



Religious Language in Suicide-Related Reddit Posts: Emotional Tone (Negativity vs. Positivity) in a Large-Scale Secondary Text Analysis

Article History:

Initial submission:	26 February 2026
First decision:	28 February 2026
Revision received:	-
Accepted for publication:	20 March 2026
Online release:	08 April 2026

Melvin A. Macuha, RPsy, ORCID No. 0000-0003-2748-3586

Psychology Program Chairperson, St. Paul University Manila, 680 Pedro Gil Street, Malate, Manila, Philippines

Abstract

Religious and spiritual concerns often occur during suicidal crises; however, little is known whether religious language in suicide-related online discourse tends to be emotionally negative (e.g., despair, guilt, abandonment) or positive (e.g., hope, comfort, meaning). Using a large and publicly available body of Reddit posts drawn from suicide-related communities, this study examined the emotional tone associated with religious language in suicide-related posts. The researcher conducted a secondary analysis of the Suicide and Depression Detection dataset (Kaggle), restricting analyses to posts labeled "suicide" (N= 116,037). Religious language was operationalized using a transparent keyword-based lexicon. Emotional tone was estimated through VADER sentiment scoring. Prevalence estimates were reported and two logistic regression models were tested: (a) whether religious language predicted high negativity and (b) whether religious language predicted high positivity, controlling for post length. Religious language appeared in 10.91% of suicide-labeled posts and was associated with heightened emotional polarity and a significantly more negative overall sentiment. Controlling for post length, religious language predicted greater odds of high negativity (OR = 1.50, 95% CI [1.44, 1.57], $p < .001$) and lower odds of high positivity (OR = 0.79, 95% CI [0.75, 0.83], $p < .001$). Findings suggest that, within suicide-related Reddit discourse, religious language is more frequently embedded in negative emotional framing than in positive emotional expression. Ethical safeguards for sensitive public data research are discussed, along with implications for spiritually integrated crisis communication and future cross-cultural studies.

Keywords: Digital mental health, Reddit, Religion, Sentiment analysis, Spirituality, Suicide, VADER



Copyright © 2026. The Author/s. Published by VMC Analytik's Multidisciplinary Journal News Publishing Services. Religious Language in Suicide-Related Reddit Posts: Emotional Tone (Negativity vs. Positivity) in a Large-Scale Secondary Text Analysis © 2026 by Melvin A. Macuha is an open access article licensed under [Creative Commons Attribution \(CC BY 4.0\)](https://creativecommons.org/licenses/by/4.0/). This permits the copying, redistribution, remixing, transforming, and building upon the material in any medium or format for any purpose, even commercially, provided that appropriate credit is given to the copyright owner/s through proper acknowledgement and standard citation.

INTRODUCTION

Suicide remains a major public health concern worldwide, and many individuals turn to online spaces to disclose distress, seek validation, and request support during suicidal crises (Bjärehed, 2023; Sørensen, 2023; Yeskuatov, 2022). Research has shown that suicide-related communities and related mental health forums contain rich linguistic signals that can be studied to understand emotional tone, thematic content, and help-seeking behavior (Yan et al., 2024; Zhu et al., 2024). Alongside psychological distress, religious and spiritual themes often arise in crisis narratives and meaning-making efforts, including prayer, perceived divine abandonment, spiritual guilt, or hope for transcendence.

Trends in suicidology looks into religion, where it is frequently discussed as a protective factor against suicidal behaviors (Poorolajal et al., 2022). However, clinical and psychosocial research also emphasizes the presence of religious and spiritual struggles and negative religious coping, which may aggravate suicidality for some individuals (Lemke et al., 2023; Pulgar et al., 2022). These seemingly conflicting patterns highlight a central issue that religion can be experienced as comfort or conflict; and, this duality may appear in the language used during suicidal crises, including online discourses.

While a growing body of work uses natural language processing (NLP) to characterize suicide-related online discourse (Dan et al., 2025; Lundeen et al., 2024), relatively fewer

studies directly ask a focused descriptive question, “when people invoke religion during suicide-related talk online, does it tend to be emotionally negative or emotionally positive?” Addressing this question can help clarify whether religious language in crisis discourse is more consistent with spiritual struggle (e.g., guilt, abandonment, hell, ultimate meaning conflict) over spiritual consolation (e.g., hope, prayerful coping).

Thus, this study examines the emotional tone associated with religious language in a large-scale body of suicide-labeled Reddit posts. Rather than making causal claims about religion and suicidality, the researcher focused on emotional framing within suicide-related discourse.

LITERATURE REVIEW

Suicide-related Online Discourse and Sentiment/ Linguistic Analysis. Recent work has increasingly investigated suicide-related texts and user behaviors in online communities. Studies have examined linguistic features and risk signals among users posting in r/SuicideWatch (Yan et al., 2024) and explored emotional tones and topics among overlapping users between depression-related and suicide-related communities (Zhu et al., 2024). Research has also emphasized the complexity of online crisis contexts and the need to interpret sentiment and themes cautiously, especially where vulnerability concerns may exist (Lundeen et al., 2024). Adding to these observational approaches, a qualitative work has developed frameworks to explain risk and protective factors in suicide crisis forums (Chancellor et al., 2021).

Beyond predictive modeling, methodological reviews have synthesized how Reddit-based datasets are leveraged for suicide ideation detection and sentiment profiling. Yeskuatov (2022) critically examined approaches to suicide detection using Reddit posts, highlighting the widespread use of lexicon-based and machine learning sentiment classifiers to quantify emotional expression.

The review emphasized both the strengths of large-scale linguistic modeling and the ethical complexities inherent in interpreting naturally occurring crisis narratives. These computational approaches demonstrate that suicide-related posts often contain distinct affective signatures, including heightened negative sentiment and lexical concentration around hopelessness, self-devaluation, and perceived burdensomeness.

Research has also examined the interactive dimensions of suicide-related discourse, moving beyond isolated posts to analyze communicative patterns within peer-to-peer digital spaces. Sørensen et al. (2023) investigated how individuals experiencing suicide ideation communicate within mediated online groups, finding that posts frequently combine disclosure of distress with explicit requests for validation, reassurance, or advice. Their findings underscore that the expressions of distress are often embedded within relational exchanges shaped by community norms. Complementing this perspective, Bjärehed et al. (2023) conducted thematic analyses of responses to help-seeking posts related to depression and suicide ideation, identifying linguistic patterns of empathy, normalization, and experiential sharing. These studies illustrate that online suicide discourse is characterized by both high emotional intensity and socially constructed meaning-making processes.

As such, the literature demonstrates that suicide-related online discourse is linguistically distinct, emotionally intense, and methodologically accessible through sentiment and computational analysis. However, while prior research has extensively examined general affective tone and risk-related linguistic markers, fewer studies have isolated specific semantic domains, such as religious or spiritual language, and examined how they are emotionally framed within suicide discourse.

Religion/Spirituality, Suicide Ideation, and Spiritual Struggle. A meta-analysis found that religious affiliation and participation are

generally associated with reduced risk across multiple suicidal outcomes, which support a protective association at the population level (Poorolajal et al., 2022). However, emerging literature underscores that religion and spirituality may also involve conflict and distress, particularly through religious/spiritual (R/S) struggles involving ultimate meaning, divine relations, moral concerns, or interpersonal religious tensions. Bidirectional longitudinal links between suicide ideation and ultimate-meaning R/S struggles among psychiatric outpatients were reported that highlighted clinical relevance of spiritual struggle in risk assessment (Lemke et al., 2023). Studies have also examined negative religious coping as associated with suicidality indicators (Pulgar et al., 2022) and explored religious commitment, faith struggles, and related constructs in the context of suicide intent (Mason et al., 2023).

More recent qualitative study has explored spirituality-related concerns in suicidal contexts and emphasizes that spiritual needs may be under-identified in standard assessments like spiritual needs among individuals with suicide ideation (Ghahramani et al., 2024). Similarly, qualitative work has explored spiritual fading and spiritual dimensions in suicide-related experiences (Ghahramani et al., 2025). As such, contemporary literature supports a nuanced view, where religious/spiritual themes may function as comfort or conflict depending on individual and contextual factors.

Sentiment analysis tools for social media text. In social media sentiment research, VADER is widely used due to its design for short, informal text and its interpretable scoring (Jain et al., 2023). More recent evaluations compare automated sentiment tools and note performance differences across domains that support careful methodological reporting when sentiment is central to inference (Gandy et al., 2025). These considerations are important in suicide-related discourse, where sarcasm, ambivalence, and crisis language can complicate sentiment measurement.

Despite growth in computational analyses of suicide-related forums, fewer studies center a focused descriptive and inferential question about the emotional valence of religious language within suicide-related talk. This study addresses that gap using a transparent, reproducible approach that combines lexicon-based religious language detection with sentiment scoring and regression modeling.

Theoretical Framework. This study is primarily grounded in Terror Management Theory (TMT) (Greenberg et al., 1997) and informed by current developments in the psychology of religion and suicide research. Terror Management Theory argues that awareness of mortality produces existential anxiety, which individuals regulate through adherence to cultural worldviews, including religious belief systems that provide meaning, coherence, and symbolic or literal immortality. Religion, in this framework, functions as an anxiety-buffering system that helps individuals manage death-related cognitions.

In the context of suicide-related discourse, the mortality relevance is not merely abstract but immediate and personally relevant. Individuals posting in suicide-related forums are often confronting their own potential death. According to TMT, religious language may emerge in such contexts as a coping attempt to manage existential threat, reaffirm meaning, or negotiate beliefs about afterlife, divine judgment, forgiveness, or suffering.

However, TMT also allows for ambivalence. When religious worldviews are experienced as punitive, rejecting, or invalidating, they may fail to buffer anxiety and may even aggravate distress. Thus, religious discourse during suicide ideation may manifest as either meaning-making and hope-oriented language, or guilt, condemnation, hopelessness, or spiritual struggle.

Furthermore, the study is also informed by Meaning-Making Theory (Park, 2010), which proposes that individuals facing distress attempt to reconcile situational meaning (e.g.,

suffering, suicide ideation) with global meaning systems (e.g., religious beliefs). Discrepancies between the two may result in spiritual struggle, while successful integration may promote adaptive coping.

With this, Terror Management Theory explains why religion may be used during suicide-related discourse (mortality salience and existential threat), while Meaning-Making Theory explains how religious language may take on either positive or negative emotional tone depending on whether reconciliation or struggle occurs.

Thus, this study conceptualizes religious language in suicide-related posts as a psychologically meaningful response to existential distress, whose emotional tone may reflect either buffering or aggravating processes.

Conceptual Framework. This study conceptualizes suicide-related online posts as linguistic manifestations of acute psychological distress. Within these posts, the presence of religious language (e.g., references to God, prayer, heaven, hell, church, Allah) is treated as an indicator of engagement with a religious meaning system.

The conceptual model proposes that the presence of religious language predicts variation in the emotional tone of suicide-related discourse. Emotional tone is operationalized through sentiment polarity measures, including negativity and positivity indices derived from computational text analysis.

Specifically, the framework theorizes that religious language (independent variable) is associated with increased likelihood of high negativity in posts, that religious language is associated with decreased likelihood of high positivity, and that post length is included as a control variable given that longer posts may naturally contain greater emotional elaboration.

The model does not assume that religion is inherently protective or harmful. Instead, it treats religious language as a contextual meaning-making variable whose emotional expression may vary depending on the individual's psychological state.

With this, the conceptual framework positions religious language as a predictor of emotional tone in suicide-related discourse, controlling for structural characteristics of the length of the post.

Current Study. Although religion is frequently discussed as either a protective or risk-related factor in suicide research, less is known about how religion is emotionally expressed within suicide-related discourse in online environments. In particular, while prior studies have examined religious coping, spiritual struggle, and suicidality in clinical and survey-based samples, fewer investigations have quantitatively analyzed how religious language is framed within large-scale suicide-related social media bodies. Guided by Terror Management Theory and Meaning-Making Theory, the present study treats religious language as an indicator of engagement with existential belief systems during moments of acute distress. Rather than assuming religion to be uniformly protective or harmful, this study examines whether religious language in suicide-labeled Reddit posts is associated with more negative or more positive emotional expression. By combining lexicon-based identification of religious language with sentiment analysis and regression modeling, the study aims to clarify the emotional tone surrounding religious references in suicide-related online discourse. More specifically, the present study addresses the following research questions:

1. What proportion of suicide-labeled Reddit posts contain religious language?
2. Do suicide-labeled posts with religious language differ in average sentiment from those without religious language?

3. Controlling for post length, does religious language predict higher odds of high negativity?
4. Controlling for post length, does religious language predict higher odds of high positivity?

METHODOLOGY

Design. The present study employed a quantitative and secondary data analysis design using computational text analytics through a reproducible computational workflow in Google Colaboratory (Colab). The research questions of the study are descriptive and inferential in nature and focus on naturally occurring linguistic behavior in suicide-related discourse.

With this, a large-scale body of suicide-labeled Reddit posts provides ecologically valid data capturing real-world expressions of suicide ideation and existential concern.

Moreover, the use of secondary data is ethically and practically justified. Suicide-related disclosures are sensitive, and directly recruiting individuals experiencing suicide ideation for research on religious language may pose ethical and safety concerns. Analyzing de-identified, publicly available posts minimizes participant burden and reduces potential harm while still enabling rigorous empirical investigation.

Thus, the combination of large-scale analysis, sentiment scoring, and regression modeling provides an appropriate and methodologically defensible framework for answering the research questions.

Dataset and sampling. The researcher used the Suicide and Depression Detection dataset hosted on Kaggle, described as posts collected from the “SuicideWatch” and “depression” subreddits via the Pushshift API (Kaggle dataset page). Analyses were restricted to posts labeled as “suicide” within the dataset, resulting in N= 116,037 suicide-labeled posts, as produced in the Colab workflow.

Furthermore, the researcher utilized a general dataset, not Philippines-only, for it would require reliable geolocation inference, which is not directly available for Reddit posts. Given these constraints, a large general body of data to answer a foundational descriptive question about emotional framing of religious language in suicide-related talk was used. The researcher presents cross-cultural generalizability as a limitation and recommend future work that explicitly samples Philippine contexts, like PH subreddit ecosystems or locally collected body of data, subject to platform rules and ethics.

Variables. The primary independent variable in this study is religious language, operationalized as a binary indicator reflecting the presence or absence of religion-related terms within each suicide-labeled Reddit post. Religious language was identified using a predefined lexicon of keywords representing institutional religion (e.g., church, mosque), theological references (e.g., God, Allah, Jesus), spiritual practices (e.g., pray, rosary), and eschatological concepts (e.g., heaven, hell). Posts containing at least one matching term were coded as 1 (religious language present), and those without were coded as 0 (religious language absent). This operationalization treats religious language as a linguistic marker of engagement with religious or spiritual meaning systems during suicidal discourse.

The primary dependent variables reflect emotional tone, operationalized through sentiment analysis. Sentiment scores were derived using the VADER SentimentIntensityAnalyzer, which produces three polarity components (negative, neutral, positive) and a composite compound score. In the present study, mean negative and positive scores were examined descriptively, while the compound score served as the basis for constructing high-negativity and high-positivity binary outcome variables. These operational definitions allow examination of whether religious language predicts increased probability of emotionally intense negative or positive expression.

A control variable, length of post, was included in the regression analyses. Post length was computed as the total number of characters in each post. This variable was included to account for wordiness and narrative elaboration, as longer posts may have greater opportunity to include affective language. Controlling for post length enables more precise estimation of the association between religious language and emotional tone independent of structural text differences.

Computational Text Analysis Procedure. All analyses were conducted using Python in Google Colaboratory (Colab), a cloud-based Python computational environment suitable for large-scale text processing. The use of Colab ensured that the entire analytic flow, from preprocessing to regression modeling, was documented, executable, and replicable. The following subsections describe the step-by-step procedure used to answer the research questions.

The dataset was imported into a pandas DataFrame, a structured tabular data object that enables efficient manipulation, filtering, and transformation of large textual data. The DataFrame structure allowed systematic filtering of suicide-labeled posts and facilitated the creation of derived variables used in subsequent analyses.

Posts labeled as “suicide” in the dataset were subgrouped to create the analytic sample. Basic preprocessing procedures were applied to ensure textual consistency, including conversion of text to string format and handling of missing values. Post length was computed as a structural control variable using character or token counts, as longer posts may naturally contain more sentiment-bearing words and therefore inflate polarity scores.

This structured data management approach using pandas is widely used in computational social science research due to its scalability, reproducibility, and compatibility with statistical modeling libraries.

Identification of Religious Language. Religious language was operationalized using a keyword-based lexical detection procedure. A case-insensitive regular expression (regex) pattern was constructed to identify posts containing religious or spiritual terms (e.g., “God,” “faith,” “pray,” “church,” “heaven,” “sin”). Regex pattern matching is a transparent and replicable method for detecting predefined lexical categories within large text data.

Each post was evaluated against the regex pattern, and a binary variable (“religion_flag”) was created to indicate the presence (1) or absence (0) of religious language. Lexicon-based detection was selected for its interpretability and methodological transparency, allowing clear operational definition of religious references. While more complex machine learning classifiers could be employed, the keyword-based approach ensures conceptual clarity and reproducibility, which is particularly important in exploratory computational research.

Sentiment Analysis. Emotional tone was quantified using the Valence Aware Dictionary and Sentiment Reasoner (VADER), a rule-based sentiment analysis tool specifically optimized for social media text. VADER produces four polarity components for each text: negative, neutral, positive, and a composite compound score representing overall emotional valence.

VADER was selected because it is widely validated for short, informal, and social media-based text, making it appropriate for Reddit posts. Its lexicon and rule-based adjustments account for capitalization, punctuation emphasis, degree modifiers, and negation handling, which enhances its sensitivity to affective expression in online discourse.

For each post, negative, positive, neutral, and compound polarity scores were extracted. Descriptive comparisons focused on negative and positive polarity components, while the compound score served as an overall indicator of emotional orientation.

Construction of High-Negativity and High-Positivity Indicators. To examine extreme emotional expression, binary outcome variables were constructed based on the distribution of compound sentiment scores. Posts falling within the upper quartile of negative compound values were classified as “high negativity,” whereas posts in the upper quartile of positive compound values were classified as “high positivity.” This threshold-based classification enabled logistic regression modeling of emotionally intense posts, rather than relying solely on continuous sentiment comparisons.

This approach is methodologically appropriate when examining whether specific linguistic features (e.g., religious language) increase the likelihood of emotionally extreme expression.

Statistical Analyses. Descriptive statistics were first computed to determine the proportion of suicide-labeled posts containing religious language. Group comparisons of mean sentiment polarity (negative, positive, compound) were then conducted to assess differences in emotional tone between posts with and without religious language.

To address inferential research questions, logistic regression analyses were performed using the statsmodels library in Python. Logistic regression was selected because the outcome variables (high negativity and high positivity) were binary. The primary predictor was religious language presence, and post length was included as a control variable to account for structural differences in text size. Odds ratios (OR) and 95% confidence intervals were computed to facilitate interpretability of effect sizes. Statistical significance was evaluated at conventional alpha levels.

Appropriateness of the Computational Design. The use of computational text analysis methods was particularly appropriate for the present research questions, which required examination of emotional tone across more than 100,000 suicide-related posts. Traditional qualitative coding would not have been feasible

at this scale. The combination of lexicon-based religious detection, validated sentiment analysis, and regression modeling provides a reproducible and scalable framework for examining linguistic correlates of distress in large digital datasets.

By integrating transparent lexical identification procedures with validated sentiment scoring and inferential statistical modeling, the analytic workflow ensures both interpretability and methodological rigor.

Ethical considerations. This study involved secondary analysis of publicly available and de-identified data obtained from a dataset repository containing Reddit posts collected from suicide-related communities. No direct interaction with users occurred, and no attempts were made to identify, trace, or contact individuals whose posts were included in the dataset. All analyses were conducted at the aggregate level, and usernames, timestamps, or metadata that could reasonably facilitate re-identification were neither analyzed nor reported. Because the dataset consisted of publicly accessible text and contained no direct identifiers, the research was conducted in accordance with ethical standards for secondary analysis of publicly available online data.

Given the highly sensitive nature of suicide-related discourse, particular care was taken to minimize potential harm. The study did not reproduce verbatim excerpts that could be searchable or linked back to specific users. Where illustrative language may be discussed in future dissemination, paraphrasing will be used to prevent traceability. The analytic procedures focused on linguistic patterns rather than individual risk classification, thereby avoiding stigmatization or labeling of specific users.

Data were stored within secured cloud-based computational environments with restricted access, and no redistribution of raw data occurred beyond the original repository source. The research design avoided any intervention,

prediction targeting, or automated flagging of individuals, limiting the study strictly to descriptive and inferential analysis of aggregated textual features.

In addition, the study adhered to emerging best practices in digital mental health research, recognizing that online disclosures, although public, may be perceived by users as occurring within semi-private support contexts. Accordingly, the study was guided by principles of respect for persons, beneficence, and minimization of harm. The findings are presented in a manner that prioritizes aggregate trends over individual narratives and avoids sensationalization of suicide-related content.

Finally, computational assistance tools were used to support analytic Python coding, specifically ChatGPT (OpenAI, 2026); all analytic decisions, model specifications, and interpretations were independently reviewed and verified by the researcher. Responsibility for the integrity and ethical conduct of the study rests entirely with the researcher.

RESULTS

RQ1. What proportion of suicide-labeled Reddit posts contain religious language? Out of the 116,037 suicide-labeled Reddit posts included in the analytic sample, 12,657 posts (10.91%) contained at least one religious or spiritual keyword. This indicates that approximately one in ten suicide-related posts explicitly referenced religion or spirituality. The prevalence estimate suggests that religious meaning systems are present, though not dominant, within suicide-related online discourse.

Table 1
Prevalence of Religious Language in Suicide-Labeled Reddit Posts (N = 116,037)

Religious Language	n	%
Present	12,657	10.91
Absent	103,380	89.09

RQ2. Do suicide-labeled posts with religious language differ in average sentiment from those without religious language? Descriptive comparisons of sentiment polarity components revealed systematic differences between posts containing religious language and those without such references. Posts containing religious language exhibited higher mean negative polarity (M= 0.179) compared to posts without religious language (M= 0.168). At the same time, religious posts also demonstrated slightly higher positive polarity (M= 0.118) relative to non-religious posts (M= 0.110). Despite this modest elevation in positive polarity, the overall compound sentiment score for religious posts was substantially more negative (M= -0.551) than that of non-religious posts (M= -0.406).

Table 2
Mean Sentiment Polarity Scores by Presence of Religious Language

Religious Language	Neg (M)	Pos (M)	Compound (M)
Absent	0.168	0.110	-0.406
Present	0.179	0.118	-0.551

Note. Sentiment scores derived from VADER polarity outputs.

These findings are consistent with broader computational analyses showing elevated negative affect in suicide-related communities (Yan et al., 2024; Zhu et al., 2024) but extend prior work by isolating religious language as a specific linguistic feature associated with greater negativity.

RQ3. Controlling for post length, does religious language predict higher odds of high negativity? A logistic regression model was conducted to examine whether religious language predicted high negativity while controlling for post length. The model was statistically significant, and religious language emerged as a significant predictor ($\beta = 0.407, p < .001$). The odds ratio indicated that posts containing religious references were 1.50 times more likely to fall within the high-negativity category compared to posts without religious language (OR= 1.50, 95% CI [1.44, 1.57]). Post length was also statistically

significant but demonstrated a comparatively small effect size.

Table 3
Logistic Regression Predicting High Negativity (Controlling for Post Length)

Predictor	β	SE	OR	95% CI for OR	p
Religious Language	0.407	—	1.50	[1.44, 1.57]	< .001
Post Length	0.0005	—	1.0005	—	< .001

Note: OR = Odds Ratio. High negativity defined based on sentiment distribution threshold.

These findings suggest that the association between religious language and negative emotional tone is not attributable solely to wordiness. Rather, religious language independently predicts greater emotional negativity in suicide-related discourse. This pattern aligns with research linking religious/spiritual struggles and negative religious coping to heightened suicide ideation and distress (Lemke et al., 2023; Pulgar et al., 2022), though it does not contradict broader meta-analytic findings of religiosity as protective at population levels (Poorolajal et al., 2022).

RQ4. Controlling for post length, does religious language predict higher odds of high positivity?

A second logistic regression model examined whether religious language predicted high positivity (upper quartile of sentiment scores), controlling for post length. Religious language was significantly associated with lower odds of high positivity ($\beta = -0.237, p < .001$). The odds ratio (OR = 0.79, 95% CI [0.75, 0.83]) indicates that posts containing religious references were approximately 21% less likely to fall within the high-positivity category compared to posts without religious language. Post length again showed statistical significance with minimal substantive impact.

This complementary model reinforces the pattern observed in SOP 3. Religious language in suicide-related posts appears more strongly associated with negative emotional expression than with positive emotional tone.

Table 4
Logistic Regression Predicting High Positivity (Controlling for Post Length)

Predictor	β	SE	OR	95% CI for OR	p
Religious Language	-0.237	—	0.79	[0.75, 0.83]	< .001
Post Length	0.0005	—	1.0005	—	< .001

DISCUSSION

The present study examined the emotional framing of religious language within suicide-related Reddit posts, using large-scale sentiment analysis and regression modeling. Religious language was present in approximately 10.91% of suicide-labeled posts, indicating that while not abundant, religious and spiritual constructs are meaningfully invoked during expressions of suicidal distress. Descriptive analyses revealed that posts containing religious references exhibited higher negative polarity and slightly elevated positive polarity, but a substantially more negative overall compound sentiment compared to posts without religious language. Logistic regression models controlling for post length demonstrated that religious language significantly predicted greater odds of high negativity and lower odds of high positivity. These patterns remained strong even when adjusting for wordiness, suggesting that the association between religious language and emotional tone cannot be attributed solely to text length or narrative elaboration.

These findings contribute to the growing computational literature on suicide-related discourse in online communities. Prior research has highlighted the prevalence of negative affect and distress signals in forums such as r/SuicideWatch (Yan et al., 2024) and in overlapping depression-suicide subreddits (Zhu et al., 2024), but few studies have isolated how specific semantic categories, such as religion, relate to emotional expression. The present results extend this work by demonstrating that religious language is associated with intensified affective polarity, particularly heightened negativity. At the same

time, the moderate elevation in positive polarity suggests emotional ambivalence rather than a unidirectional negative effect, reflecting the complexity of affective experience in crisis contexts.

From a theoretical perspective, the findings align with Expectations from Terror Management Theory (Greenberg et al., 1997), which argues that awareness of mortality activates culturally grounded meaning systems, including religion, that individuals use to interpret existential threat. In the context of suicide ideation, individuals invoking religious language may be engaging with spiritual frameworks in an attempt to make sense of suffering. However, such engagement does not uniformly confer emotional comfort; instead, it may be entangled with perceptions of moral failure, divine judgment, or existential conflict. This interpretation is consistent with research on negative religious coping and spiritual struggle, which has been associated with increased psychological distress and suicidal indicators (Lemke et al., 2023; Pulgar et al., 2022). Thus, religious language in suicide-related discourse may reflect unresolved meaning-making efforts, internal conflict, or attempts to negotiate suffering in ways that are emotionally intense and predominantly negative.

It is important to contextualize these findings alongside broader suicide research that often reports protective associations of religiosity at the population level (Poorolajal et al., 2022). The present study does not contradict population-level findings; rather, it highlights that within active expressions of suicidal distress, observable in online discourse, religious language may function as a linguistic marker of emotional struggle rather than as a source of consolation. This distinction underscores the value of examining how religious language is framed and experienced during crisis rather than assuming uniform effects based on global religiosity.

In conclusion, the evidence suggests that religious language in suicide-related Reddit

posts is associated with more negative emotional framing and reduced likelihood of positive emotional expression, even after controlling for structural text features. These results emphasize the importance of emotional nuance when interpreting religious discourse in contexts of acute psychological distress. Researchers and practitioners should consider not only the mere presence of religious language but its affective valence and contextual meaning in crisis narratives.

Limitations and Recommendations. Several limitations necessitate consideration. First, the study relied on keyword-based identification of religious language. Although lexicon-based approaches are transparent and replicable, they may fail to capture implicit spiritual themes, metaphorical language, or culturally specific religious expressions. Second, sentiment analysis was conducted using VADER, a lexicon-based model optimized for social media text. While appropriate for large-scale analysis, it cannot fully capture nuanced theological meanings, sarcasm, contextual irony, or complex narrative arcs. Third, the dataset was not geographically restricted and does not represent a specific cultural or national context. Religious expression and suicide discourse may vary substantially across cultural settings, limiting contextual specificity. Fourth, the cross-sectional and observational design precludes causal inference; religious language may co-occur with negativity without directly contributing to it. Finally, suicide labeling in the dataset reflects computational classification rather than clinical diagnosis, which may introduce heterogeneity in the operationalization of suicidality.

Despite these limitations, the study offers important implications for multiple groups beyond academic researchers.

For mental health professionals, particularly those working in crisis intervention or online counseling contexts, the findings underscore the importance of assessing not only whether clients identify as religious, but how religion is emotionally framed in their narratives.

Religious language may signal spiritual struggle, guilt, or existential distress rather than comfort. Clinicians should therefore explore the affective tone of religious references and assess whether they function as protective meaning systems or sources of additional burden.

For faith leaders and pastoral counselors, the results highlight the need to recognize that individuals invoking religious themes in contexts of suicidal distress may be experiencing spiritual conflict rather than simple faith-based reassurance. Faith communities may benefit from training in recognizing signs of spiritual struggle and integrating evidence-based mental health literacy into pastoral care.

For digital platform moderators and suicide prevention organizations, automated detection systems could incorporate religious language markers as potential contextual signals of heightened emotional intensity. However, interventions should be carefully designed to avoid misclassification or over-pathologizing religious expression.

For policymakers and suicide prevention advocates, the study suggests that culturally sensitive prevention strategies should integrate discussions of spiritual struggle and religious meaning-making, particularly in communities where religious identity is salient.

Future research should extend this work using contextualized natural language processing models, like transformer-based architectures, to capture deeper semantic nuance. Qualitative thematic analyses of religiously coded posts may clarify whether references primarily reflect condemnation, longing for transcendence, appeal to divine help, or moral injury. Longitudinal modeling within online communities could examine whether religious language predicts changes in distress trajectories over time. Cross-cultural studies may further explore whether the emotional valence of religious discourse varies across

socioreligious contexts, including that of the Philippines.

With this, while religion has often been conceptualized as protective in suicide research, the present findings suggest that within active suicide-related discourse, religious language may reflect intensified emotional struggle and complex meaning-making processes. Recognizing this complexity is essential for clinicians, faith leaders, digital health systems, and suicide prevention initiatives seeking to respond effectively and compassionately to individuals in crisis.

Author contributions. The researcher conceptualized and designed the study, conducted data management and computational analyses, interpreted the findings, and drafted and revised the manuscript.

Conflict of interest. The author declares no conflict of interest.

Funding source. This research received no external funding.

Artificial intelligence use. Computational assistance tools were used to support portions of analytic Python coding, specifically ChatGPT (OpenAI, 2026); however, all analytic decisions, model specifications, and interpretations were independently reviewed and verified by the researcher.

Ethics approval statement. Ethics approval was not required because the study involved secondary analysis of publicly available, de-identified data with no direct participant interaction.

Data availability statement. The dataset analyzed is publicly available through the Kaggle repository: <https://www.kaggle.com/datasets/nikhileswarkomati/suicide-watch>.

Publisher's disclaimer. The views expressed in this article are those of the authors and do not necessarily reflect the views of the publisher.

The publisher disclaims any responsibility for errors or omissions.

analysis. *Research Square*.
<https://doi.org/10.21203/rs.3.rs-5200153/v1>

REFERENCES

- Bjärehed, J., Grönberg, H., & Jarlvik, E. (2023). The different faces of online support: A thematic analysis of responses to help-seeking related to depression and suicidal ideation in public online forums. *Cyberpsychology: Journal of Psychosocial Research on Cyberspace*, *17*(5), Article 5. <https://doi.org/10.5817/CP2023-5-5>
- Chancellor, S., Birnbaum, M. L., Caine, E. D., Silenzio, V., & De Choudhury, M. (2021). Suicide risk and protective factors in online support forum posts: Annotation scheme development and validation. *JMIR Mental Health*, *8*(11), e24471. <https://doi.org/10.2196/24471>
- Dan, E., (2025). Exploring suicide factors in online discourse: Sentiment and thematic analysis of r/SuicideWatch and related data. *ACM Transactions on the Web*, *19*(4), 1–26. <https://doi.org/10.1145/3716546>
- Gandy, L. M., Ivanitskaya, L. V., Bacon, L. L., & Bizri-Baryak, R. (2025). Evaluating automated sentiment analysis methods on social media data. *JMIR Formative Research*, *9*, e57395. <https://doi.org/10.2196/57395>
- Ghahramani, M., Memaryan, N., Ghahari, S., & Malakouti, K. (2024). When spirituality fades: a qualitative exploration of spiritual deficiencies in suicide ideation and attempt. *Mental Health, Religion & Culture*, *27*(6), 571–592. <https://doi.org/10.1080/13674676.2024.2431708>
- Ghahramani, M., Memaryan, N., Ghahari, S., & Malakouti, K. (2025). The spiritual needs of individuals with suicide ideation or suicide attempt: a qualitative content analysis. *Research Square*. <https://doi.org/10.21203/rs.3.rs-5200153/v1>
- Greenberg, J., Solomon, S., & Pyszczynski, T. (1997). Terror management theory of self-esteem and cultural worldviews: Empirical assessments and conceptual refinements. *Advances in Experimental Social Psychology*, *29*, 61–139. [https://doi.org/10.1016/S0065-2601\(08\)60016-7](https://doi.org/10.1016/S0065-2601(08)60016-7)
- Jain, R., Kumar, A., Nayyar, A., Dewan, K., Garg, R., Raman, S., & Ganguly, S. (2023). Explaining sentiment analysis results on social media texts through visualization. *Multimedia tools and applications*, *82*(15), 22613–22629. <https://doi.org/10.1007/s11042-023-14432-y>
- Kaggle. (n.d.). *Suicide and Depression Detection (SuicideWatch and depression subreddit posts; collected via Pushshift API)* [Data set]. Kaggle. <https://www.kaggle.com/datasets/nikhilswarkomati/suicide-watch>
- Lemke, A. W., Davis, E. B., Voytenko, V. L., Cowden, R. G., Chen, Z. J., McConnell, J. M., Pargament, K. I., Phillips, K. P., Marseilles, R., & Wolff, R. P. (2023). Religious/Spiritual Struggles and Suicide Ideation Among Adult Psychiatric Outpatients: A 12-Month Longitudinal Study. *Journal of Affective Disorders Reports*, *14*, 100640. <https://doi.org/10.1016/j.jadr.2023.100640>
- Lundeen, L. A., McCall, J. R., Bradshaw, A. S., LeBlanc, E. L., & Humphrey, L. (2024). Digital Catharsis or Harmful Exposure? A Thematic Analysis of Self-Directed Violence Reddit Posts. *Social Media + Society*, *10*(3). <https://doi.org/10.1177/20563051241263562>

- Mason, K., Moore, M., Palmer, J., & Yang, Z. (2023). Religious Commitment and Intent to Die by Suicide during the Pandemic. *Religions, 14*(10), 1226. <https://doi.org/10.3390/rel14101226>
- OpenAI (2026). ChatGPT (GPT-5.2) [Large language model]. <https://chat.openai.com/>
- Park, C. L. (2010). Making sense of the meaning literature: An integrative review of meaning making and its effects on adjustment to stressful life events. *Psychological Bulletin, 136*(2), 257–301. <https://doi.org/10.1037/a0018301>
- Poorolajal, J., Goudarzi, M., Gohari-Ensaf, F., & Darvishi, N. (2022). Relationship of religion with suicidal ideation, suicide plan, suicide attempt, and suicide death: a meta-analysis. *Journal of research in health sciences, 22*(1), e00537. <https://doi.org/10.34172/jrhs.2022.72>
- Pulgar, C. A., Shaikh, A. N., Shannonhouse, L., & Davis, D. (2022). The Role of Religious Coping in Understanding the Suicide Risk of Older Adults during COVID-19. *Religions, 13*(8), 731. <https://doi.org/10.3390/rel13080731>
- Sørensen, J. B., Thomassen, J. L., Konradsen, F., Meyrowitsch, D. W., Vildekilde, T., Karstad, O. M., Ploug, T., & Kingod, N. R. (2023). Online with suicidal ideation: How individuals communicate in and perceive a peer-to-peer mediated social media group. *Mental Health & Prevention, 32*, 200303. <https://doi.org/10.1016/j.mhp.2023.200303>
- Yan, Y., Li, J., Liu, X., Li, Q., & Yu, N. X. (2024). Identifying Reddit Users at a High Risk of Suicide and Their Linguistic Features During the COVID-19 Pandemic: Growth-Based Trajectory Model. *Journal of Medical Internet Research, 24*. <https://doi.org/10.2196/48907>
- Yeskuatov, E., Chua, S. L., & Foo, L. K. (2022). Leveraging Reddit for Suicidal Ideation Detection: A Review of Machine Learning and Natural Language Processing Techniques. *International journal of environmental research and public health, 19*(16), 10347. <https://doi.org/10.3390/ijerph191610347>
- Zhu, J., Jin, R., Kenne, D. R., Phan, N., & Ku, W. S. (2024). User Dynamics and Thematic Exploration in r/Depression During the COVID-19 Pandemic: Insights from Overlapping r/SuicideWatch Users. *Journal of medical Internet research, 26*, e53968. <https://doi.org/10.2196/53968>