The Impact of Listening to Music on Stress Level for Anxiety, Depression, and PTSD: Mixed-Effect Models and Propensity Score Analysis

Mazin Abdalla, Parya Abadeh, Zeinab Noorian, Amira Ghenai, Fattane Zarrinkalam, *Member, IEEE*, and Soroush Zamani Alavijeh

***Abstract*—The intersection of music and mental health has gained increasing attention, with previous studies highlighting music's potential to reduce stress and anxiety. Despite these promising findings, many of these studies are limited by small sample sizes and traditional observational methods, leaving a gap in our understanding of music's broader impact on mental health. In response to these limitations, this study introduces a novel approach that combines generalized linear mixed models (GLMM) with propensity score matching (PSM) to explore the relationship between music listening and stress levels among social media users diagnosed with anxiety, depression, and posttraumatic stress disorder (PTSD). Our research not only identifies associative patterns between music listening and stress but also provides a more rigorous examination of potential causal effects, taking into account demographic factors such as education level, gender, and age. Our findings reveal that across all mental health conditions, music listening is significantly associated with reduced stress levels, with an observed 21.3% reduction for anxiety, 15.4% for depression, and 19.3% for PTSD. Additionally, users who listened to music were more likely to report a zero stress score, indicating a stronger relaxation effect. Further, our analysis of demographic variations shows that age and education level influence the impact of music on stress reduction, highlighting the potential for personalized interventions. These findings contribute to a deeper understanding of music's therapeutic potential, particularly in crafting interventions tailored to the diverse needs of different populations.**

***Index Terms*—Mental disorder, music intervention, social media analysis, stress level analysis.**

I. Introduction

OVER the past decade, there has been a growing interest in using digital traces from social media to identify the mental states of users [1], [2], [3]. Previous research has primarily focused on analyzing textual content, visual data, and online activity traces to either detect the onset of mental health conditions [3], [4] or to differentiate behavioral dispositions and linguistic patterns used across various mental disorders [5]. These studies have consistently highlighted a strong association between users' mental health and the language used in their posts, with specific vocabulary choices, emotional expressions, and psychometric attributes serving as key predictors of psychological disorders [6]. Such research findings show their potential in complementing the findings of clinical lab trials, demonstrating the value of large-scale social media analysis in advancing mental health research. Recently, in a preliminary research conducted by [7], researchers explored the association between different mental health disorders manifested in social media and music preferences, particularly by analyzing the language used in music lyrics. The results highlight how the language and emotion conveyed in music lyrics differ between users with mental health conditions and those without, suggesting that music preferences and the emotional content of lyrics may provide further insights into mental health.

Recent studies have shown that music plays a significant role in mental health interventions. Recent work has demonstrated the potential of music listening in reducing stress and anxiety, particularly during the COVID-19 pandemic [8]. Research has explored various dimensions of this relationship, including the impact of different music genres on emotion [9], the connection between music preferences and psychological disorders [7], and the role of music engagement in managing stress and anxiety in individuals with different mental health conditions [10], [11]. These findings underscore the potential of music as a tool for promoting mental well-being and mitigating the negative effects of stressors. However, current studies have predominantly been conducted on a limited number of patients under controlled clinical settings or through traditional observational cohort studies relying on questionnaires and self-reported surveys. Thus, despite their valuable insights, they suffer from limitations such as small sample sizes and a potential bias associated with user selection. Moreover, they lack exploiting large-scale datasets that combine various factors such as music preferences, social media activity, and behavioral patterns, leaving a gap in understanding the broader, real-world impact of music listening on stress levels.

Our contribution is multifold. First, we propose a novel approach that integrates generalized linear mixed models (aka GLMM) [12] with propensity score matching (PSM) [13] to investigate the relationship between music listening and stress

Received 29 August 2024; revised 21 February 2025; accepted 10 April 2025. *(Corresponding author: Zeinab Noorian.)*

Mazin Abdalla, Zeinab Noorian, and Amira Ghenai are with the Toronto Metropolitan University, Toronto ON M5B 2K3, Canada (e-mail: znoorian@torontomu.ca).

Parya Abadeh and Fattane Zarrinkalam are with the University of Guelph, Guelph ON N1G 2W1, Canada (e-mail: znoorian@torontomu.ca).

Soroush Zamani Alavijeh is with the McMaster University, Hamilton ON L8S 4L8, Canada (e-mail: znoorian@torontomu.ca).

Digital Object Identifier 10.1109/TCSS.2025.3561073

levels in individuals with mental health disorders, specifically focusing on anxiety, depression, and PTSD from social media data. GLMM allows us to account for repeated measures and individual differences, while PSM helps control for confounding variables, enhancing the robustness of our findings and addressing both association and potential causality within observational data. Second, we are among the first to leverage social media data to study the intersection of music listening and mental health, providing a large-scale, real-world perspective on how music influences stress levels. Our study takes a broad approach by analyzing its effects across an extensive and varied timeframe, ensuring that findings are not limited to singular stress-inducing events. Furthermore, our study not only examines the general impact of music on stress but also explores variations across different demographic groups, including education level, gender, and age, offering deeper insights into personalized music-based interventions.

The research aims to address the following research questions.

1) *RQ1:* Does listening to music have a positive effect on reducing stress levels in individuals with mental health disorders, and how does this effect vary among different disorders such as depression, anxiety, and PTSD?
2) *RQ2:* How does the stress level of users with different mental disorders who listen to music vary across different demographic groups, including education level, gender, and age?

The remainder of the article is organized as follows. Section II reviews the related work. Section III describes the data and the research methodology. Section IV presents the results corresponding to research questions. Section V is dedicated to discussing and presenting the limitations of our study. Finally, Section VI sheds light on future work and concludes the article.

II. Related Works

The stress-reducing effects of music are influenced by factors such as genre, tempo, and emotional content, as well as individual differences and the listening context [14]. Clinically, music has proven effective in reducing anxiety and pain during medical procedures and in improving stress management overall [15].

The complex interplay between mental health and musical engagement, such as listening habits, choices, and preferences, is also a significant area of research [7], [16]. Studies have investigated how individuals with various mental health conditions use music as a tool for managing emotions [17]. For example, emotional reliance on music can intensify during episodes of depression and psychological distress [18]. However, the impact of music on emotions is not always positive and can sometimes worsen pathological symptoms [19].

Research has shown that individuals experiencing high levels of distress may find that music exacerbates their negative moods rather than alleviates them [19]. Additionally, music choices can affect mood and well-being, with some individuals using music in maladaptive ways that lead to rumination and social withdrawal [20], [21]. Specifically, those with depression often gravitate towards sad music [22], [23].

Music therapy has also shown efficacy in treating various mental health disorders [21], [24].

Music-based interventions have been found to significantly improve symptoms of depression [25], sleep quality [26], and quality of life (QOL) [24].

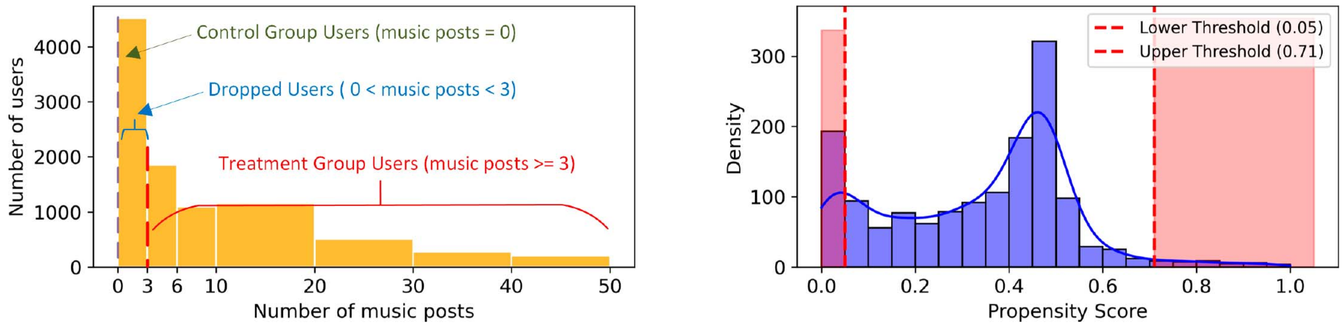
It has also been effective in reducing anxiety and depression levels in individuals with attention deficit hyperactivity disorder (ADHD) and eating disorder [10], [27]. However, most existing studies on mental health have been conducted in clinical settings or relied on traditional methods such as questionnaires and self-reports [28], often focusing on depression and involving small participant groups [29]. To address these limitations, researchers are increasingly turning to social media data to study and detect a broader range of mental health conditions, including depression, anxiety, and stress [3], [30].

Social media offers a rich source of user-generated content that can be analyzed using natural language processing, machine learning, and other computational techniques to identify linguistic patterns, behavioral cues, and emotional states linked to mental health concerns [3], [4].

This method provides a complementary perspective to traditional clinical approaches by offering real-time, large-scale data on mental health experiences and expressions. However, most existing studies focus solely on processing user-generated content on social media to identify psychological disorders, neglecting other types of information, such as music shared by users. Building on clinical research on the interplay between mental health and musical interactions, a recent study [7] has explored the connection between users' self-reported mental health disorders on social media and their musical preferences. Researchers conducted a comprehensive analysis of Twitter data to examine the linguistic characteristics of music favored by individuals with various mental health conditions, comparing these patterns to those of a control group without such conditions.

The findings reveal significant differences in the linguistic and semantic features of music tracks between affected users and the control group, as well as among users with different psychological disorders (Table IV in Appendix A summarizes previous studies and compare them from different technical perspectives). Despite these insights, previous studies have key limitations that limit their applicability. Below, we identify these shortcomings and how our approach addresses them.

1) *Small Sample Sizes and Lack of Generalizability:* Many existing studies rely on self-reported surveys or small-scale clinical trials (e.g., [9], [19], [31]). While these studies provide valuable insights, they often suffer from selection bias, limiting their generalizability to broader populations [10], [24], [25], [29], [32]. We overcome this limitation by utilizing a large-scale dataset of 13 597 Twitter users with self-reported mental health diagnoses (anxiety, depression, PTSD), offering a more representative and ecologically valid analysis of music's impact on stress.
2) *Lack of Causal Insights in Observational Studies:* Most prior research primarily explores correlation-based relationships between music and mental health outcomes (e.g., [7], [19], [24]), rather than establishing causal



(a) (b)

Fig. 1. (a) Frequency of users for number of music posts. (b) Propensity score distribution for anxiety group (shaded region represents those dropped in our analysis).

links. The absence of robust methodologies to control for confounding variables reduces the reliability of conclusions drawn from these studies. We incorporate PSM to control for confounders such as users' activity and psycholinguistic features. By simulating a quasiexperimental design, our study moves beyond mere association to provide stronger evidence of a potential causal effect of music on stress reduction.

3) *Lack of Consideration for Demographic Variability:* Prior research largely generalizes the impact of music across all individuals without accounting for demographic differences (e.g., [11], [15], [26], [28]). However, stress responses to music may vary by age, gender, and education level, which remain underexplored in existing literature. We conduct a demographic-based analysis to examine variations in stress reduction across different age groups, genders, and education levels. This allows us to provide personalized insights into how music-based interventions can be tailored for different populations.
4) *Reliance on Traditional Statistical Methods:* Several previous studies (e.g., [24], [28], [31]) rely on basic statistical correlations and randomized controlled trials (RCTs) [9], [10], [11], [25], [26] rather than advanced modeling techniques to analyze music's effects on stress. These approaches often fail to account for individual variability and repeated measures in user behavior. We employ generalized linear mixed models (GLMM), which account for intraindividual correlations and repeated measures, ensuring greater statistical rigor and robustness in our findings.

By integrating large-scale data, PSM for bias reduction, GLMM for statistical robustness, and NLP-based stress assessment, our study provides a rigorous, scalable, and unbiased evaluation of how music influences stress in individuals with mental health conditions. These advancements contribute significantly to computational social science and mental health research.

III. Research Methodology

A. Dataset

We followed the data collection methodology similar to the approaches used by [7], [33]. Our study focused on Twitter users who publicly shared their self-reported diagnoses related to mental health disorders, specifically depression, anxiety, and PTSD. These three disorders were chosen as they are among the most common mental health conditions [34].

We identified disorder groups by processing tweets that explicitly mention a diagnosis using high-precision patterns between December 2019 and February 2022 (see Table I in [7]. Then, we manually annotated the data to remove users with multiple conditions and false positive, achieving Cohen's Kappa score of 0.99 among three annotators. Following the steps presented [7], we collected users' timelines and analyzed them to observe any indications of music engagement. This process included searching for tweets containing links to music from platforms such as Spotify, SoundCloud, and Apple Music. Fig. 5 depicts the data collection process and Table VII in Appendix C summarizes the number of identified users with the three targeted mental health disorders (depression, anxiety, and PTSD) as well as the total number of their tweets and the number of music-related tweets within each disorder group.

B. Control and Treatment Group

In this study, we focused on individuals with mental health conditions, aiming to compare those who engaged with music (treatment group) to those who did not (control group). Out of 13 597 users with diagnosed conditions, 3693 shared at least one music post while 9904 did not. To ensure the presence of a consistent pattern of music listening, we established a threshold requiring at least three music posts in users' timelines.[1] Fig. 1(a) illustrates this process, showing how we included only users with more than three music posts while excluding those below this threshold. Fig. 2(a) shows the complete process of our framework.

Treatment users were randomly matched with control users based on two criteria: 1) sharing the same mental health diagnosis; and 2) having diagnosis dates within ten days of each other. Each Treatment user was matched with up to two control users who met these criteria. The distribution of participants across mental health disorders is depicted in Table VIII in Appendix C.

[1]While some users shared the same track multiple times, the overall music shared remained diverse.

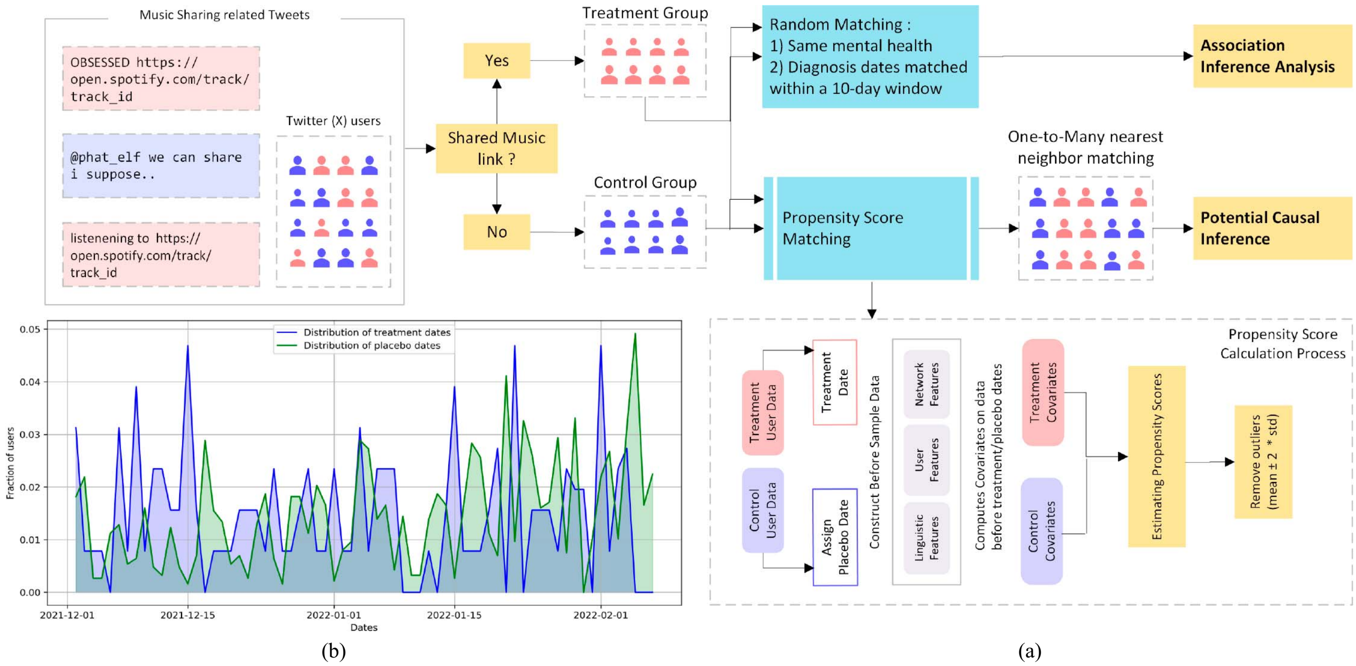


Fig. 2. (a) Causal and association inferences process. (b) Distribution of treatment and placebo dates for all disorders.

C. Propensity Score Calculation Process

1) Constructing Before and After Samples: In this study, we aimed to assess the potential causal impact of music listening on stress levels among Twitter users with mental health conditions using PSM. The first music postdate was the "treatment date" for Treatment users, while a "placebo date" was assigned to the Control group to match the timing [5].

Before assigning placebo dates to Control users, we ensured each date met three criteria: 1) it was after the user joined Twitter; 2) it was after the user's self-reported diagnosis date; and 3) the user's timeline had more than the average number of tweets (i.e., 300) before the date. Control users without a suitable placebo date were excluded from the analysis ($n = 109$). To ensure comparability, we non-parametrically matched treatment and placebo date distributions within each mental health condition. The Kolmogorov–Smirnov test confirmed the similarity ($D < 0.03$ for all disorders), mitigating potential temporal confounds. This process split user timelines into "before" and "after" periods for both groups, allowing stress level comparisons over time. Fig. 2(b), shows the distribution of treatment and placebo dates for all disorders.

2) Computing Covariates: The covariates in this study are selected to capture psycholinguistic, behavioral, and social dimensions of user activity on social media. Inspired by prior research in mental health studies [5], we categorize these covariates into three main groups: *linguistic features*, *user activity metrics*, and *network engagement* indicators. *Linguistic features* are derived from the linguistic inquiry and word count (LIWC) framework [35], widely used for detecting psychological and cognitive patterns in text. From 72 LIWC categories, we refine our selection to 12 high-level categories, including linguistic dimensions, psychological processes, social interactions, cognitive and perceptual processes, biological references, drives, time orientation, relativity, personal concerns, informal language, and grammatical variations. These features provide insights into users' mental states and emotional expressions.[2] User activity metrics and network engagement features capture behavioral patterns and online interactions. *User activity* metrics include tweet frequency, retweets, replies, likes, and weekly/monthly posting trends. *Network engagement* features measure follower and following counts and account longevity, offering insights into a user's social presence and connectivity.

3) Propensity Score Analysis: We used PSM to reduce confounding and estimate the causal effect of music listening on stress. PSM, a method for addressing selection bias in observational studies, estimates the likelihood of treatment assignment (music listening) based on observed covariates, allowing for the creation of comparable treatment and control groups. A logistic regression model calculated propensity scores using pretreatment covariates, including user activity details, user network behaviors, and linguistic features. Outliers beyond ± 2 standard deviations were removed. Propensity score distributions, shown in Fig. 1(b) for anxiety, were similar across other mental health disorders. We then applied one-to-many nearest neighbor matching [36] without replacement to pair Treatment and Control group members, balancing covariate distributions and reducing selection bias. This matching approach was chosen because it effectively pairs individuals with similar propensity scores, minimizing differences in observed characteristics between groups. Nearest-neighbor matching also increases statistical power by maximizing the number of Control users

[2]Due to page limitations, please refer to [35] for a complete description of each LIWC category.

available for comparison, leading to a more precise estimation of the treatment effect while maintaining a reasonable balance in covariate distributions (see Table VIII in Appendix C for the distribution of participants across disorders after matching).

4) Quality of Matching: To assess the quality of matching between the treatment and control groups, we used the standardized mean difference (SMD) as our primary metric. SMD is a widely used measure for evaluating covariate balance in causal inference studies [5], [37], as it quantifies the difference in means between groups while accounting for pooled standard deviations. We adopted the SMD threshold of 0.20 [37] to determine acceptable balance between covariates. After applying PSM, we calculated SMD values for all covariates. The results showed that all covariates fell below the threshold, confirming that the matching process effectively minimized confounding effects and ensured comparability between the groups.

D. Defining and Measuring the Variables of Interest

1) Stress Score: The emotion and language in Twitter posts can indicate stress and anxiety levels [5], [37]. To assess stress levels, we employed TensiStrength [38], a widely used tool for detecting stress and relaxation in social media text. TensiStrength analyzes text using predefined stress and relaxation lexicons, assigning scores from –1 (no stress) to –5 (extreme stress). It incorporates spelling correction, booster words (e.g., very stressed), and negation handling (e.g., not happy) to refine sentiment detection. The model also considers punctuation emphasis (e.g., stressed!!!) and elongated words (e.g., sooo tired) and emoticons to classify stress levels more accurately. The classifier has been validated in prior studies on social media-based stress detection [39], [40]. Before using the TensiStrength API, we cleaned and preprocessed tweet as recommended [38], and only English tweets were analyzed. To simplify interpretation, we rescaled these scores to a 0–4 scale, where 0 represents the lowest stress level and 4 represents the highest. The stress score for the Treatment group was computed from users' social media posts following music engagement. To perform our analysis, we identified the time points when users listened to music and collected all tweets posted within the next 24 h. The final stress score was determined by applying a majority vote across all tweets during this period. For the control group, each treatment user's music listening event was matched with a Control user from the same mental disorder group who had posted within the same 24-h window and their stress score was computed using the same method as the treatment group.

2) Social-Demographic Information Inference: This section outlines the methodology for inferring age, gender, and education level from Twitter user data.

Age and Gender Inference: To infer age and gender, we built a labeled dataset from explicit demographic information in user profiles. For age, we used regular expressions on cleaned user descriptions to extract explicit mentions of age (e.g., "23 years old") and categorized users into "Gen-Z" (under 25) and "Millennials" (25 and older). For gender, we first scanned user descriptions for gender-specific terms such as "woman/girl" or "man/boy." Then, we matched first names against the U.S. Social Security Administration (SSA) Baby Names database, excluding names with less than a 95% gender assignment probability to reduce ambiguity.

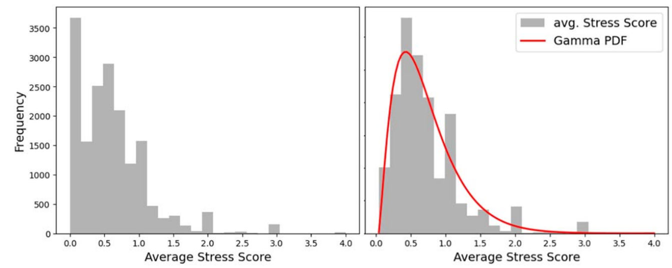


Fig. 3. Distribution of stress score: anxiety as an example.

The labeled dataset was used to train machine learning models to predict demographic attributes for users who didn't explicitly disclose them. We engineered features from Twitter data, including user metadata (account age, followers, tweeting frequency), linguistic features (language patterns, emoji use, lexical diversity), and textual features. These features were chosen based on their proven effectiveness in inferring demographics in previous studies [5]. We trained an XGBoost classifier using an 80% training set and 10-fold cross-validation to optimize the model and prevent overfitting. The age inference model achieved 79.9% accuracy, with F1-scores of 0.83 for Gen-Z and 0.67 for Millennials. The gender inference model had a 64% accuracy, with F1-scores of 0.65 for females and 0.64 for males. Our models' performance aligns with previous studies on demographic inference from social media data [41], validating our feature selection and training approach.

Education level: We inferred education level using the automated readability index (ARI), which estimates the U.S. grade level required to comprehend text and has been used in prior studies [5]. Users with ARI scores below 9 were classified as "Middle/Elementary School," and those scoring 9 or above as "High School/College." To validate ARI outcomes, we calculated the Flesch–Kincaid grade level, confirming ARI's robustness with a Spearman correlation of 0.97.

E. Statistical Analysis

1) Model Selection and Rationale: The primary objective of this study is to investigate the impact of music listening on stress levels among individuals with depression, anxiety, and PTSD. Given that each user in our dataset has multiple sessions of music listening, we needed a statistical model that accounts for individual differences and the repeated measures present in the data. To meet these requirements, we employed a generalized linear mixed-effects model (GLMM) [12], which is well-suited for modeling continuous outcomes like stress levels while accounting for intraindividual correlations due to repeated measurements.

Considering the significant number of zero stress scores in our dataset, as shown in Fig. 3, we utilized a zero-inflated GLMM (ZIGLMM) [42]. This model addresses the excess zeros by modeling the data as a combination of two processes: one generating the zero outcomes and another generating positive stress levels, thereby improving the accuracy of parameter estimates and overall model fit [43].

2) Model Specification: To address our research questions on the impact of music listening on stress levels in individuals with mental disorders, we structured our models as follows:

Fixed Effects: The primary variable in our study–music listening– was modeled as a binary fixed effect (0 = no music listening, 1 = music listening). This fixed effect allows us to estimate the average effect of music listening on stress levels across the population, directly addressing RQ1. Additionally, we included demographic variables such as age, gender, and education level as fixed effects. These demographic factors are crucial in addressing RQ2 to explore their interaction with music listening on stress and control for potential confounding influences.

Random Effects: We included random intercepts at the user level to account for individual variability not captured by fixed effects, such as personal history and lifestyle, ensuring unbiased estimates of the effects of music listening and demographics on stress.

We developed two models to address our research questions.

1) *Model 1:* This model assesses the association between music listening and stress levels across the entire dataset. Users in the treatment and control groups were randomly matched within each disorder category. The ZIGLMM for this analysis includes a random intercept for each user and fixed effects for the covariates such as age, gender, and education.
2) *Model 2:* This model targets the subset of users matched through PSM to explore potential causal relationships between music listening and stress within different mental disorder groups. PSM was conducted separately for each disorder to control for confounders, reducing selection bias by 15%, 13%, and 17% of users in the anxiety, depression, and PTSD groups, respectively. Like model 1, the ZIGLMM includes a random intercept and fixed effects.

3) Formalization: We formalized the above-mentioned models to address RQ1 and RQ2 as follows.

Given that the stress scores Y_{ij} are modeled using a Zero-Inflated Gamma distribution as follows:

$$Y_{ij} \sim \text{Gamma}(\mu_{ij}, \phi). \tag{1}$$

For RQ1, we modeled the linear predictor (aka the conditional model) for the mean stress score of user i (μ_{ij}) as

$$\log(\mu_{ij}) = \beta_0 + \beta_1 \text{Group}_i + u_i. \tag{2}$$

The zero-inflation component was represented as

$$\log\left(\frac{\pi_{ij}}{1-\pi_{ij}}\right) = \gamma_0 + \gamma_1 \text{Group}_i + u_i \tag{3}$$

where

1) Group_i represents the mental disorder group (anxiety, PTSD, or depression);
2) β_1,γ_1 are the fix effect coefficients for group_i;
3) π_{ij} represents the probability of a zero stress score;
4) u_i is the random intercept for user i.

We formalized RQ2 to understand the impact of demographic variables on the stress level as follows:

$$\log(\mu_{ij}) = \alpha_0 + \sum_{m=1}^{M} \alpha_m X_{im} + u_i. \tag{4}$$

The zero-inflation component is specified as follows:

$$\log\left(\frac{\pi_{ij}}{1-\pi_{ij}}\right) = \delta_0 + \sum_{m=1}^{M} \delta_m X_{im} + u_i \tag{5}$$

where

1) X_{im} represents the value of the mth demographic variable (such as gender, age group, or education level) for user i;
2) δ_m and α_m are the fixed effect coefficients corresponding to the mth demographic variable.

4) Model Fit Evaluation: To assess the validity of our model, we examined Pearson residuals versus fitted values to evaluate linearity and heteroscedasticity [44]. The residuals were randomly scattered around zero, indicating that the model appropriately captured variance with minimal bias. Additionally, we assessed zero-inflation to ensure that the model correctly accounted for the structure of the data. These validation checks confirm that our statistical model is well-fitted and effectively captures the relationship between music listening and stress.

IV. Results

A. Basic Statistics

We analyzed user activity and music engagement patterns to understand behavioral differences between the treatment and control groups. Table V in Appendix B shows that treatment group users post more tweets on average, with PTSD users being the most active. The control group has shorter tweets on average. Table VI in Appendix B highlights that PTSD users have the highest average number of music sessions per user (26.23), followed by those with anxiety (21.03) and depression (20.18). The PTSD group also shows the highest variability in session counts. Spotify is the most popular platform across all disorders. Fig. 4 in Appendix B reveals that users with lower education levels dominate across all disorders, with the PTSD group having the highest proportion of users with higher education. The gender distribution shows more females than males across all disorders, with this trend consistent across groups. The age distribution indicates that Gen-Z and younger individuals are the majority, especially in the depression and anxiety groups.

B. RQ1: Impact of Listening to Music on Stress in Different Mental Disorders

In this section, we present the results of two models, each addressing a different aspect of the relationship between music listening and stress in individuals with various mental disorders.

1) Model 1: We explored the *association* between music listening and stress levels across different mental disorders by comparing the treatment group (music listeners) with the control group (nonmusic listeners).

TABLE I
CONDITIONAL AND ZERO-INFLATION MODELS FOR RANDOM MATCHING- RQ1

	Conditional Model					Zero-Inflation Model				
All Disorders										
Predictors	**Estimate**	**SE**	**CI**	**Pr(> |Z|)**	**Significance**	**Estimate**	**SE**	**CI**	**Pr(> |Z|)**	**Significance**
(Intercept)	–0.248	0.007	[–0.26, –0.23]	0.000	Highly significant	–3.514	0.077	[–3.67, –3.36]	0.000	Highly significant
Group	–0.194	0.012	[–0.22, –0.17]	0.000	Highly significant	0.631	0.089	[0.46, 0.81]	0.000	Highly significant
Anxiety										
Predictors	**Estimate**	**SE**	**CI**	**Pr(> |Z|)**	**Significance**	**Estimate**	**SE**	**CI**	**Pr(> |Z|)**	**Significance**
(Intercept)	–0.239	0.015	[–0.27, –0.21]	0.000	Highly significant	–3.363	0.151	[–3.66, –3.07]	0.000	Highly significant
Group	–0.217	0.024	[–0.26, –0.17]	0.000	Highly significant	0.606	0.176	[0.26, 0.95]	0.001	Highly significant
Depression										
Predictors	**Estimate**	**SE**	**CI**	**Pr(> |Z|)**	**Significance**	**Estimate**	**SE**	**CI**	**Pr(> |Z|)**	**Significance**
(Intercept)	–0.275	0.010	[–0.29, –0.26]	0.000	Highly significant	–3.517	0.105	[–3.72, –3.31]	0.000	Highly significant
Group	–0.173	0.016	[–0.21, –0.14]	0.000	Highly significant	0.632	0.121	[0.39, 0.87]	0.000	Highly significant
PTSD										
Predictors	**Estimate**	**SE**	**CI**	**Pr(> |Z|)**	**Significance**	**Estimate**	**SE**	**CI**	**Pr(> |Z|)**	**Significance**
(Intercept)	–0.198	0.014	[–0.23, –0.17]	0.000	Highly significant	–3.667	0.168	[–4.00, –3.34]	0.000	Highly significant
Group	–0.215	0.024	[–0.26, –0.17]	0.000	Highly significant	0.645	0.195	[0.26, 1.03]	0.001	Highly significant

Table I shows the estimated effect of Treatment on stress scores, with the control group serving as the reference. Overall, and across all disorders, the intercept is highly significant in the conditional model, indicating that the baseline stress score for the control group is statistically different from zero.

The results indicate that music listening is associated with a statistically significant reduction in stress across all disorders. Specifically, the Treatment group exhibited a 17.6% reduction in stress scores compared with the control group ($\beta = -0.194$, 95% CI $[-0.22, -0.17]$, $p < 0.001$). This reduction suggests that music listening has a measurable and meaningful calming effect on individuals with mental health disorders.

For individuals with anxiety, music listening resulted in a 19.5% reduction in stress ($\beta = -0.217$, 95% CI $[-0.26, -0.17]$, $p < 0.001$). This finding aligns with prior research on music's role in reducing physiological markers of stress, such as cortisol levels and heart rate, particularly in high-anxiety situations. This suggests that integrating personalized music-based interventions could provide significant benefits to individuals experiencing anxiety symptoms.

The depression group showed a 15.9% reduction in stress ($\beta = -0.173$, 95% CI $[-0.21, -0.14]$, $p < 0.001$). Given that depression is often accompanied by persistent low mood and negative affect, the stress-relieving effect of music could contribute to improved emotional regulation.

For those with PTSD, the 19.3% reduction in stress ($\beta = -0.215$, 95% CI $[-0.26, -0.17]$, $p < 0.001$) highlights the potential of music as a tool for alleviating symptoms related to trauma. The significant reduction across these groups reinforces the potential of music listening as a scalable, accessible intervention for diverse mental health conditions.

The zero-inflation model also supports these findings. Music listeners were 87.9% ($\beta = 0.631$, 95% CI *[*0.46, 0.81*]*, $p < 0.001$) more likely to report a zero stress score compared with nonlisteners, indicating moments of complete emotional relief. This result suggests that music listening not only reduces stress but may also help individuals achieve brief periods of relaxation and calm, which could be vital for mental health recovery.

Additionally, the model's overall results, as well as those across individual disorders, reveal notable variability among users, as indicated by the variance and standard deviation of the random effects (authors) in both the conditional and zero-inflation components (see Table IX in Appendix C).

2) Model 2: We explored the potential *causal* relationship between music listening and stress reduction by analyzing propensity score-matched users. From Table II we notice that, the intercept in the conditional model is highly significant, consistent with the results observed in random matching.

Across all disorders, the treatment group exhibited 17.8% less stress compared with the control group ($\beta = -0.196$, 95% CI $[-0.22, -0.17]$, $p < 0.001$).

For disorder specific findings, music listening was associated with a 21.3% reduction in stress for individuals with anxiety ($\beta = -0.240$, 95% CI $[-0.29, -0.19]$, $p < 0.001$), a 15.4% reduction for those with depression ($\beta = -0.167$, 95% CI $[-0.20, -0.13]$, $p < 0.001$), and a 19.3% reduction for those with PTSD.

In all cases, the zero-inflation model also showed positive "group" coefficients across all disorders, suggesting that users who listened to music are more likely to report a zero-stress score; however, these results were not significant.

Additionally, we observed a notable variability in both stress scores (std = 0.344) and zero-inflation (std = 2.344) across different authors (see Table IX in Appendix C for details), indicating individual differences in how stress is experienced and reported by users, which justifies using a generalized mixed linear model.

We visualized the average stress scores distribution across the anxiety disorder for both control and treatment groups (see Fig. 6 in Appendix C). The Treatment group's histogram shows a lower and narrower peak, indicating that music listeners generally experienced lower and more consistent stress levels compared with nonlisteners. Similar patterns were observed for the other two

TABLE II
CONDITIONAL AND ZERO-INFLATION MODELS FOR PSM- RQ1

	Conditional Model					Zero-Inflation Model				
All Disorders										
Predictors	**Estimate**	**SE**	**CI**	**Pr(> |Z|)**	**Significance**	**Estimate**	**SE**	**CI**	**Pr(> |Z|)**	**Significance**
(Intercept)	–0.249	0.008	[–0.27, –0.23]	0.000	Highly significant	–2.859	0.065	[–2.99, –2.73]	0.000	Highly significant
Group	–0.196	0.013	[–0.22, –0.17]	0.000	Highly significant	0.114	0.082	[–0.05, 0.27]	0.161	Not significant
Anxiety										
Predictors	**Estimate**	**SE**	**CI**	**Pr(> |Z|)**	**Significance**	**Estimate**	**SE**	**CI**	**Pr(> |Z|)**	**Significance**
(Intercept)	–0.217	0.016	[–0.25, –0.18]	0.000	Highly significant	–2.767	0.129	[–3.02, –2.51]	0.000	Highly significant
Group	–0.240	0.026	[–0.29, –0.19]	0.000	Highly significant	0.133	0.164	[–0.19, 0.45]	0.417	Not significant
Depression										
Predictors	**Estimate**	**SE**	**CI**	**Pr(> |Z|)**	**Significance**	**Estimate**	**SE**	**CI**	**Pr(> |Z|)**	**Significance**
(Intercept)	–0.282	0.011	[–0.30, –0.26]	0.000	Highly significant	–2.914	0.090	[–3.09, –2.74]	0.000	Highly significant
Group	–0.167	0.017	[–0.20, –0.13]	0.000	Highly significant	0.145	0.112	[–0.07, 0.36]	0.194	Not significant
PTSD										
Predictors	**Estimate**	**SE**	**CI**	**Pr(> |Z|)**	**Significance**	**Estimate**	**SE**	**CI**	**Pr(> |Z|)**	**Significance**
(Intercept)	–0.208	0.017	[–0.24, –0.17]	0.000	Highly significant	–2.829	0.133	[–3.09, –2.57]	0.000	Highly significant
Group	–0.214	0.027	[–0.27, –0.16]	0.000	Highly significant	0.017	0.175	[–0.33, 0.36]	0.924	Not significant

mental health disorders, which suggest that music listening effectively reduces stress and has a uniform impact across individuals.

Note that, we have conducted additional experiments analyzing the impact of different music features (i.e., tempo, sentiment valence, and speechiness) on stress levels across different disorder groups. Due to page limitations, we have included the detailed results of these experiments in Tables XII and XIII in Appendix C, providing further insights into the role of specific music features in stress reduction.

C. RQ2: Analysing Stress Variation by Demographics in Music-Listening Mental Health Groups

In this section, we present the results of our analysis on the association between demographic factors (age, gender, and education level) and stress levels among users who listen to music.[3] As can be seen in Table III, the intercept in both the conditional and zero-inflation models is highly significant, which establishes a baseline stress score for the reference groups (Gen-Z, female, and high education level). The analysis of demographic factors reveals notable variations in stress reduction. Age and education significantly influence how individuals respond to music, whereas gender showed no significant impact.

Age Differences: Millennials and older users experienced approximately 12.3% higher stress levels than Gen-Z users ($\beta = 0.116$, 95% CI [0.07, 0.16], $p < 0.001$). This finding may reflect differences in coping mechanisms, life experiences, or music preferences between age groups. For older adults, integrating familiar or nostalgic music into interventions could potentially enhance its stress-reducing effects.

[3]We did not use causal inference analysis because the focus was on examining the impact of demographics on stress levels in different mental disorder groups who listen to music. Without a control group, causal inference wasn't applicable, so the analysis was conducted on all treatment group users within the context of random matching.

Education: Users with lower education levels reported 23.9% lower stress scores compared with those with higher education ($\beta = -0.273$, 95% CI [–0.33, –0.22], $p < 0.001$). This counterintuitive result warrants further investigation. It may suggest that individuals with lower education levels have different coping strategies or that their music preferences align more closely with stress-reducing genres or styles. This insight could guide future research into socio-economic factors and stress resilience.

Disorder-Specific Demographic Patterns: For individuals with anxiety, older age groups experienced 14.2% higher stress ($\beta = 0.133$, 95% CI [0.04, 0.22], $p < 0.01$). Lower education levels correlated with 24% lower stress scores ($\beta = -0.274$, 95% CI [–0.39, –0.16], $p < 0.001$). This suggests that interventions targeting older adults should incorporate tailored music selections that better meet their preferences and stress responses.

The depression group, older users experienced 17.5% higher stress levels than younger users ($\beta = 0.161$, 95% CI [0.10, 0.22], $p < 0.001$). Similarly, users with lower education levels saw a 26.1% reduction in stress ($\beta = -0.303$, 95% CI [–0.38, –0.22], $p < 0.001$).

For individuals with PTSD, education level was the only significant demographic factor, with lower education levels resulting in a 20.5% reduction in stress scores ($\beta = -0.229$, 95% CI [–0.32, –0.14], $p < 0.001$). Table XI in Appendix C summarizes our key findings for RQ1 and RQ2.

V. DISCUSSION AND IMPLICATIONS

The primary aim of this study was to investigate the impact of music listening on stress levels in individuals with various mental disorders, specifically anxiety, depression, and PTSD.

TABLE III
CONDITIONAL AND ZERO-INFLATION MODELS FOR RANDOM MATCHING- RQ2

	Conditional Model					Zero-Inflation Model				
All Disorders										
Predictors	**Estimate**	**SE**	**CI**	**Pr(> |Z|)**	**Significance**	**Estimate**	**SE**	**CI**	**Pr(> |Z|)**	**Significance**
(Intercept)	–0.252	0.029	[–0.31, –0.20]	0.000	Highly significant	–3.195	0.175	[–3.54, –2.85]	0.000	Highly significant
Gender	0.034	0.018	[0.00, 0.07]	0.057	Not quite significant	0.088	0.103	[–0.11, 0.29]	0.392	Not significant
Age	0.116	0.021	[0.07, 0.16]	0.000	Highly significant	0.600	0.119	[0.37, 0.83]	0.004	Highly significant
Education	–0.273	0.028	[–0.33, –0.22]	0.000	Highly significant	0.527	0.162	[0.21, 0.85]	0.001	Significant
Anxiety										
Predictors	**Estimate**	**SE**	**CI**	**Pr(> |Z|)**	**Significance**	**Estimate**	**SE**	**CI**	**Pr(> |Z|)**	**Significance**
(Intercept)	–0.272	0.061	[–0.39, –0.15]	0.000	Highly significant	–3.017	0.350	[–3.70, –2.33]	0.000	Highly significant
Gender	0.059	0.037	[–0.01, 0.13]	0.117	Not significant	–0.030	0.208	[–0.44, 0.38]	0.885	Not significant
Age	0.133	0.045	[0.04, 0.22]	0.003	Significant	0.834	0.242	[0.36, 1.31]	0.001	Highly significant
Education	–0.274	0.059	[–0.39, –0.16]	0.000	Highly significant	0.437	0.335	[–0.22, 1.09]	0.192	Not significant
Depression										
Predictors	**Estimate**	**SE**	**CI**	**Pr(> |Z|)**	**Significance**	**Estimate**	**SE**	**CI**	**Pr(> |Z|)**	**Significance**
(Intercept)	–0.235	0.043	[–0.32, –0.15]	0.000	Highly significant	–3.289	0.256	[–3.79, –2.79]	0.000	Highly significant
Gender	0.040	0.025	[–0.01, 0.09]	0.107	Not significant	0.243	0.138	[–0.03, 0.51]	0.078	Not quite significant
Age	0.161	0.031	[0.10, 0.22]	0.000	Highly significant	0.686	0.166	[0.36, 1.01]	0.000	Highly significant
Education	–0.303	0.042	[–0.38, –0.22]	0.000	Highly significant	0.516	0.238	[0.05, 0.98]	0.030	Marginally significant
PTSD										
Predictors	**Estimate**	**SE**	**CI**	**Pr(> |Z|)**	**Significance**	**Estimate**	**SE**	**CI**	**Pr(> |Z|)**	**Significance**
(Intercept)	–0.237	0.049	[–0.33, –0.14]	0.000	Highly significant	–3.105	0.333	[–3.76, –2.45]	0.000	Highly significant
Gender	0.003	0.037	[–0.07, 0.07]	0.945	Not significant	–0.123	0.229	[–0.57, 0.33]	0.592	Not significant
Age	0.012	0.039	[–0.06, 0.09]	0.752	Not significant	0.236	0.245	[–0.24, 0.72]	0.335	Not significant
Education	–0.229	0.047	[–0.32, –0.14]	0.000	Highly significant	0.573	0.302	[–0.02, 1.16]	0.058	Not quite significant

Additionally, we explored how demographic factors such as age, gender, and education level influence this relationship. Results showed insights into the effects of music on stress, as well as the demographic variations in these effects.

Our study found that listening to music significantly reduces stress levels across individuals with anxiety, depression, and PTSD, with reductions ranging from 15.4% to 21.3% depending on the disorder. This degree of reduction is comparable to improvements observed in therapeutic interventions such as mindfulness practices and cognitive-behavioral therapy. For individuals experiencing high stress, this reduction can provide critical moments of relief that contribute to long-term recovery. Music listeners were also more likely to report zero stress, particularly in the *association* analysis where they were 87.9% more likely to do so compared with nonlisteners. This finding is particularly noteworthy as it suggests that music may help individuals achieve temporary states of emotional calm–moments that are essential for coping with chronic stress and improving overall quality of life. These moments of zero stress might reflect periods of full recovery or resilience, highlighting music's unique ability to offer immediate, tangible relief.

These findings align with existing research that highlights music's ability to lower physiological stress markers, such as cortisol, and improve emotional regulation, particularly in those with anxiety. For example, a study by [15] demonstrated that listening to relaxing music before a stressor significantly reduced cortisol levels and helped individuals recover from stress more quickly. Similarly, Lin et al. [11] found that music therapy reduced symptoms of anxiety and depression in cancer patients, highlighting the therapeutic potential of music in clinical settings. However, our study adds nuance by showing that the impact of music varies across different disorders and demographic groups. For instance, older adults (Millennials and older) experienced higher stress levels compared with Gen-Z users, despite listening to music, contrasting with some prior research, such as that [45], which suggests older adults might benefit more from music due to nostalgic connections and life experience. Further, the association between lower education levels and reduced stress is somewhat counterintuitive, as higher education is generally associated with better health outcomes. This finding might reflect a more complex interplay of socioeconomic factors and stress coping mechanisms that warrant further investigation.

One possible explanation for this pattern is that individuals with lower education levels may have different stress-coping mechanisms or rely more on music as an accessible and affordable stress management tool compared with clinical alternatives. Alternatively, they may engage with musical genres that have stronger emotional resonance or calming effects.

These findings have several concrete implications for the development of music-based interventions aimed at reducing stress. Given that music listening was shown to have *casual* impact in reducing stress across individuals with anxiety, depression, and PTSD, healthcare providers could incorporate personalized music therapy sessions into treatment plans for these conditions. For instance, in clinical settings, therapists could curate playlists with calming or uplifting music specifically designed to alleviate symptoms of anxiety, such as rapid heartbeat or muscle tension. These playlists could be made available through mobile apps that patients can use during moments of high stress, ensuring that the therapeutic benefits of music are accessible outside of therapy sessions. These findings also have implications for public health initiatives. Community-based programs could promote music listening as a low-cost, scalable intervention for stress management, particularly for underserved populations that may lack access to traditional mental health services. By integrating these insights into mobile health apps and online platforms, healthcare providers can extend the reach of music-based interventions to broader audiences.

However, the demographic differences observed in our study suggest that a one-size-fits-all approach may not be effective. For example, Millennials and older users, who reported higher stress levels, might benefit from interventions that incorporate familiar or nostalgic music, which research has shown to be particularly effective in reducing stress for these age groups [45]. In contrast, Gen-Z users might respond better to more contemporary music that resonates with their cultural context [46]. The differences in stress reduction between disorders suggest that more specialized interventions are needed. For example, individuals with Anxiety might benefit from music that emphasizes slow tempos and steady rhythms, which are known to induce relaxation and reduce physiological arousal [47]. On the other hand, interventions for depression might focus on music that is more uplifting and energizing, helping to counteract the low mood and lethargy that are common in this condition. Tailoring the type of music to the specific symptoms of each disorder could enhance the effectiveness of these interventions.

A. Limitation and Privacy

While our study provides valuable insights into the impact of music on stress reduction across various mental health conditions, several limitations should be acknowledged. First, the study relies on self-reported data for stress levels, which may be subject to bias and inaccuracies. Self-reporting can be influenced by a range of factors, including memory recall and social desirability, which may affect the validity of the results. Another limitation is the generalizability of the findings. The study sample may not fully represent the broader population, particularly regarding demographic diversity such as age, cultural background, and socioeconomic status. As a result, the effectiveness of music as a stress-reduction tool may vary across different populations, and further research is needed to explore these differences. Next, a key limitation of our study is the reliance on social media text analysis for stress assessment, which lacks the precision of clinical evaluations. While TensiStrength [38] provides scalable insights, integrating consultation with medical professionals and physiological stress measures (e.g., cortisol levels) could improve the clinical relevance. Future work could combine social media analysis with clinical validation to offer a more comprehensive and robust understanding of how music affects stress. Additionally, our study is limited by the incompleteness of demographic data. While we analyzed key demographic factors such as age, gender, and education level, other important variables—such as ethnicity, occupation, and socioeconomic background—were not included due to their limited availability in social media data. These factors could influence both stress levels and the effects of music on stress reduction, and their omission may impact the generalizability of our findings. Future studies could integrate more comprehensive demographic data, possibly through survey-based approaches or linkage with external datasets, to better understand the variability in music's impact on stress across different populations.

Our study involved the analysis of user data, which raises several important ethical considerations. First, we are committed to ethical data sharing and will make the dataset available upon request and approval, ensuring that the use of the data aligns with ethical guidelines and protects participant privacy. We have also anonymized all user data, ensuring that individual identities could not be traced back to the dataset.

VI. Conclusion and Future Work

Our study highlights the potential for music to be a powerful tool in managing stress across various mental health conditions. The associative patterns between music listening and stress levels across the general population of users with mental disorders provide valuable insights. Additionally, by rigorously controlling for confounders and matching users based on their propensity to engage in music listening, we offer a more robust examination of potential causal effects. Together, these approaches allow us to draw more reliable inferences about the role of music listening in modulating stress levels, considering both within-subject correlations and potential biases present in observational data. To maximize its impact, future research should refine these interventions by considering demographic and disorder-specific factors. For instance, exploring different music genres, cultural preferences, and personalized playlists could identify the most effective strategies for diverse populations. Longitudinal studies should examine not only the immediate effects but also the long-term benefits of regular music engagement, particularly for chronic conditions such as depression and PTSD. Further, integrating music therapy with other treatments, such as cognitive-behavioral therapy, could be further studied to assess how such combinations enhance overall effectiveness in stress management.

Additionally, while this study focuses on Twitter data for identifying mental health conditions and analyzing music engagement, future work should explore applying this methodology to other platforms that offer both indicators. Platforms such as Reddit, YouTube, and Spotify-integrated social media (e.g., Instagram stories) can provide a richer context for detecting mental health disorders and music listening behavior. YouTube comments and playlists offer insights into users' emotional states through both music engagement and language patterns in user-generated content. Similarly, integrating data from Spotify's shared listening history on social media could enhance the accuracy of music-related interventions. Future research should combine multimodal data sources, such as text, images, and shared music links, to build a more comprehensive framework for identifying mental health states and their relationship with music engagement across different platforms. Finally, future research will combine social media data with controlled studies to validate our findings and address data variability. While social media offers large-scale, real-time insights into mental health trends, controlled settings can provide greater precision and help establish causal relationships. This mixed-method approach will enhance the reliability and depth of our conclusions.

APPENDIX A
RELATED WORKS ADDITIONAL INFORMATION

TABLE IV
COMPARISON OF PREVIOUS STUDIES WITH THE CURRENT STUDY

Study	Data Source	Data Size	Variables Considered	Methodology	Limitation
Lin et al. (2011)	Clinical experiment	98 participants	Anxiety level, biobehavioural indicators	RCT	Small sample size, lack of demographic variability (female-only), reliance on traditional statistical methods.
Thoma et al. (2013)	Clinical experiment	60 participants	Cortisol levels, physiological response	Physiological study	Small sample size, lack of generalizability, lack of demographic variability (female-only), reliance on traditional statistical methods.
Sakka & Juslin (2018)	Survey-based	77 participants	Emotional regulation strategies	Survey-based statistical analysis	Small sample size, lack of generalizability, reliance on traditional statistical methods.
Li (2023)	Clinical experiment	90 participants	Depression, urinary cortisol	RCT, pre-post assessments	Small sample size, lack of generalizability.
Alavijeh et al. (2023)	Twitter dataset	3999 users	Music preferences, linguistic analysis	NLP-based correlation analysis	Lack of causal insights.
Malakoutikhah et al. (2023)	Survey-based	46 participants	Physiological parameters, emotion	RCT, pre-post assessments	Small sample size (undergrad students only), reliance on traditional statistical methods.
J.-I. Park et al. (2023)	Clinical experiment	36 participants	Depression, stress	RCT	Small sample size (ADHD-only), reliance on traditional statistical methods.
Akhshabi et al. (2024)	Clinical experiment	32 participants	Stress	Survey-based statistical analysis	Small sample size, lack of demographic variability (female-only), reliance on traditional statistical methods.
Chen et al. (2024)	Clinical experiment	45 participants	Depression, anxiety	Survey-based statistical analysis	Small sample size (medical students only), reliance on traditional statistical methods.
Morgan & Marroquín (2024)	Survey-based	146 participants	Music as Emotion Regulation, depression, anxiety	Online survey	Small sample size (US adults only), reliance on traditional statistical methods, Lack of causal insights.
Wang et al. (2024)	Clinical experiment	112 participants	Anxiety, depression, QOL, clinical satisfaction	Observational study	Small sample size (elderly patients only), reliance on traditional statistical methods, Lack of causal insights.
Ugurlu et al. (2024)	Clinical experiment	61 participants	Menopausal symptoms, depression, sleep quality	RCT, pre-post assessments	Small sample size, lack of demographic variability (females-only), reliance on traditional statistical methods.
Xiaodan Wang et al. (2024)	Clinical experiment	66 participants	Emotion, heart rate, stress level	RCT, pre-post assessments	Small sample size, lack of generalizability.
Current Study	**Twitter dataset**	**13 597 users**	**Stress**	**GLMM + PSM for causal inference**	Self-reported bias.

APPENDIX B
SUMMARY STATISTICS OF OUR DATASET

TABLE V
SUMMARY OF TWEETING ACTIVITY IN DIFFERENT MENTAL DISORDERS

			Tweets			Words in Tweets	
Group	**Disorder**	**Users**	**Avg/User**	**Min::Max::Std**	**Total Tweets**	**Avg/Tweet**	**Min::Max::Std**
Control	Anxiety	782	903.38	1::3142::936.5	706 442	14.71	1::82::13.0
	Depression	1686	926.92	1::3248::929.5	1 562 783	14.47	1::102::12.8
	PTSD	587	1071.84	1::3204::950.4	629 170	15.77	1::92::13.5
Treatment	Anxiety	457	1811.01	22::3061::779.1	827 631	13.88	1::81::12.3
	Depression	987	1815.18	19::3212::798.4	1 791 580	13.5	1::109::12.1
	PTSD	381	1925.2	16::3102::739.8	733 502	15.44	1::74::13.4

TABLE VI
SUMMARY OF MUSIC LISTENING ACTIVITY IN DIFFERENT MENTAL DISORDERS

		Music Sessions					Online Listening Platforms			
Disorder	Users	Earliest	Latest	Avg./User	Min::Max::Std	Total	Spotify	Apple	SoundCloud	Others
Anxiety	457	2013-04-19	2022-02-09	21.03	3::908::52.33	9609	6894	1784	893	38
Depression	987	2011-09-27	2022-02-07	20.18	3::841::42.10	19 914	15 886	2979	1012	37
PTSD	381	2012-01-31	2022-02-10	26.23	3::833::67.26	9994	6973	2325	657	39

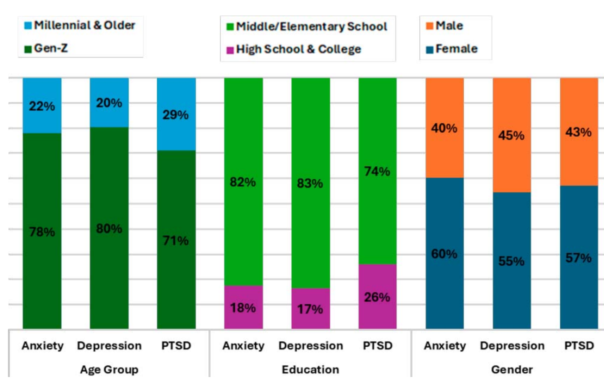


Fig. 4. Demographic information distribution across disorders.

APPENDIX C
ADDITIONAL EXPERIMENTAL RESULTS AND DATA COLLECTION PROCESS

Summary of the random effects from conditional and zero-inflation models for different mental health disorders for random matching and using PSM (shown in Table IX) for RQ1. Table X also show these stats for RQ2. The tables present the number of observations, authors, variance, and standard deviation for both conditional and zero-inflation components.

Tables XII and XIII present the influence of key music features on users' stress levels. We analyzed Tempo [the speed of a musical piece, measured in beats per minute (BPM)], sentiment valence (a measure of how positive or negative a piece of music is), and speechiness (the presence of spoken words in a track). All results were statistically significant, with notable differences observed between users exposed to high versus low levels of tempo, sentiment valence, and speechiness. Overall, users who listened to high-tempo music experienced greater stress reduction compared with those who listened to low-tempo music. Similarly, music with higher sentiment valence and greater speechiness was generally associated with lower stress levels.

TABLE VII
SUMMARY OF USER INFORMATION ACROSS DIFFERENT MENTAL HEALTH GROUPS BEFORE MATCHING

Disorder Groups	User Count	Tweet Count	Music Count
Anxiety	3309	5 336 796	10 354
Depression	6785	11 129 492	21 318
PTSD	3503	5 891 646	10 396
Total	13 597	22 357 934	42 068

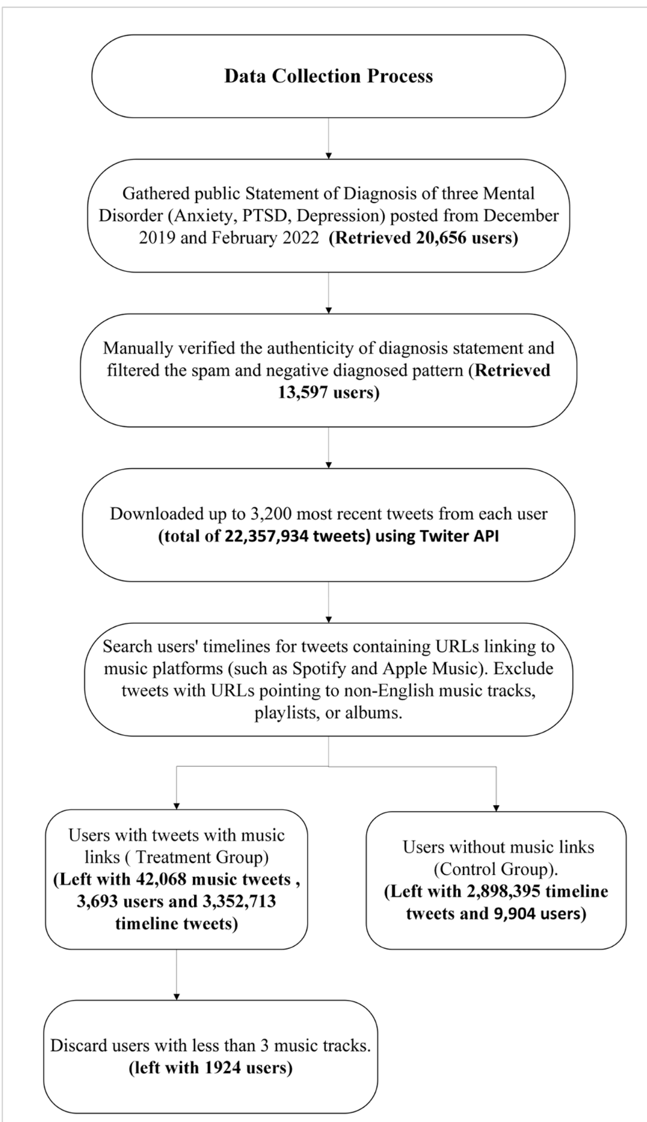


Fig. 5. Dataset collection process.

TABLE VIII
DISTRIBUTION OF USERS ACROSS MENTAL HEALTH GROUPS IN RANDOM MATCHING AND PSM

Disorder/Group		Random Matching	Propensity Score Matching
Anxiety	Control	988	782
	Treatment	494	457
Depression	Control	2056	1686
	Treatment	1028	987
PTSD	Control	804	587
	Treatment	402	381

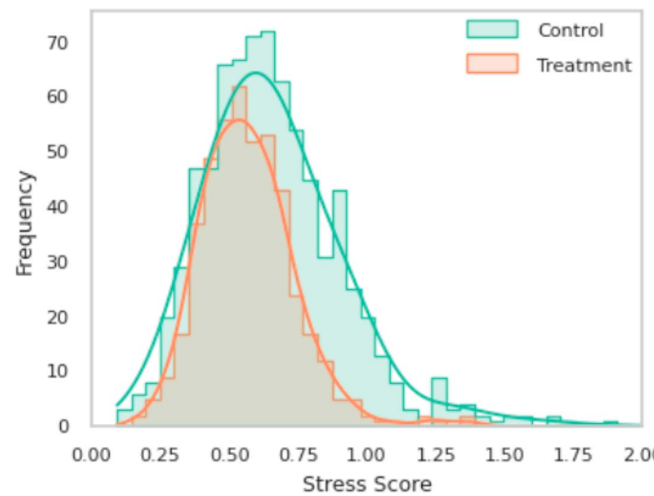


Fig. 6. Stress score distribution: anxiety as an example.

TABLE IX
CONDITIONAL AND ZERO-INFLATION MODELS FOR RANDOM AND PSM- RQ1

Random Matching						
Disorder	Observations	Random Effects				
		Authors	Conditional		Zero-Inflation	
			Variance	Std	Variance	Std
All	70 269	5110	0.118	0.344	5.492	2.344
Anxiety	17 088	1303	0.122	0.350	5.485	2.342
Depression	35 297	2710	0.122	0.350	5.303	2.303
PTSD	17 884	1097	0.103	0.320	5.886	2.426
Propensity Score Matching						
Disorder	Observations	Random Effects				
		Authors	Conditional		Zero-Inflation	
			Variance	Std	Variance	Std
All	35 767	1850	0.107	0.327	2.816	1.678
Anxiety	8665	465	0.114	0.338	2.906	1.705
Depression	17 985	986	0.109	0.330	2.698	1.643
PTSD	9117	399	0.090	0.300	2.922	1.709

TABLE X
CONDITIONAL AND ZERO-INFLATION MODELS- RQ2

Disorder	Observations	Random Effects				
		Authors	Conditional		Zero-Inflation	
			Variance	Std	Variance	Std
All	69 987	4482	0.128	0.357	4.421	2.103
Anxiety	16 969	1145	0.131	0.362	4.557	2.135
Depression	35 009	2423	0.129	0.359	4.366	2.090
PTSD	18 009	914	0.119	0.345	4.355	2.087

TABLE XI
KEY FINDINGS ON THE IMPACT OF MUSIC ON STRESS

Key Finding	Description
1) Music listening reduces stress across mental health disorders	Listening to music significantly reduces stress levels in individuals with anxiety, depression, and PTSD. Stress reduction ranged from 15.4% to 21.3%, depending on the disorder.
2) Music listeners more likely to report zero stress	Individuals who listened to music were 87.9% more likely to report zero stress compared with nonlisteners, indicating a strong association between music listening and lower stress levels.
3) Variability in stress reduction across disorders	The casual impact of music on stress reduction varied by disorder: anxiety (21.3% reduction), PTSD (19.3% reduction), and depression (15.4% reduction).
4) Demographic factors influence stress levels	• *Age:* Older users experienced higher stress levels than younger users for depression and anxiety groups. • Education: Lower education levels were associated with lower stress levels is all three disorder groups. • Gender: No significant effect on stress levels.

TABLE XII
EFFECT OF MUSIC FEATURES ON STRESS REDUCTION

Features	Predictors	High		Low	
		Estimate	95% CI	Estimate	95% CI
Tempo	(Intercept)	–0.26892	[–0.29, –0.25]	–0.28175	[–0.30, –0.26]
	Group	**–0.21628**	[–0.25, –0.19]	**–0.18969**	[–0.22, –0.16]
Valence	(Intercept)	–0.27025	[–0.29, –0.25]	–0.27287	[–0.29, –0.26]
	Group	**–0.22641**	[–0.26, –0.19]	**–0.19630**	[–0.23, –0.17]
Speechiness	(Intercept)	–0.27914	[–0.30, –0.26]	–0.27161	[–0.29, –0.26]
	Group	**–0.21309**	[–0.25, –0.18]	**–0.20471**	[–0.23, –0.18]

Note: The bold entries represent the main variable and its estimate.

TABLE XIII
EFFECT OF MUSIC FEATURES ON STRESS REDUCTION PER DISORDER

Disorder	Features	Predictors	High		Low	
			Estimate	95% CI	Estimate	95% CI
Anxiety	Tempo	(Intercept)	–0.24965	[–0.29, –0.21]	–0.27680	[–0.31, –0.24]
		Group	**–0.26935**	[–0.33, –0.20]	**–0.20216**	[–0.26, –0.14]
	Valence	(Intercept)	–0.25270	[–0.29, –0.22]	–0.26321	[–0.30, –0.23]
		Group	**–0.24974**	[–0.31, –0.19]	**–0.24435**	[–0.31, –0.18]
	Speechiness	(Intercept)	–0.27704	[–0.33, –0.23]	–0.25755	[–0.29, –0.22]
		Group	**–0.23689**	[–0.32, –0.16]	**–0.23693**	[–0.29, –0.18]
Depression	Tempo	(Intercept)	–0.29471	[–0.32, –0.27]	–0.29949	[–0.32, –0.28]
		Group	**–0.19378**	[–0.24, –0.15]	**–0.18694**	[–0.23, –0.15]
	Valence	(Intercept)	–0.29655	[–0.32, –0.27]	–0.29512	[–0.32, –0.27]
		Group	**–0.22204**	[–0.27, –0.18]	**–0.17333**	[–0.21, –0.13]
	Speechiness	(Intercept)	–0.29846	[–0.33, –0.27]	–0.29842	[–0.32, –0.28]
		Group	**–0.20857**	[–0.26, –0.16]	**–0.18725**	[–0.23, –0.15]
PTSD	Tempo	(Intercept)	–0.22609	[–0.26, –0.19]	–0.24069	[–0.28, –0.20]
		Group	**–0.20712**	[–0.27, –0.15]	**–0.18101**	[–0.25, –0.12]
	Valence	(Intercept)	–0.22690	[–0.26, –0.19]	–0.22746	[–0.26, –0.19]
		Group	**–0.20401**	[–0.27, –0.14]	**–0.19977**	[–0.26, –0.14]
	Speechiness	(Intercept)	–0.24001	[–0.28, –0.20]	–0.21998	[–0.25, –0.19]
		Group	**–0.18773**	[–0.26, –0.11]	**–0.20833**	[–0.27, –0.15]

Note: The bold entries represent the main variable and its estimate.

REFERENCES

[1] S. Chancellor and M. De Choudhury, "Methods in predictive techniques for mental health status on social media: a critical review," *NPJ Digit. Med.*, vol. 3, no.1, p. 43, 2020.

[2] S. Ghosh and T. Anwar, "Depression intensity estimation via social media: A deep learning approach," *IEEE Trans. Comput. Social Syst.*, vol. 8, no. 6, pp. 1465–1474, Dec. 2021.

[3] X. Xu, et al., "Mental-LLM: Leveraging large language models for mental health prediction via online text data," in *Proc. ACM InterAct., Mobile, Wearable Ubiquitous Technol.*, vol. 8, no. 1, pp. 1–32, 2024.

[4] D. Zhang, L. Zhou, J. Tao, T. Zhu, and G. Gao, "Ketch: a knowledge-enhanced transformer-based approach to suicidal ideation detection from social media content," *Inf. Syst. Res.*, vol. 36, no. 1, pp. 572–599, 2024.

[5] G. Verma, A. Bhardwaj, T. Aledavood, M. De Choudhury, and S. Kumar, "Examining the impact of sharing COVID-19 misinformation online on mental health," *Sci. Rep.*, vol. 12, no. 1, p. 8045, 2022.

[6] E. A. Ríssola, M. Aliannejadi, and F. Crestani, "Beyond modelling: Understanding mental disorders in online social media," in *Proc. ECIR*, 2020.

[7] S. Z. Alavijeh, F. Zarrinkalam, Z. Noorian, A. Mehrpour, and K. Etminani, "What users' musical preference on twitter reveals about psychological disorders," *Inf. Process. Manage.*, vol. 60, no. 3, 2023, Art. no. 103269.

[8] D. Vidas, J. L. Larwood, N. L. Nelson, and G. A. Dingle, "Music listening as a strategy for managing COVID-19 stress in first-year university students," *Front. Psychol.*, vol. 12, 2021, Art. no. 647065.

[9] A. Malakoutikhah, M. Dehghan, A. Ghonchehpour, M. A. Afshar, and P. Zakeri, "A randomized controlled trial on the effects of different music genres on physiologic parameters and emotion," *Irish J. Med. Sci. (1971)*, vol. 192, no. 5, pp. 2313–2322, 2023.

[10] J.-I. Park et al., "Effects of music therapy as an alternative treatment on depression in children and adolescents with ADHD by activating serotonin and improving stress coping ability," *BMC Complementary Med. Therapies*, vol. 23, no. 1, p. 73, 2023.

[11] M.-F. Lin, Y.-J. Hsieh, Y.-Y. Hsu, S. Fetzer, and M.-C. Hsu, "A randomised controlled trial of the effect of music therapy and verbal relaxation on chemotherapy-induced anxiety," *J. Clin. Nursing*, vol. 20, no. 7-8, pp. 988–999, 2011.

[12] B. Bolker et al., "Generalized linear mixed models: a practical guide for ecology and evolution," *Trends Ecol. Evol.*, vol. 24, no. 3, pp. 127–135, 2009.

[13] J. Liu, L. Xiao, and B. Wang, "varying effects of residential built environment on travel behavior of internal migrants and locals," *Travel Behav. Soc.*, vol. 34, 2024, Art. no. 100692.

[14] M. De Witte, A. Spruit, S. van Hooren, X. Moonen, and G.-J. Stams, "Effects of music interventions on stress-related outcomes: a systematic review and two meta-analyses," *Health Psychol. Rev.*, vol. 14, no. 2, pp. 294–324, 2020.

[15] M. V. Thoma, R. La Marca, R. Brönnimann, L. Finkel, U. Ehlert, and U. M. Nater, "The effect of music on the human stress response," *PLoS One*, vol. 8, no. 8, 2013, Art. no. e70156.

[16] T. L. Golden et al., "The use of music in the treatment and management of serious mental illness: A global scoping review of the literature," *Front. Psychol.*, vol. 12, p. 880, Mar. 2021.

[17] D. E. Gustavson, P. L. Coleman, J. R. Iversen, H. H. Maes, R. L. Gordon, and M. D. Lense, "Mental health and music engagement: review, framework, and guidelines for future studies," *Transl. Psychiatry*, 2021.

[18] S. Garrido and E. Schubert, "Moody melodies: Do they cheer us up? a study of the effect of sad music on mood," *Psychol. Music*, vol. 43, no. 2, pp. 244–261, 2015.

[19] R. M. Morgan and B. Marroquín, "Music listening as emotion regulation: Associations with other emotion regulation strategies and symptoms of depression and anxiety," *Musicae Scientiae*, vol. 28, no. 3, pp. 591–605, 2024.

[20] S. C. Kanagala, T. Schäfer, D. M. Greenberg, and A. Gabińska, "Depression symptoms relationship with music use: investigating the role of trait affect, musical ability, music preferences," *Music Sci.*, 2021.

[21] Q. Tang, Z. Huang, H. Zhou, and P. Ye, "Effects of music therapy on depression: a meta-analysis of randomized controlled trials," *PLoS One*, vol. 15, no. 11, 2020, Art. no. e0240862.

[22] S. Garrido and E. Schubert, "Music and people with tendencies to depression," *Music Perception: Interdisciplinary J.*, vol. 32, no. 4, pp. 313–321, 2015.

[23] K. Wilhelm, I. Gillis, E. Schubert, and E. L. Whittle, "On a blue note: depressed peoples' reasons for listening to music," *Music Med.*, vol. 5, no. 2, pp. 76–83, 2013.

[24] L. Wang, P. Liu, and X. He, "Personalized music therapy for elderly patients with chronic renal failure to improve their quality of life and mental health: a retrospective study," *Noise Health*,, vol. 26, no. 120, pp. 8–13 2024.

[25] X. Wang, Z. Xie, and G. Du, "Research on the intervention effect of vibroacoustic therapy in the treatment of patients with depression," *Int. J. Mental Health Promotion*, vol. 26, no. 2, 2024.

[26] M. Ugurlu, M. V. Şahin, and O. H. Oktem, "The effect of music on menopausal symptoms, sleep quality, and depression: a randomized controlled trial," *Revista da Associação Médica Brasileira*, vol. 70, no. 2, 2024, Art. no. e20230829.

[27] P. Shah, E. Mitchell, S. Remers, S. Van Blyderveen, and H. Ahonen, "The impact of group music therapy for individuals with eating disorders," *Approaches: Interdisciplinary J. Music Therapy*, vol. 15, no. 1, 2023.

[28] M. Akhshabi, N. Namjoo, and H. Khakbaz, "Effectiveness of music interventions in occupational stress of women," *J. Iran. Med. Council*, vol. 7, no. 2, pp. 351–361, 2024.

[29] Q. Chen et al., "Music-based therapeutic interventions for medical school students with emotional regulation and mental health: a pre-post cohort study," *Front. Psychol.*, vol. 15, 2024, Art. no. 1401129.

[30] U. Naseem, M. Khushi, J. Kim, and A. G. Dunn, "Hybrid text representation for explainable suicide risk identification on social media," *IEEE Trans. Comput. Social Syst.*, vol. 11, no. 4, pp. 4663–4672, Aug. 2024.

[31] L. S. Sakka and P. N. Juslin, "Emotion regulation with music in depressed and non-depressed individuals: Goals, strategies, and mechanisms," *Music Sci.*, vol. 1, 2018, Art. no. 2059204318755023.

[32] K. Li, "Research on the therapeutic effect of college students' depression based on music therapy," *J. Educ., Humanities Social Sci.*, vol. 7, pp. 206–211, 2023.

[33] A. Cohan, B. Desmet, A. Yates, L. Soldaini, S. MacAvaney, and N. Goharian, "SMHD: A large-scale resource for exploring online language usage for multiple mental health conditions," in *Proc. ACL*, 2018.

[34] E. A. Ríssola, D. E. Losada, and F. Crestani, "A survey of computational methods for online mental state assessment on social media," *ACM Trans. Comput. Healthcare*, vol. 2, no. 2, pp. 1–31, 2021.

[35] J. W. Pennebaker, R. L. Boyd, K. Jordan, and K. Blackburn, "The development and psychometric properties of LIWC2015," *Tech. Rep.*, 2015.

[36] C. Gosselin and B. Boller, "The impact of retirement on executive functions and processing speed: findings from the Canadian longitudinal study on aging," *Aging, NeuroPsychol., Cogn.*, vol. 31, no. 1, pp. 1–15, 2024.

[37] K. Saha, B. Sugar, J. Torous, B. Abrahao, E. Kıcıman, and M. D. Choudhury, "A social media study on the effects of psychiatric medication use," in *Proc. WSM*, 2019.

[38] M. Thelwall, "Tensistrength: stress and relaxation magnitude detection for social media texts," *Inf. Process. Manage.*, vol. 53, no. 1, pp. 106–121, 2017.

[39] C. Gupta, Sangeeta, N. S. Gill, and P. Gulia, "A deep learning-based framework for analyzing stress factors among working women," in *Proc. Int. Conf. Innov. Data Analytics*, Singapore: Springer, 2022, pp. 225–234.

[40] X. Tao et al., "Towards an Understanding of the Engagement and Emotional Behaviour of Mooc Students Using Sentiment and Semantic Features," *Comput. Educ., Artif. Intell.*, vol. 4, 2023, Art. no. 100116.

[41] B. Onikoyi, N. Nnamoko, and I. Korkontzelos, "Gender prediction with descriptive textual data using a machine learning approach," *Nlp J.*, vol. 4, 2023, Art. no. 100018.

[42] M. E. Brooks et al., "glmmTMB balances speed and flexibility among packages for zero-inflated generalized linear mixed modeling," *R J.*, 2017.

[43] M. E. Brooks et al., "Modeling zero-inflated count data with glmmTMB," *BioRxiv*, p. 132753, 2017.

[44] K. Burak and A. Kashlak, "Bootstrapping generalized linear models to accommodate overdispersed count data," *Stat. Papers*, vol. 65, no. 6, pp. 3769–3788, 2024.

[45] P. Laukka, "Uses of music and psychological well-being among the elderly," *J. Happiness Stud.*, vol. 8, pp. 215–241, 2007.

[46] K. S. McFerran and S. Saarikallio, "Depending on music to feel better: Being conscious of responsibility when appropriating the power of music," *Arts Psychother.*, vol. 41, no. 1, pp. 89–97, 2014.

[47] L. Bernardi, C. Porta, and P. Sleight, "Cardiovascular, cerebrovascular, and respiratory changes induced by different types of music in musicians and non-musicians: the importance of silence," *Heart*, vol. 92, no. 4, pp. 445–452, 2006.