenberg, 1997). Today, SITBs are commonly viewed as manipulative, attention-seeking, and disruptive (Klonsky et al., 2014; Inckle, 2011; 2020; Ammerman et al., 2021). These stigmatizing assumptions continue to persist in clinical and research settings despite a robust body of literature indicating that emotional disorders (e.g., anxiety, depression), not diagnoses in the Disruptive, Impulse-Control, and Conduct Disorder category, are most comorbid with SITBs, and that SITBs most often serve affect regulation purposes rather than social reinforcement purposes (Bentley et al., 2021; Hooley & Franklin, 2018).

Fears of reinforcing SITBs, combined with patient knowledge of stigmas associated with SITBs, often drive SITB concealment and treatment delays. Research consistently demonstrates that individuals engaged in SITBs worry over how they will be viewed by their clinician and if their treatment will be impacted as a result (Fox et al., 2021; Mitten et al., 2016; Robinson, 2020). In recent qualitative work with those who self-injure, Long found that individuals engaged in self-harm wanted people to know that they are "not monsters … just really sad sometimes" (2017). Likewise, individuals engaged in SITBs frequently report that their clinicians do not understand their self-harm, and many report that they avoid disclosing SITBs to their provider over fears that support will be withdrawn, that they will be involuntarily hospitalized, and/or (for minors) that their parents will be involved (Fox et al., 2021; Mitten et al., 2016; Robinson, 2020). Individuals engaged in SITBs also report being shamed for their behavior, being made to wait longer for treatment, and being told that they are less deserving of treatment (than patients not presenting for SITB issues) (MacDonald et al., 2020). These experiences and expectations of stigma from health providers result in treatment avoidance, delay, and dropout (MacDonald et al., 2020; Staniland et al., 2021).

SITBs therefore sit at an interesting nexus of care and purity concerns in mental health spaces, in which providers often have strong disgust-driven negative bias against SITBs and worries that providing care may lead to SITB reinforcement. Moreover, to date, SITB interventions have been developed, tested, and implemented without input from those with lived experience in SITBs. But given the low rates of treatment presentation and efficacy, research indicates that the field's current conceptualization of SITB treatment, rooted in purity concerns and care fears, is incongruent with the lived experience and treatment needs of those engaged in SITBs (Beale, 2022; emmaontheedge, 2020; Fisher, 2022). Individuals with lived SITB experience have long called for treatments that are person-centered and that affirm the values of individuals engaged in SITBs (emmaontheedge, 2020; Inckle, 2020; Pembroke, 2002). Most treatments for SITBs (e.g., dialectical behavioral therapy, cognitive behavioral therapy) were originally developed for other symptoms, leading a growing number of researchers to consider how SITB-specific treatments may better meet the needs of those engaged in SITBs (Andover et al., 2017; Fox et al., 2020). However, SITB-specific treatments are in their infancy, and it is important that as these interventions are developed, they are guided by the values and experiences of those with lived experience in SITBs rather than by implicit disgust-driven anti-SITB biases and fears of reinforcing SITBs.

2. The current study

The current study seeks to center the voices and the values of those

engaged in SITBs by using MFT to analyze conversations in the two largest SITB e-communities on Reddit over a seven-year period. The field's treatment of SITBs has been ineffective, largely due to implicit disgust against SITBs and fears of inadvertently reinforcing SITBs, both of which manifest in the negative treatment of those engaged in SITBs, leading to negative treatment experiences, treatment delays, and even increased self-injury urges. In recent years, a growing number of researchers have begun to call for the creation of SITB-specific treatments that are congruent with the values and experiences of those engaged in SITBs. Research that identifies values and experiences that resonate with individuals engaged in SITBs is therefore urgently needed to inform this new wave of SITB interventions. SITB e-communities offer individuals engaged in SITBs a place to access support instantly and anonymously and offer researchers insight to the real-time experiences and needs of those engaged in SITBs (Preston & West, 2023). In this study, we used natural language processing methods (topic modeling, neural network-based classifier) to extract latent conversation topics and moral concerns from 1.7 M messages on the two largest SITB groups on Reddit. Once conversation topics and moral concerns were extracted, we fit a linear regression model to describe the relationship between likes on Reddit, moral concerns, and latent conversation topics. To our knowledge, no research has examined moral concerns underlying any type of SITB discussions nor examined how moral concerns interact with content of SITB discussions.

3. Methods

Fig. 1 shows a visualization of the research pipeline used in this study. First, we collected and cleaned text data from the two largest suicide/self-harm subreddits over a seven-year period. Next, we used latent Dirichlet allocation (LDA; Blei et al., 2003) to determine latent topics of conversation in these text data. Then, we used a neural network based language model and classifier (small BERT; Turc et al., 2019) to identify moral concerns present in these text data. Finally, we fit linear regression models to describe the relationship between likes on Reddit, moral concerns, and latent conversation topics.

31. Transparency and openness

We report how we determined our analytic sample, all data exclusions, and all manipulations used in the study. There are no measures to report. The dataset for this study is not publicly available to maximize privacy of these SITB e-communities and to comply with internal review board protocol for this study. We have no previously published or currently in press works stemming from this same dataset. Analytic code used in these analyses is available at: https://github.com/goytoom/eSITB. Data were analyzed using python and R, see below for specific packages used. This study's design and its analysis were not preregistered.

3.2. Data

Reddit. Reddit is a heavily trafficked social media platform that has >430 million active users and billions of site visits (Dean, 2021). Reddit allows individuals to view, subscribe, and contribute to a variety of subreddits, or content-specific chats. In subreddits, individuals have the option to make their own posts, comment on the posts of other users, and "upvote" and "downvote" other content. Content that receives many upvotes are indicative of general community support whereas content that receives many downvotes signals community rejection. Content that receives many upvotes automatically gets moved to the top of the subreddit homepage thread, making them the first thing individuals see when logging on to the site.

Data collection. All posts and all comments (collectively referred to as "messages") during the period of 1/1/2014–6/1/2021 were collected from the largest suicide (>400,000 members) and self-harm (>100,000 members) subreddits, via Pushshift (Baumgartner et al., 2020). For each post and comment, we retrieved the time point of the messages, the

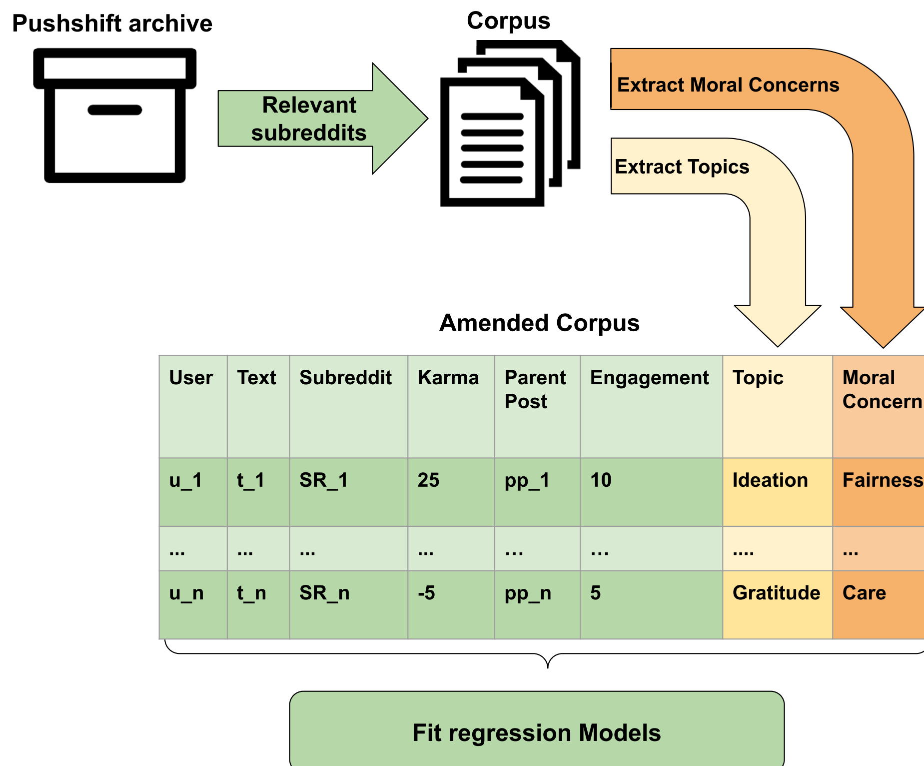

Fig. 1. Research pipeline utilized in this study
1) gather the messages from the two subreddits via the Pushshift archives
2) extract the relevant information (e.g., texts, karma scores, etc.)
3) extract topics and moral concerns using trained machine learning models and
4) fit a linear regression model on our data predicting karma score by message topic, expressed moral concern and control variables.

"karma" count (difference between upvote and downvotes, informally, "likes"), the user's id and the id of the subreddit in which the message was published. For all comments, we also retrieved the id of the original post to match posts and subsequent responses together. Both subreddits are heavily moderated to keep the contents of the subreddits focused on peer-support and peer-exchange. The largest suicide subreddit requires content to be non-judgmental peer-support and any abuse/"tough love", pro-suicide messages, religious proselytizing and trolling are forbidden and removed. On the largest self-harm subreddit, no glorification of self-harm or any demeaning/triggering messages are permitted. Thus, the collected messages from these two subreddits are robust depictions of peer-conversations in the SITB community. All analytic procedures were approved by the authors' institutional review board.

Data cleaning and preprocessing. First, we aggregated all messages and their information (e.g., username, karma, etc.) into a unified dataset. For all posts, we combined the title and the body into one text. We then removed all stopwords (e.g., "and", "with") based on the NLTK (Loper & Bird, 2002) and gensim (Rehurek & Sojka, 2012) libraries in python. We further removed any punctuation and symbols from the texts, including URLs and emails. Any message that did not have at least five words after cleaning were removed, to decrease noise in the model. The final data set used in the analyses below contained 1,688,280 messages.

3.3. Topic modeling

We used latent Dirichlet allocation (LDA) to determine the topics of peer-conversations on SITB e-communities. LDA is a popular topic modeling technique used to extract latent topics from a wide range of text data that would be impossible to discern via manual inspection (e.g., Curiskis et al., 2020). LDA is a hierarchical Bayesian model that represents each text in a corpus (or each comment/post on Reddit) as a mixture of latent topics, with each topic itself being modeled as a mixture of underlying words (Blei et al., 2003). Using statistical contextual metrics, LDA groups words into topics and uses unsupervised learning to determine the topic distributions from the corpus itself (Blei, 2012). LDA uses the statistical co-occurrence of words in a text corpus to group words with connected meanings into topics and then determines which topics are in a given text based on the occurrence of the respective "topic words" in the text.

In the current study, the text from all collected posts and comments were combined into a unified corpus of SITB Reddit data. We used the gensim library in python to train the LDA model (Rehurek & Sojka, 2012). In LDA, the number of topics must be prespecified, and models with different topic numbers are trained and evaluated to determine the optimal model. We trained topic models for a range of topic numbers (2–30) and evaluated the models using both a mathematical and qualitative approach. First, we used the normalized pointwise mutual information coherence score (c_{NPMI}) to evaluate models with higher c_{NPMI} scores indicating a higher likelihood of the topics being interpretable to a human (Bouma, 2009; Newman et al., 2011). The c_{NPMI} indicated that models with 26–30 topics were most likely to be interpretable to a human. Next, a clinician trained in SITB research reviewed output from each of the 26–30 topic models and determined that the 26-topic model provided the most clinical insights while keeping parsimony. Once the final model was determined, we used it to identify the most probable topic for each message and the keywords and exemplary messages for each topic.

3.4. Moral concern extraction

To identify underlying moral concerns in the SITB conversations on Reddit, we trained a neural network-based classifier based on the messages' text embeddings. Embeddings are quantitative representations of language that allow computational models to perform subsequent tasks that depend on meaning (for a full discussion of embeddings, see Mikolov et al., 2013). Modern embedding models are usually deep contextual neural networks, trained on large-scale language data and have been used to predict moral language on Facebook and even train chatbots to administer behavioral therapy for depression (Pola & Sheela Rani Chetty, 2021). We used a "BERT"-based (Bidirectional Encoder Representations from Transformers, Devlin et al., 2019) model to determine the moral concerns in each Reddit message using the message's text as input. See Fig. 2 for a structural visualization of a BERT model. Specifically, we employed a pre-trained BERT model called "small BERT" (Turc et al., 2019) and added a downstream classification layer to the language model to predict whether a message contained moral vs non-moral language, and whether moral messages express the five MFT values: care, fairness, loyalty, authority, or purity concerns. We simultaneously trained the classification layer and fine-tuned the embedding layers on the Moral Foundations Twitter Corpus (Hoover et al., 2020), a large, annotated corpus containing over 35,108 tweets and each tweet's moral framing according to Moral Foundations framework (Graham et al., 2013). The classifier achieved a cross-validated F_1 score of 0.84 for moral/non-moral message classification and 0.71 when predicting the actual foundations. We applied the trained model on all collected reddit messages to determine whether each message contained a moral concern and whether they explicitly expressed care, fairness, loyalty, authority, or purity concerns. Messages with moral concerns that could not be assigned to any foundation, due to a lack of clear reference/markers, were labeled as "Thin-Moral" (see Trager et al., 2022 for a detailed explanation of "Thin-Morality").

3.5. Statistical analysis strategy

Once we identified latent conversation topics and moral content within the dataset, we fit a linear regression model to the data using the "stats" library in R (R Development Core Team, 2010) to predict "karma" (likes) based on message topic, moral concerns expressed in the messages, and the interaction of topic and moral concern for each Reddit message. Since Reddit determines which posts to show "at the top" of a subreddit based on initial engagement of a new post (i.e., number of comments), we controlled for the number of comments under a post (or under the parent post for comments) to reduce bias due to post visibility. To mitigate bias from different post formats we also controlled for message length (word count). The final data set consisted of 1,688,280 messages and included "karma" (likes), message topic, moral concerns (binary variable for each of the five MFT foundations), message text, message word count and number of comments (for comments, we used the number of comments under the parent post). For the message topics, we set the reference level to "negative thoughts" as general negative thoughts are consistent across SITB engagement. Setting "negative thoughts" as the reference level allowed us to examine if other, potentially more specific, message topics resonate more (or less) with those in these SITB e-communities. For the moral concerns variable, we set the reference value to zero, meaning that they were compared against messages that do not express moral concerns. Lastly, we also checked whether the inclusion of topics and moral concerns significantly adds to explainability of community resonance by fitting a series of models while incrementally adding the predictors. The reference model (M0) only contains the control variables. M1 adds message topics, M2 adds moral concerns, M3 adds message topics and moral concerns and M4 adds message topics, moral concerns and their interaction to the baseline (see Table 7). The following command was used to fit the full model in R: *lm(score ~ Topic_Name * (care + fairness + loyalty + authority + purity) + n_words + engagement)*.

4. Results

4.1. Topics

See Table 1 for a summary of all topics, including topic name,

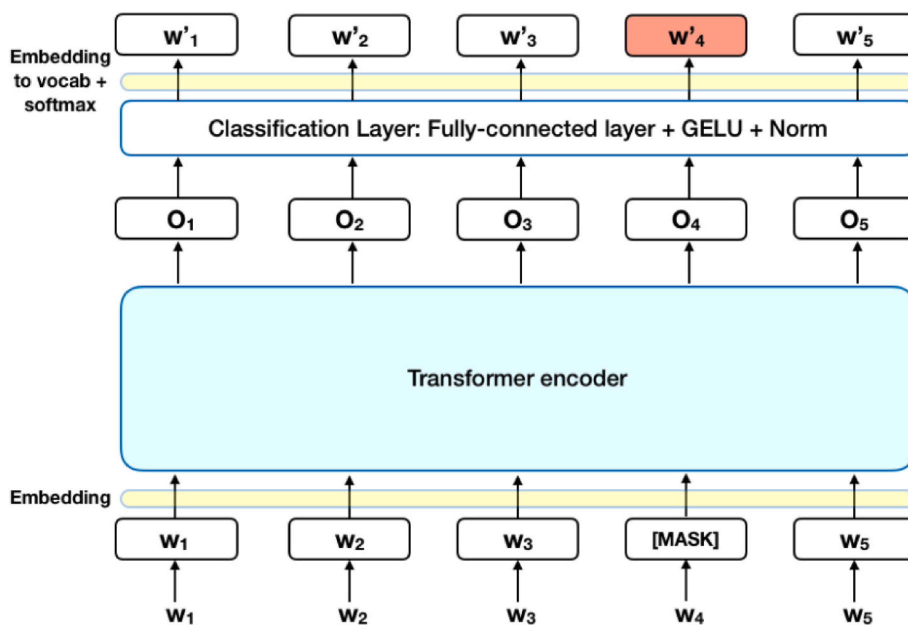

Fig. 2. High-level structural representation of a BERT embedding model (Horev, 2018).
The model is trained on a large-scale language dataset to predict randomly masked words in the training data based on the context provided by the remaining words (semi-supervised). The inputs are first encoded, embedded, and then fed into a neural network (classification) layer for the prediction task. The model then fits the embeddings according to the prediction goal. Thus, after training, the model can create context sensitive embeddings of language inputs. In our study, we added an additional classification layer to the BERT model to fine tune the model to moral concerns using an annotated corpus of moral messages (MFTC; Hoover et al., 2020).

frequency, keywords, and a shortened exemplary message from the corpus. The most frequently discussed topic was "Ideation" (11.5%), which deals primarily with general suicide ideation. Two other topics ("Immediate Help Seeking" and "Describing Suicide") also contain themes related to ideation, and these themes most often occurred in specific contexts. For instance, "Immediate Help Seeking" usually occurred contexts that evoked urgency and sympathy, while "Describing Suicide" often captured discussions of suicidal ideation alongside concrete suicide plans.

Other frequent topics included: 1) "Negative Thoughts" (5.5%), which broadly centered around any negative thoughts but especially negative, intrusive, and depressive thoughts; 2) "NSSI experiences/urges" (6.0%), which includes details of past NSSI experiences and urges (e.g., cutting oneself and the consequences thereof); 3) and "Wishing Well" (5.7%), in which users sympathize and encourage each other on their path to recovery. Overall, the topics ranged in content and included discussions of SITB engagement, personal and familial relationships (e.g., personal, familial, romantic), academic and work experiences, NSSI experiences, and professional care and immediate help seeking. It is important to note that all topics meaningfully related to SITB engagement: conversations in these SITB e-communities are focused and specific and allow users a place to discuss the complexities of SITB engagement and environmental triggers and access and offer support. Table 6 further shows, for each possible topic in a post, the most frequent topic in its responses (comments). Most posts (19 out of 26) are responded to with their original post topic, indicating that our modeling approach was successful in capturing meaningful topics that are consistent across conversations. For the seven exceptions, the most frequent topics in the responses were all related topics that were reasonable responses in that context. For instance, "Family Abuse", "Immediate Help Seeking", and "Loved Ones" (usually about issues with loved ones) were mostly responded to with "Wishing Well".

4.2. Moral concerns

Table 2 shows a summary of the moral concerns containing a definition, exemplary message from the corpus and the frequency of occurrence in the corpus for each type of moral concern. In total 42.8% of messages contained a moral concern. 15.4% of the messages were classified as "thin moral", indicating the presence of unspecified moral content. Among all messages with identified moral concerns, 93.7% contained care, 3% fairness, 3.3% loyalty, 1.2% purity, 0.2% authority, concerns. Fig. 3 shows the distribution of moral concerns over the corpus for both posts and comments. The care concerns expressed related to a broad range of messages but often appeared in relation to advice on how to avoid physical and emotional pain or with sharing experiences that are causing physical and emotional pain. Fairness concerns were often expressed in messages dealing with reliving/reflecting on past negative experiences, such as (sexual) assault, or thinking about the future. Loyalty concerns centered around messages relating to interpersonal relations especially familial relationships and community. Authority concerns were mostly expressed in messages relating to issues with authorities, rebelling against authority figures (especially within family) and politics. Finally, purity concerns were mostly expressed in messages relating to religious experiences or bodily issues.

4.3. Statistical analysis

The model comparison, depicted in Table 7, shows that topics, moral concerns, and the interactions each add significantly to prediction of community resonance (defined as by "likes"). Thus, results discussed below reflect the full model as outlined in the analysis strategy section. Tables 3 and 4 show the summary of all model coefficients. For the message topic variable, results indicated that most topics receive significantly more "likes" compared to "Negative Thoughts" when controlling for all other variables. The only topics that did not have statistically significant main effects were "Academic Concerns", "Society/Philosophy", and "Wishing Well". The strongest significant main

Table 1
Distribution of Topics, Keywords, and Examples (note examples are edited to preserve anonymity).

Topic (%)	Keyword	Example
Ideation (11.5)	life, live, die, end, think, even, kill, anymore, never, reason	I gotta die I gotta die I gotta die …
Seeking Change (6.8)	life, thing, time, make, try, think, well, change, work, good	I'm stuck vicious cycles. I was happy and ok, then some bad things happened to me. I got through them. But then I really screwed up
Talking to Friends (6.0)	tell, say, talk, think, really, friend, sorry, make, ask, thing	I got hammered and told my ex the world would be better without me. They told my friends. I'm pissed, but don't know if I should be
NSSI Experiences/Urges (6.0)	cut, clean, deep, stop, really, time, make, bad, start, arm	… I cut myself with scissors for the first time on Friday. It was on my left upper thigh, below my underwear line …
Wishing Well (5.7)	hope, well, good, try, find, make, help, happy, hard, stay	I get it. I, tried to get help and couldn't find what I needed. I do still believe in a future for us although we don't know each other so, please, if you're still here, keep trying …
People Care (5.7)	people, care, help, think, good, say, make, really, try, thing	It's people like you that really make this world a better place. And I love people like you because you're the underdog …
Negative Thoughts (5.5)	help, thought, think, suicidal, way, also, thing, experience, feeling, depression	It could also be a sort of intrusive thought, it's harmless but can be unsettling. If we really are against something, we're more likely to get intrusive thoughts of us doing that thing
Immediate Help Seeking/ Offering (5)	need, help, delete, talk, kill, chat, free, message, pm, tonight	Help me Help me Help me Help me
Gratitude (5)	thank, write, read, well, make, really, day, glad, thing, good	Just wanted to say thanks for taking time out of your days to respond. I've been feeling better the last couple of weeks.
Professional Care (4)	help, therapist, mental, call, therapy, doctor, hospital, health, try, see	If you're about to die, then what happens is you call 911 and those specifics aren't more important than the fact that *it can bring you closer to getting the help that you need* …
Society & Philosophy (3.7)	life, people, world, make, human, live, believe, mean, thing, way	What the human race sees as natural is totally irrelevant to nature as a whole. You think the stars give a shit about what we insist upon be the truth …
Suicide Methods/Attempts (3.5)	sleep, take, day, eat, tired, time, try, pill, drink, wake	Three days ago I took about 50 pills of OTC super strong pain relievers with Acetaminophen (it was Walgreens brand, not Tylenol) and two days ago I took another 50.
Employment/Financial Experiences (3.2)	job, work, money, pay, live, find, even, make, time, try	I was doing great in my new city. I had a job at an firm right out of college for 14 months then I was laid off …
Positive Life Experiences (3)	school, play, game, friend, video, high, make, people, fun, really	You're interested in a real addiction that could actually bring you joy? Try gaming or music. You will be shocked at the friends you will make …
Reflections (2.9)	thing, think, come, time, walk, day, back, see, look, moment	I dreamt i was shot in the head by a friend i love. I kept trying to open my eyes and they'd open and everyone stared at me like i was crazy …
Romantic Experiences (2.7)	love, year, girl, old, guy, date, life, still, time, woman	no one will ever love me no one will ever love me no one will ever love me
Loved Ones (2.4)	mom, home, brother, dog, sister, family, room, leave, house, cat	my dad got drunk again. so i have to stay home and watch my sister so my mom can go get my brother from school. My dad is an alcoholic but he hasn't gone to a meeting in 8 months …
Academic Concerns (2.4)	school, college, fail, year, high, class, study, grade, suicide, think	I remember when i was in college. Exams really do suck. They really really suck
NSSI Concealment (2.3)	self, harm, scar, wear, see, cut, hide, ask, long, cover	… I was just wondering is it inappropriate to wear a swimsuit that shows my scars around them
Positive Distractions (2.3)	listen, music, day, cry, really, watch, sometimes, good, time, thing	Nirvana is pretty good, even though they're well known. It can be fun to analyze the lyrics in some of their songs
Describing Suicide (2.1)	pain, brain, way, death, die, think, take, try, damage, painful	I want to die too. Two nights ago I partial suspension hanging but I couldn't get myself to faint at all. Then I pressed down on my carotid artery a bunch but I still couldn't faint.
Positive Relationships (2)	friend, people, love, person, alone, relationship, think, even, make, good	The One is mythological, there's NOT just one ideal person for everyone. Humans can be happy and satisfied with lots of different people and in lots of different relationships …
Family Abuse (1.8)	parent, kid, child, mother, abuse, family, year, wife, father, time	My exwife abused me and i dated her for 6 years before we married …
Negative Emotionality (1.2)	fuck, shit, man, fucking, dude, life, stupid, even, piece, suck	I'M JUST DAMN EMPTY AND SO FUCKING NUMB
Body Image Concerns (1.2)	look, die, body, ugly, face, even, think, make, weight, see	NOOOO. I'M HIDEOUS AND UGLY SO UGLYYYYYY
Talking with Others (1)	talk, hate, friend, people, like, _, much, make, family, help	Are you doing alright? If you're not here anymore, then I'm sorry. If you are still here, then here's a hug from me

effects were as follows: "NSSI Concealment" ($b = 8.3, p < 0.001$), "NSSI Experiences and Urges" ($b = 5.6, p < 0.001$, and "Negative Emotionality" ($b = 4.15, p < 0.001$). For the moral concerns variable, the only significant main effect was for care concerns ($b = 2.63, p < 0.001$) when controlling for all other variables, indicating that care concerns resonate more than other moral concerns within these SITB communities.

Findings also included several significant interaction effects (see Table 5 for a detailed overview of the significant interaction effects including an exemplary message from the corpus for each interaction). The presence of care concerns statistically increased the likelihood of a message receiving increased karma in the following topics: "Negative Emotionality" ($b = 4.91, p < 0.001$), "Gratitude" ($b = 4.06, p < 0.001$), "Talking with Others", ($b = 2.65, p < 0.01$) "Academic Concerns" ($b = 2.1, p < 0.001$), "Positive Distractions" ($b = 1.8, p < 0.01$), "Body Image Concerns" ($b = 1.7, p < 0.01$), and reduced it for: "Positive relationships" ($b = -.85, p < 0.05$), "Seeking Change" ($b = -1.2, p < 0.001$), "Wishing Well" ($b = -1.4, p < 0.001$), and "NSSI Experiences/Urges" ($b = -1.4, p < 0.001$).

For the remaining moral foundations, the model produced fewer significant interactions. Fairness interacted with both "Talking to Others" ($b = 22.3, p < 0.001$) and "Describing Suicide" ($b = 14.1, p < 0.01$) to increase likes. Loyalty interacted with both "NSSI Experiences/Urges" ($b = 19.6, p < 0.001$) and "Gratitude" ($b = 9.5, p < 0.001$) to increase likes. Purity interacted with both "Talking with Others" ($b = 21.6, p < 0.001$) and "Describing Suicide" ($b = 11.3, p < 0.01$) to increase karma. There were no significant interactions between authority and any of the message topics.

Table 2
Distribution, Definitions, and Examples of Moral Values (note examples are edited to preserve anonymity).

Moral Concern	Frequency (% of Moral Messages)	Definition	Exemplary Message
Care	25.4% (93.7%)	Intuitions about avoiding emotional and physical damage to another individual. It underlies virtues of kindness, gentleness, and nurturing.	It's there on the resources page, under helplines. The Samaritans are one of the most popular helplines for distressed people (anyone can talk to them, not only suicidal people). They operate in the UK but people from other countries can email them. But remember: they are counselors, not therapists. So they'll listen but they won't give their own personal input or advice …
Fairness	0.8% (3%)	Intuitions about fair treatment and outcome for all individuals and groups. It underlies virtues of social justice and equality.	I could have false hope that i'd get a good job or win the lottery and my parents would be really sad but i think they would get it. I'm just 19 but i have the mind of an old person. Im really sick of the human race and how unfair it is.
Loyalty	0.9% (3.3%)	Intuitions about cooperating with ingroups and competing with outgroups. It underlies virtues of patriotism and self-sacrifice for the group.	… yo what you just put out was an act of community. We need more of that in this group. Community brings people back and holds us all together
Authority	0.1% (0.2%)	Intuitions about deference toward legitimate authorities and high-status individuals. It underlies virtues of leadership and respect for tradition.	… When I was 19 my way of getting more control was by rebelling and dating a lot at university. I probs can't recommend that since it didn't end great for me lol. Try just pushing the boundaries a bit, the little wins that get taken for granted can help you feel more in control …
Purity	0.3% (1.2%)	Intuitions about deference toward legitimate authorities and high-status individuals. It underlies virtues of leadership and respect for tradition.	I get what you're describing. I have a binging problem but it's new. I know what youre talking about in that feeling afterward where you feel gross and disgusting like 20 min later. Ive been making myself vomit after but im bad at it. I feel disgusting …
Thin-Morality	15.4%	Moral content without clear or only weak markers of any intuition	Basically can't do even the easiest things right. I always end up screwing things up more.
None	57.1%	Absence of the aforementioned intuitions	I don't think I'm that speedy! But I'll try

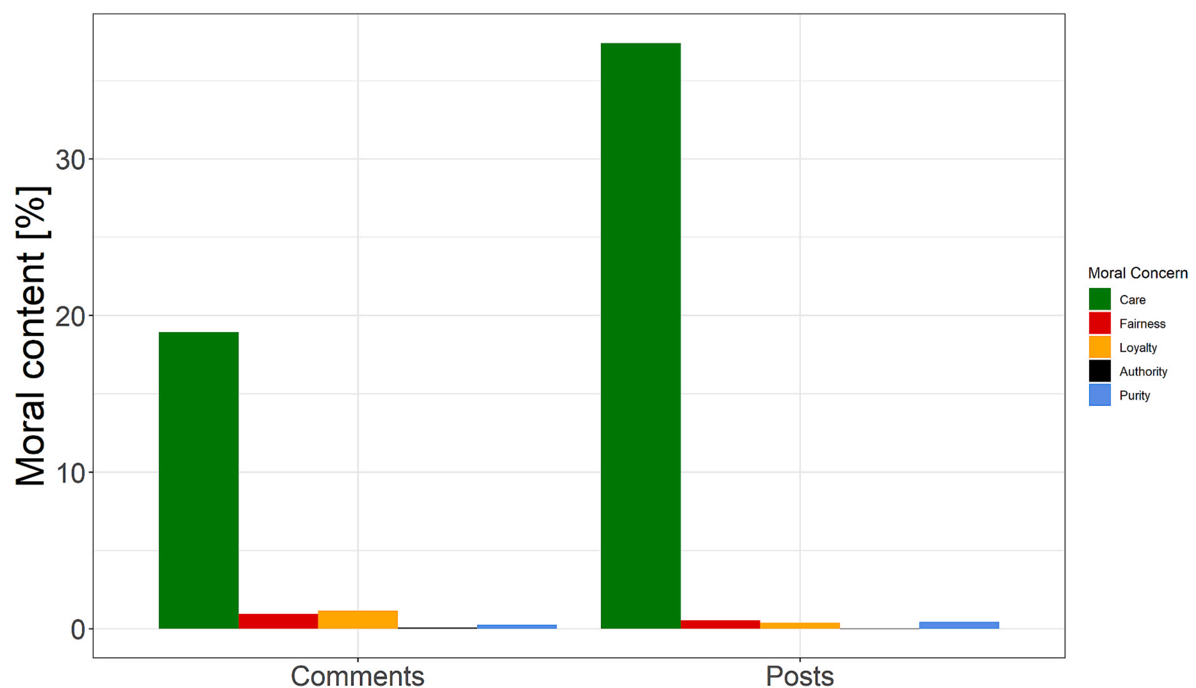

Fig. 3. Distribution of moral concerns across all messages.

Table 3
Main effects.

Variable	Beta	95% CI	p-value
Negative Thoughts	–	–	–
Academic Concerns	0.31	−0.09, 0.71	0.12
Body Image Concerns	2.2	1.7, 2.7	<0.001
Describing Suicide	2.9	2.5, 3.4	<0.001
Employment/Financial Experiences	−0.4	−0.76, −0.04	0.03
Family Abuse	0.92	0.39, 1.4	<0.001
Gratitude	0.86	0.55, 1.2	<0.001
Ideation	3.1	2.8, 3.4	<0.001
Immediate Help-Seeking/Offering	2.8	2.5, 3.1	<0.001
Loved Ones	3.4	3.0, 3.8	<0.001
Negative Emotionality	4.2	3.7, 4.7	<0.001
NSSI Concealment	8.3	7.9, 8.8	<0.001
NSSI Experiences/Urges	5.6	5.3, 5.9	<0.001
People Care	0.57	0.25, 0.89	<0.001
Positive Distractions	1.2	0.79, 1.6	<0.001
Positive Life Experiences	−0.8	−1.2, −0.40	<0.001
Positive Relationships	0.32	−0.04, 0.68	0.085
Professional Care	1.5	1.1, 1.8	<0.001
Reflections	1.1	0.71, 1.4	<0.001
Romantic Experiences	1	0.65, 1.4	<0.001
Seeking Change	−0.3	−0.59, 0.00	0.051
Society & Philosophy	0.08	−0.29, 0.45	0.7
Suicide Methods/Attempts	1.9	1.5, 2.2	<0.001
Talking to Friends	1.7	1.4, 2.0	<0.001
Talking with Others	0.61	0.08, 1.2	0.024
Wishing Well	0.09	−0.23, 0.41	0.6
Care	2.6	2.2, 3.1	<0.001
Fairness	−0.22	−2.6, 2.2	0.9
Loyalty	0.5	−1.5, 2.5	0.6
Authority	−0.21	−6.6, 6.2	>0.9
Purity	−0.79	−7.2, 5.6	0.8
Number of Words	0.01	0.01, 0.01	<0.001

5. Discussion

To the best of our knowledge, this is the first study that examines moral concerns underlying SITB discussions and examines how moral concerns interact with content of SITB discussions. Our results indicate that users discuss a variety of topics in these SITB e-communities, ranging from general offerings of support and disclosures of sadness to conversations of attempts to highly specific conversations about self-harm, body image, work concerns, and family. These results are consistent with past research into SITB e-communities demonstrating that users seek out a variety of support and information in these spaces (Lavis & Winter 2020). Compared to general messages expressing sadness, messages with more specific topics (e.g., family abuse, ideation, discussions of loved ones, etc.) received significantly more likes. One possible explanation for this finding is that general negative thoughts and sadness are baseline shared experiences in this group, and members prefer and engage more with specific discussions of symptoms and experiences.

The "NSSI Experiences and Urges" topic and "Negative Emotionality" topic had the strongest main effect in predicting likes. Messages in the "NSSI Experiences and Urges" topic included detailed discussions of self-harm (e.g., methods, urges) and strategies for wound care. Similarly, the next strongest main effects in predicting likes were found for "Ideation", "Describing Suicide", and "Immediate Helpseeking/Offering", which include detailed, vivid, and graphic descriptions of suicidal, attempt, and aftermath, as well as in-the-moment sharing and solicitation of advice. These findings suggest that discussion of the complex issues surrounding SITBs, is valued by members of SITB communities. Messages in the "Negative Emotionality" topic tended to be raw and unfiltered expressions of suffering. Often filled with explicit language, these messages vividly describe in-the-moment experiences of psychological pain. Interestingly, the most frequent type of response to a "Negative Emotionality" post was a message in the "People Care" topic. Thus, it appears that not only do members of these SITB e-communities value the expression of psychological suffering, but they also seek to alleviate that suffering by offering specific expressions of support in the comments. Similar relationships were found between several post and comment topic pairs, such as "Family Abuse" and "Wishing Well" or "Ideation" and "Seeking Change", and provide quantitative evidence to support anecdotal and qualitative evidence that people experiencing suicidal ideation and/or engaging in SITBs use SITB e-communities to access support (Brown et al., 2020). They also demonstrate that individuals engaged in SITBs do not use e-communities to simply share general baseline experiences and have them validated, but rather that

Table 4
Interaction effects.

Topic Name	Care Interaction			Fairness Interaction			Loyalty Interaction			Purity Interaction		
	Beta	95% CI	p-value	Beta	95% CI	p-value	Beta	95% CI	p-value	Beta	95% CI	p-value
Academic Concerns	2.1	1.2, 3.0	<0.001	0.78	−3.7, 5.2	0.7	2.1	−4.2, 8.4	0.5	−1	−18, 16	>0.9
Body Image Concerns	1.7	0.56, 2.9	0.004	−0.08	−6.1, 5.9	>0.9	5.1	−5.7, 16	0.4	2.7	−4.2, 9.6	0.4
Describing Suicide	−0.02	−0.80, 0.75	>0.9	22	15, 30	<0.001	−4.7	−17, 7.8	0.5	11	3.3, 19	0.005
Employment/Financial Experiences	0.4	−0.45, 1.2	0.4	−0.17	−3.2, 2.9	>0.9	−1.2	−4.4, 2.0	0.5	7.7	−5.8, 21	0.3
Family Abuse	−0.39	−1.2, 0.43	0.4	−0.68	−3.9, 2.5	0.7	−2.3	−5.8, 1.3	0.2	2.9	−5.8, 12	0.5
Gratitude	4.1	3.1, 5.0	<0.001	−0.56	−5.7, 4.6	0.8	9.5	6.4, 13	<0.001	5.3	−7.5, 18	0.4
Ideation	−0.17	−0.70, 0.36	0.5	1.4	−1.8, 4.6	0.4	−2	−5.2, 1.3	0.2	1.6	−5.1, 8.3	0.6
Immediate Help-Seeking/Offering	0.56	−0.12, 1.2	0.11	1.1	−4.4, 6.6	0.7	−0.21	−3.6, 3.1	>0.9	2.4	−8.9, 14	0.7
Loved Ones	−0.14	−0.94, 0.66	0.7	−2.6	−7.6, 2.4	0.3	−3.7	−8.7, 1.3	0.15	2.5	−7.6, 13	0.6
Negative Emotionality	4.9	3.4, 6.4	<0.001	3.4	−1.6, 8.3	0.2	−4	−15, 7.0	0.5	6.3	−1.1, 14	0.1
NSSI Concealment	−0.59	−1.4, 0.17	0.13	−1.1	−7.1, 5.0	0.7	3.4	−5.3, 12	0.4	8.5	−2.8, 20	0.14
NSSI Experiences/Urges	−1.4	−2.1, −0.76	<0.001	0.53	−5.2, 6.2	0.9	20	13, 26	<0.001	5.4	−4.0, 15	0.3
People Care	−0.62	−1.3, 0.02	0.056	0.41	−2.6, 3.4	0.8	−0.37	−2.6, 1.9	0.7	4.7	−3.2, 13	0.2
Positive Distractions	1.8	0.64, 2.9	0.002	−0.28	−9.4, 8.8	>0.9	2.8	−5.2, 11	0.5	−3.5	−18, 11	0.6
Positive Life Experiences	0.26	−1.1, 1.6	0.7	−0.94	−7.7, 5.8	0.8	−0.9	−5.1, 3.3	0.7	5.1	−7.8, 18	0.4
Positive Relationships	−0.85	−1.5, −0.18	0.012	0.36	−3.1, 3.8	0.8	−1.1	−3.4, 1.1	0.3	1.5	−6.1, 9.0	0.7
Professional Care	−0.15	−0.84, 0.53	0.7	0.35	−3.2, 3.9	0.8	−0.66	−3.9, 2.6	0.7	0.33	−12, 13	>0.9
Reflections	0.2	−0.72, 1.1	0.7	−0.85	−7.1, 5.3	0.8	−2.2	−8.1, 3.6	0.5	0.41	−11, 12	>0.9
Romantic Experiences	0.06	−0.73, 0.86	0.9	−0.28	−3.4, 2.9	0.9	−2.2	−5.5, 1.2	0.2	3.4	−5.0, 12	0.4
Seeking Change	−1.2	−1.8, −0.51	<0.001	−0.05	−3.4, 3.3	>0.9	−1.3	−4.5, 1.8	0.4	0.5	−8.7, 9.7	>0.9
Society & Philosophy	−0.01	−0.70, 0.69	>0.9	0.88	−1.9, 3.7	0.5	−1.9	−5.6, 1.8	0.3	2.6	−3.9, 9.1	0.4
Suicide Methods/Attempts	−0.51	−1.3, 0.28	0.2	2.1	−5.2, 9.4	0.6	−0.83	−8.8, 7.1	0.8	2.9	−7.3, 13	0.6
Talking to Friends	0.55	−0.13, 1.2	0.11	−0.09	−2.8, 2.7	>0.9	−2	−5.0, 0.90	0.2	4.8	−3.7, 13	0.3
Talking with Others	2.7	0.82, 4.5	0.005	14	4.6, 24	0.004	3.1	−4.6, 11	0.4	22	9.3, 34	<0.001
Wishing Well	−1.4	−2.0, −0.72	<0.001	−0.15	−3.7, 3.4	>0.9	−0.47	−2.9, 2.0	0.7	0.37	−8.3, 9.1	>0.9

Note: nonsignificant "Authority" values not included for brevity.

Table 5
Significant Interaction Effects and Example Messages (examples are edited to preserve anonymity).

Interaction	Estimate	p-value	Exemplary Message
Care x Negative Emotionality	4.9	<0.001	Oh my gosh! go to the authorities. Screw everybody else, they're not the ones got raped.
Care x Gratitude	4.1	<0.001	Thanks so much i have a notebook of poems i write on hard days to help me out and keep me from suicide
Care x Talking with others	2.65	0.0046	Hi, if you're still in pain and are still here, I can chat with you and help with the feelings you're having
Care x Academic Concerns	2.1	<0.001	Lots of people fail high school and then are successful in community college or 4 year university. Or other people go into an alternative career and still thrive! Don't kill yourself for failing high school. Your life is much bigger than letters on a piece of paper.
Care x Positive Distractions	1.8	0.002	I'm sure we all know how easy it is to fall into unhealthy coping strategies. What are healthy coping strategies you have? How do they help?
Care x Body Image Concerns	1.7	0.004	Harming yourself will only make you feel more ugly. Are there any hygiene routines that help you feel good about your appearance? Why do you feel bad about being gay? And why don't like yourself? I get it tho. Often I feel bad about being into girls. Its ok if you do.
Care x Positive Relationships	−0.85	0.012	Breaking up hurts! It's normal to feel unsteady after the end of a relationship. The worst part is usually right after and then you slowly start to get used to things. Grieve the end of the relationship but don't believe that it has to be your last relationship or tell yourself that you'll never meet anyone again
Care x Seeking Change	−1.2	<0.001	I just wish I could be happy and joyful like other people … I've been trying to fix things and get better for a long time but I just don't have any control. Others control my life and my happiness and they can't change. I just can't be happy there's nothing i can imagine doing to change that
Care x Wishing Well	−1.4	<0.001	Wow that's so shitty. Even though I know I couldn't survive on my own, I wanted to run away from home. Now this conversation makes me feel smaller. I so wish I could help you even a little bit but I can't even help myself. I'm so so sorry man. And I'm really hoping that eventually you end up in a good place. Some of us deserve to be saved, and really I hope it's you cause you really do deserve it
Care x NSSI Experiences/ Urges	−1.4	<0.001	The cuts on my upper thigh, on the inner side is hurting quite a lot. It is placed above a vein but i dont think i cut through that. I remember feeling a striking pain when i first cut around that area. It has been 2 months and they still hurt a lot. I don't think i went too deep but idk …
Fairness x Talking with Others	22.3	<0.001	I don't know if talking to her will help … to make it even more insulting and hurtful, I recently found out she's cheating on me but she doesn't know I know
Fairness x Describing Suicide	14.1	0.0038	Please be honest I really don't want to be lied to–is suicide by hanging really painful and does it take a long time? I've already tied the noose even though I'm still not sure if I'm gonna go thru w it
Loyalty x NSSI Experiences/Urges	19.6	<0.001	I just got betrayed and I picked up the blade as a last resort since nothing else was working. The razor hasn't betrayed me, it's never put me down, and it's never lied to me
Loyalty x Gratitude	9.5	<0.001	Thanks so much and same to you! I read some of your post history so you can talk to me too and we can support each other.)
Purity x Talking with Others	21.6	<0.001	There's no one like me I could even talk to about this because my whole fucking state has zero people brave enough to come out since the state really hates faggots like me. I cant even talk to other queer teens about it online bc my stuff keeps getting removed from lgbteens. Im just so fucking disgusting I cant even look at myself in a mirror
Purity x Describing Suicide	11.3	0.0053	Squeezing the carotid sinus provokes a parasympathetic reaction in your body.

Table 6
Most frequent response topic to post topic.

Post Topic	Response Topic	Percentage
Negative Thoughts	Negative Thoughts	17.5
Academic Concerns	Academic Concerns	18.2
Body Image Concerns	Body Image Concerns	13.1
Describing Suicide	Ideation	8.6
Employment/Financial Experiences	Employment/Financial Experiences	22.8
Family Abuse	Wishing Well	9.7
Gratitude	Gratitude	13.9
Ideation	Seeking Change	11.3
Immediate Help Seeking/ Offering	Wishing Well	8.5
Loved Ones	Wishing Well	9.8
Negative Emotionality	People Care	10.1
NSSI Concealment	NSSI Concealment	22.4
NSSI Experiences/Urges	NSSI Experiences/Urges	33.9
People Care	People Care	18.5
Positive Distractions	Positive Distractions	20
Positive Life Experiences	Positive Life Experiences	14.5
Positive Relationships	Positive Relationships	12.6
Professional Care	Professional Care	21
Reflections	Seeking Change	8.3
Romantic Experiences	Romantic Experiences	10.2
Seeking Change	Seeking Change	16.8
Society & Philosophy	Society & Philosophy	19.1
Suicide Methods/Attempts	Suicide Methods/Attempts	12.5
Talking To Friends	Talking to Friends	14.3
Talking with Others	People Care	8.9
Wishing Well	Wishing Well	12.8

they use e-communities to discuss complex multifaceted experiences and receive support *specific to these experiences*.

Study findings demonstrate that in these SITB e-communities, expressing and receiving care across a variety of conversation topics overwhelmingly dominates the user experience. Individuals in these SITB communities were more likely to discuss care concerns than any other moral concerns (93.7% of all moral messages contained a care concern). Care messages ranged from referrals to formal care to questions about disclosing SITBs and sexuality, to empathic sharing of personal anecdotes (see example message in Table 2). Interestingly, results indicated that, depending on the message topic, the presence of care concerns either significantly increased or decreased the likelihood of that message receiving likes. In conversations in the "Negative Emotionality", "Gratitude", "Talking with Others", "Academic Concerns", "Positive Distractions", and "Body Image Concerns" topics, the presence of a care value increased the likelihood of those messages receiving "likes" (and attention). Conversely, in conversations in the "Positive Relationships", "Seeking Change", "Wishing Well", and "NSSI Experiences and Urges" topics, care concerns significantly decreased "likes" (however it is worth noting that the effect sizes were larger for the positive interactions than the negative). Thus, it appears that while individuals in these communities view and send care messages more than other types of moral-themed messages, expression of care does not indiscriminately resonate with users. Rather, conversations that resonate most with users are those that occur within specific conversation contexts and which center around care.

While care was the dominant moral value in these data, results indicated that for certain conversation topics, individuals were more likely to "like" content if it included moral values other than care purity

Table 7
Model Comparison of the Baseline Model and Models with Predictors. Table Shows That Including Each of the Used Predictors in the Full Model Adds Significantly to explained variance.

Model 1	Model 2	F statistic	P-value	Inference
M0	M1	302.57	<0.001	Topics add significantly to explained variance
M1	M3	391.56	<0.001	Concerns add significantly to explained variance beyond topics
M2	M3	258.32	<0.001	Topics add significantly to explained variance beyond concerns
M3	M4	4.4158	<0.001	Interactions of topics and concerns add significantly to explained variance beyond the individual variables

*M0 = No predictors; M1 = Topics; M2 = Concerns; M3 = Topics + Concerns; M4 = Topics + Concerns + Interaction.

(see Table 5 for examples). In particular, individuals were more likely to "like" content in "Talking with Others" and "Describing Suicide" if content included values of fairness and. Likewise, content in "Gratitude" and "NSSI Experiences/Urges" was more likely to resonate with users if it included loyalty. The effect sizes for all of these interactions were sizable, indicating that individuals engaged in SITBs highly value the inclusion of these moral values in conversations. This is notable because previous research demonstrates that online discussions often lack moral language, even when the content is contentious. For instance, a recent paper analyzing moral language in anti-vaccine online discussions relating to the Covid-19 pandemic found that moral concerns only appeared in 25% of the messages (Trager et al., 2022). Therefore, the current study suggests that content that emphasizes values may have a greater appeal to and resonance with individuals engaged in SITBs than among the general population.

5.1. Clinical implications

Study findings indicate that 1) having specific conversations about SITBs, 2) values-laden content, and 3) receiving and expressing care are core to these SITB e-communities. Considering the sheer volume of users in these communities (>100,000 members in the self-harm e-community and >400,000 members in the suicide e-community), these results provide insight to the real-time needs of those who have lived experience in SITBs at an unprecedented scale, and as such have the potential to inform STIB training and treatment across settings (e.g., private practice, emergency departments, psychiatric wards, hospital outpatient practices, etc.).

Results indicate that specific conversations about SITBs and SITB-adjacent experiences resonate more with individuals engaged in SITBs than general conversations of sadness. Faculty and supervisors developing SITB trainings may consider incorporating this information into their SITB trainings to provide clinicians with understanding of concurrent issues facing individuals engaged in SITBs (e.g., body image concerns, family abuse, employment concerns) and behaviors relevant SITB engagement (e.g., "NSSI Concealment", "NSSI Urges"). Likewise, findings indicate that there are some topics (e.g., "Describing Suicide", "Talking to Others") that resonate more with individuals when they include values of fairness and purity. Thus, facilitators may include specific training material that help clinicians incorporate these values into conversations about suicide and disclosing SITBs with patients.

Study findings extend support for caretaking needs identified in qualitative literature and indicate that those with lived SITB experience may be more likely to engage in treatment (and disclose self-injurious behaviors) when clinicians approach the behavior from a caretaking mentality (rather than a success/failure mentality). Given the historical reticence in the scientific and clinical community to express care towards those engaged in SITBs over fears of SITB reinforcement, these findings regarding care are particularly relevant for SITB treatment development and clinical training. Indeed, those with lived experience in SITBs have long expressed frustration and treatment avoidance when they are punished (either by withdrawal of support or by more coercive measures) for SITBs (Inckle, 2020; MacDonald, 2020; Pembroke, 2002). Instead, many with lived SITB experience emphasize the importance of celebrating "caretaking" even if "caretaking" manifests as cutting on the "stomach instead of the forearm … and celebrating when somebody *has* taken care, not just seeing it as a huge failure that they are self-harming" (Inckle, 2011). This information may be particularly important to incorporate into trainings for emergency department staff, as these health professionals are often the first point of contact in the healthcare system of those engaged in SITBs and past research has consistently demonstrated that those engaged in SITBs experience stigma, rejection, and censure in emergency settings (Bergen et al., 2023; Manchester, 2018; Beale, 2022; MacDonald et al., 2020; emmaontheedge, 2020).

Moreover, study findings suggest that provision of care is not indiscriminately well-received by individuals engaged in SITBs. Findings indicate that conversation topic is important for care expression to be maximally effective: our results suggest that expressions of care are most likely to resonate with those who have lived SITB experience when they occur in the context of emotion dysregulation, concerns about body image and academics, conversations about gratitude, and some, but not all, conversations about social support. For example, when discussing triggers for emotion dysregulation, prior to introducing skills, a clinician may spend more time on validation of distress or more time allowing the patient to share in detail what led to the dysregulation. Likewise, validation may be particularly important in conversations about body image and academic concerns, and this may inform how a therapist chooses which thoughts to challenge when providing cognitive-behavioral therapy to clients engaged in or with a history of SITBs. Thus, these findings can help guide content and affect areas of focus for clinicians and researchers to target as they begin to create person centered, SITB-specific treatments. Additionally, since many clinicians worry that expressing care reinforces SITBs, it is important that clinical training for SITBs address these concerns and provide clinicians with a) data to correct misinformation about SITB engagement (e.g., SITB engagement is manipulative), and b) strategies to provide the care so craved by those engaged in SITBs in an effective, non-reinforcing manner.

Given the volume of users within SITB e-communities, practitioners may also consider how providing formalized SITB e-community training may increase clinician ability to accurately assess social support and information available to patients engaging in SITBs. Knowledge of SITB e-communities may empower clinicians to 1) effectively assess e-community use at intake, 2) consider its functionality in SITB engagement (e.g., does SITB e-community use reinforce SITBs? Does it provide an otherwise isolated patient with support? Does it depend on the context?), and 3) set treatment goals with their patient that include the SITB e-communities already used by the patient (e.g., use e-community to increase social support seeking, explore alternatives to e-community to mitigate isolation).

Finally, as an increasing number of researchers (and funders) turn towards digital interventions for SITBs, findings provide insight for content and structure of digital interventions to maximize retention. People are more likely to use digital interventions when interventions are personalized to their needs (Saleem et al., 2021). Study findings yielded a range of topics ranging from general (painful emotions and sadness) to specific (abuse, employment concerns, etc.) that reflect the unique needs of individuals engaged in SITBs which may be used to help personalize SITB digital interventions. And, given the strong relationship between specific, care-focused messages and likes, when designing in-the-moment or chat-based interventions, researchers may consider how to incorporate more specific and/or more care-laden content or prompts into the intervention to increase use and retention among those engaged in SITBs. For example, a digital intervention study may examine whether using care messaging (e.g., a prompt such as "how did

you take care of your body today?" rather than a prompt like "you've been self-harm free for 17 days, don't break your streak!") increases retention among individuals engaged in SITBs.

5.2. Limitations

Several study limitations must be addressed. First, study data was anonymous and thus sample demographic information is unknown. Although an anonymous dataset limits insight to population specifics and the generalizability of these findings, anonymity is one of the driving forces behind engagement with SITB e-communities, as users often feel more comfortable and open to discussing sensitive mental health information anonymously (DeChoudhury & De, 2014). Future research should consider mixed-method approaches (e.g., qualitative focus groups, market research) to better understand the types of individuals who utilize SITB e-communities. Second, study data came from two monitored SITB e-communities, and thus it is possible that some content may have been removed by moderators, potentially limiting the generalizability of study data. Next, it is important to note that not all individuals who engage in SITBs use SITB e-communities. Therefore, generalizability of these results may be limited to individuals engaged in SITBs who also use online communities. However, because individuals engaged in SITBs are often more likely to seek support online than in person (Wilks et al., 2016), and given the volume of users in these groups, it is likely that these results generalize to a large percentage of the SITB-engaged population. Finally, this study was primarily data-driven, designed to identify key conversation topics and values and understand how the two interact to predict community resonance.

5.3. Conclusion

SITBs constitute one of the leading public health crises in the United States yet effective SITB prevention and intervention remains challenging. To date, SITB interventions have been developed without input from those with lived SITB experience, and are hampered by implicit, disgust driven anti-SITB bias and concerns that providing care may reinforce SITBs. Given the longstanding issues with SITB treatment utilization and efficacy, a growing number of researchers and clinicians have begun calling for the development of SITB-specific interventions that are more congruent with the values and needs of those engaged in SITBs to increase treatment efficacy, utilization, and retention. If we look at the experiences of those who are engaged in SITBs, it is clear that there are care needs that are not being met (Inckle, 2011, 2020; Long, 2017, pp. 89–103; Pembroke, 2002). The current study extends these findings and demonstrates that specific, care-focused content resonates most with individuals engaged in SITBs online. If our field continues to prioritize clinician worries that expressing care reinforces SITBs over patient's needs for care and nurturance, it appears unlikely that treatment utilization and efficacy rates will change.

Credit author statement

E.P: Conceptualization, Writing-Original draft preparation, Methodology, **S.A.:** Methodology, Software, Formal Analysis, Writing-Original draft preparation, **D.W.:** Writing-reviewing and editing, **A. W.:** Supervision, Writing-review and editing.

Declaration of competing interest

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.

Data availability

Data will be made available upon request. Analytic code is available at: : https://github.com/goytoom/eSITB

References

emmaontheedge. (2020). *What if we listened to pain instead of risk? Emma on the edge*. https://emmaontheedgeblog.wordpress.com/2020/09/23/example-post-3/.

Andover, M. S., Schatten, H. T., Morris, B. W., Holman, C. S., & Miller, I. W. (2017). An intervention for nonsuicidal self-injury in young adults: A pilot randomized controlled trial. *Journal of Consulting and Clinical Psychology, 85*(6), 620–631. https://doi.org/10.1037/ccp0000206

Beale, C. (2022). Magical thinking and moral injury: Exclusion culture in psychiatry. *BJPsych Bulletin, 46*(1), 16–19. https://doi.org/10.1192/bjb.2021.86

Bentley, K. H., Maimone, J. S., & Nock, M. K. (2021). In D. H. Barlow (Ed.), *Addressing self-injurious thoughts and behaviors within the context of transdiagnostic treatment for emotional disorders* (6th ed.).

Bergen, C., Lomas, M., Ryan, M., & McCabe, R. (2023). Gatekeeping and factors underlying decisions not to refer to mental health services after self-harm: Triangulating video- recordings of consultations, interviews, medical records and discharge letters. *SSM – Qualitative Research in Health.* , Article 100249. https://doi.org/10.1016/j.ssmqr.2023.100249

Blei, D. M. (2012). Probabilistic topic models. *Communications of the ACM, 55*(4), 77–84. https://doi.org/10.1145/2133806.2133826

Blei, D. M., Ng, A. Y., & Edu, J. B. (2003). Latent dirichlet allocation michael I. *Jordan. Journal of Machine Learning Research, 3*, 993–1022.

Bouma, G. (2009). Normalized (pointwise) mutual information in collocation extraction. *Proceedings of GSCL, 30*, 31–40.

Brown, R. C., Fischer, T., Goldwich, D. A., & Plener, P. L. (2020). I just finally wanted to belong somewhere"—qualitative analysis of experiences with posting pictures of self- injury on Instagram. *Frontiers in Psychiatry, 11*, 274. https://doi.org/10.3389/fpsyt.2020.00274

Brown, R. C., & Plener, P. L. (2017). Non-suicidal self-injury in adolescence. *Current Psychiatry Reports, 19*(3), 20. https://doi.org/10.1007/s11920-017-0767-9

Commons Treloar, A. J., & Lewis, A. J. (2008). Targeted clinical education for staff AttitudesTowards deliberate self-harm in borderline personality disorder: RandomizedControlled trial. *Australian and New Zealand Journal of Psychiatry, 42*(11), 981–988. https://doi.org/10.1080/00048670802415392

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), Article 102034. https://doi.org/10.1016/j.ipm.2019.04.002

Curtis, V., & Biran, A. (2001). Dirt, disgust, and disease: Is hygiene in our genes? *Perspectives in Biology and Medicine, 44*(1), 17–31. https://doi.org/10.1353/pbm.2001.0001

Dean, B. (2021). *Reddit User and growth stats*. https://backlinko.com/reddit-users.

DeChoudhury, M., & De, S. (2014). Mental health discourse on reddit: Self-disclosure, SocialSupport, and anonymity. In *Eighth international AAAI conference on weblogs and SocialMedia*. https://www.aaai.org/ocs/index.php/ICWSM/ICWSM14/paper/view/8075.

Devlin, J., Chang, M.-W., Lee, K., & Toutanova, K. (2019). *Bert: Pre-Training of deep, bidirectional Transformers for language understanding (arXiv:1810.04805)*. arXiv. https://doi.org/10.48550/arXiv.1810.04805

Fisher, J. (2022). *The fortress of mental health services*. Mad in the UK. https://www.madintheuk.com/2022/01/the-fortress-of-mental-health-services/.

Fox, K. R., Bettis, A. H., Burke, T. A., Hart, E. A., & Wang, S. B. (2021). Exploring adolescent, experiences with disclosing self-injurious thoughts and behaviors across settings. *Research on Child and Adolescent Psychopathology*, 1–13. https://doi.org/10.1007/S10802-021-00878-X, 2021.

Fox, K. R., Huang, X., Guzmán, E. M., Funsch, K. M., Cha, C. B., Ribeiro, J. D., & Franklin, J. C. (2020). Interventions for suicide and self-injury: A meta-analysis of randomized controlled trials across nearly 50 years of research. *Psychological Bulletin, 146*(12), 1117–1145. https://doi.org/10.1037/bul0000305

Gearing, R. E., & Lizardi, D. (2009). Religion and suicide. *Journal of Religion and Health, 48*(3), 332–341. https://doi.org/10.1007/s10943-008-9181-2

Graham, J., Haidt, J., Koleva, S., Motyl, M., Iyer, R., Wojcik, S. P., & Ditto, P. H. (2013). Chapter two - moral foundations theory: The pragmatic validity of moral pluralism. In P. Devine, & A. Plant (Eds.), *Advances in experimental social psychology* (Vol. 47, pp. 55–130). Academic Press. https://doi.org/10.1016/B978-0-12-407236-7.00002-4.

Gunderson, J. G., & Choi-Kain, L. W. (2019). Working with patients who self-injure. *JAMA Psychiatry, 76*(9), 976–977. https://doi.org/10.1001/jamapsychiatry.2019.1241

Henderson, N. L., & Dressler, W. W. (2019). Cultural models of substance misuse risk and moral foundations: Cognitive resources underlying stigma attribution. *Journal of Cognition and Culture, 19*(1–2), 78–96. https://doi.org/10.1163/15685373-12340049

Hom, M. A., Stanley, I. H., & Joiner, T. E. (2015). Evaluating factors and interventions that influence help-seeking and mental health service utilization among suicidal individuals: A review of the literature. *Clinical Psychology Review, 40*, 28–39. https://doi.org/10.1016/j.cpr.2015.05.006

Hooley, J. M., & Franklin, J. C. (2018). Why do people hurt themselves? A new conceptual, model of nonsuicidal self-injury. *Clinical Psychological Science, 6*(3), 428–451. https://doi.org/10.1177/2167702617745641

Hoover, J., Portillo-Wightman, G., Yeh, L., Havaldar, S., Davani, A. M., Lin, Y., Kennedy, B., Atari, M., Kamel, Z., Mendlen, M., Moreno, G., Park, C., Chang, T. E., Chin, J., Leong, C., Leung, J. Y., Mirinjian, A., & Dehghani, M. (2020). Moral foundations twitter, corpus: A collection of 35k tweets annotated for moral sentiment. *Social Psychological and Personality Science, 11*(8), 1057–1071. https://doi.org/10.1177/1948550619876629

Inckle, K. (2011). The first cut is the deepest: A harm-reduction approach to self-injury. *Social Work in Mental Health, 9*(5), 364–378. https://doi.org/10.1080/15332985.2011.575726

Inckle, K. (2020). Inequality, distress and harm-reduction: A social justice approach to self-injury. *Social Theory &, ealth, 18*(4), 224–239. https://doi.org/10.1057/s41285-020-00146-w

Kang, L. L., Rowatt, W. C., & Fergus, T. A. (2016). Moral foundations and obsessive-compulsive symptoms: A preliminary examination. *Journal of Obsessive-Compulsive and Related Disorders, 11*, 22–30.

Lavis, A., & Winter, R. (2020). #Online harms or benefits? An ethnographic analysis of the positives and negatives of peer support around self-harm on social media. *Journal of Child Psychology and Psychiatry, 61*(8), 842–854. https://doi.org/10.1111/jcpp.13245

Long, M. (2017). *'We're not monsters … we're just really sad sometimes:' hidden self-injury, stigma and help-seeking*. https://doi.org/10.1080/14461242.2017.1375862,27(1), 10.1080/14461242.2017.1375862.

Loper, E., & Bird, S. (2002). *Nltk: The Natural Language toolkit (arXiv:cs/0205028)*. arXiv. https://doi.org/10.48550/arXiv.cs/0205028

MacDonald, S., Sampson, C., Turley, R., Biddle, L., Ring, N., Begley, R., & Evans, R. (2020). Patients' experiences of emergency hospital, care following self-harm. *SystematicReview and Thematic Synthesis of Qualitative Research, 30*(3), 471–485. https://doi.org/10.1177/1049732319886566

Manchester: The University of Manchester. (2018). *The assessment of clinical risk in mental health services*. National Confidential Inquiry into Suicide and Safety in Mental Health, (NCISH).

Meheli, S., & Banerjee, D. (2022). Revisiting social stigma in non-suicidal self-injury: A, narrative review. *Consortium Psychiatricum, 3*(3). https://doi.org/10.17816/CP196. Article 3.

Mikolov, T., Chen, K., Corrado, G., & Dean, J. (2013). *Efficient estimation of word, representations in vector space (arXiv:1301.3781)*. arXiv. https://doi.org/10.48550/arXiv.1301.3781

Miron, O., Yu, K.-H., Wilf-Miron, R., & Kohane, I. S. (2019). Suicide rates among adolescents and young adults in the United States, 2000-2017. *JAMA, 321*(23), 2362–2364. https://doi.org/10.1001/jama.2019.5054

Mitten, N., Preyde, M., Lewis, S., Vanderkooy, J., & Heintzman, J. (2016). The perceptions of adolescents who self-harm on stigma and care following inpatient psychiatric treatment. *Social Work in Mental Health, 14*(1), 1–21. https://doi.org/10.1080/15332985.2015.1080783

National Institute for Health Care Excellence. (2013). *Self-harm. Quality standard*. Retrieved from: https://www.nice.org.uk/guidance/qs3.

Newman, D., Bonilla, E. V., & Buntine, W. (2011). Improving topic coherence withRegularized topic models. *Advances in Neural Information Processing Systems, 24*. https://proceedings.neurips.cc/paper/2011/hash/5ef698cd9fe650923ea331c15af3b160-Ab stract.html.

Nock, M. K., Joiner, T. E., Gordon, K. H., Lloyd-Richardson, E., & Prinstein, M. J. (2006). Non-suicidal self-injury among adolescents: Diagnostic correlates and relation to suicide attempts. *Psychiatry Research, 144*(1), 65–72. https://doi.org/10.1016/j.psychres.2006.05.010

Pembroke, L. (2002). *Cutting the risk: Self-harm, self care and risk reduction*. The National SelfHarm Network. http://studymore.org.uk/ctr.pdf.

Pola, S., & Sheela Rani Chetty, M. (2021). Behavioral therapy using conversational chatbot for depression treatment using advanced RNN and pretrained word embeddings. *Materials Today: Proceedings*. https://doi.org/10.1016/j.matpr.2021.02.521

Preston, E. G., & West, A. E. (2023). What do my (online) friends think? *A Topic Modeling Approach to Identifying Patterns of Response to Self-Injurious Behaviors on Reddit. Archives of Suicide Research*. https://doi.org/10.1080/13811118.2023.2193594

Pury, C. L. S., & Mineka, S. (1997). Covariation bias for blood-injury stimuli and aversive outcomes. *Behaviour Research and Therapy, 35*(1), 35–47. https://doi.org/10.1016/S0005-7967(96)00075-7

R Development Core Team. (2010). *A language and environment for statistical computing: Reference index*. R Foundation for Statistical Computing. http://www.polsci.wvu.edu/duval/PS603/Notes/R/fullrefman.pdf.

Rehurek, R., & Sojka, P. (2012). *Gensim–python framework for vector space modelling*. https://radimrehurek.com/gensim/.

Robinson, L. (2020). *A qualitative study into people's experiences of interventions and support for non-suicidal self-injury (NSSI): Stigma, shame, and society*. Doctoral Thesis, University of the West of England. https://uwerepository.worktribe.com/output/1491411.

Rowe, S. L., French, R. S., Henderson, C., Ougrin, D., Slade, M., & Moran, P. (2014). Help-seeking behaviour and adolescent self-harm: A systematic review. *Australian and New Zealand Journal of Psychiatry, 48*(12), 1083–1095. https://doi.org/10.1177/0004867414555718

Saleem, M., Kühne, L., Santis, K. K. D., Christianson, L., Brand, T., & Busse, H. (2021). Understanding engagement strategies in digital interventions for mental health, promotion: Scoping review. *JMIR Mental Health, 8*(12), Article e30000. https://doi.org/10.2196/30000

Sarabian, C., Curtis, V., & McMullan, R. (2018). Evolution of pathogen and parasite avoidance behaviours. *Philosophical Transactions of the Royal Society B: Biological Sciences, 373*(1751), Article 20170256. https://doi.org/10.1098/rstb.2017.0256

Schoppmann, S., Schröck, R., Schnepp, W., & Büscher, A. (2007). Then I just showed her my arms. . ." Bodily sensations in moments of alienation related to self-injurious behaviour, A hermeneutic phenomenological study. *Journal of Psychiatric And Mental Health Nursing, 14*(6), 587–597. https://doi.org/10.1111/j.1365-2850.2007.01150.x

Staniland, L., Hasking, P., Boyes, M., & Lewis, S. (2021). Stigma and nonsuicidal self-injury: Application of a conceptual framework. *Stigma and Health, 6*(3), 312–323. https://doi.org/10.1037/SAH0000257

Trager, J., Ziabari, A. S., Davani, A. M., Golazazian, P., Karimi-Malekabadi, F., Omrani, A., & Dehghani, M. (2022). *The moral foundations reddit corpus*. https://doi.org/10.48550/arXiv.2208.05545. *arXiv preprint arXiv:2208.05545*.

Turc, I., Chang, M.-W., Lee, K., & Toutanova, K. (2019). *Well-read students learn better: On the importance of pre-training compact models (arXiv:1908.08962)*. arXiv. https://doi.org/10.48550/arXiv.1908.08962

Wackwitz, v. R. (1992). https://law.justia.com/cases/virginia/supreme-court/1992/911384-1.html.