Music Listening, Mental Health, and Stress: A Computational Framework for Personalized Analysis and Recommendation

PARYA ABADEH, University of Guelph, Canada
ZEINAB NOORIAN, Toronto Metropolitan University, Canada
FATTANE ZARRINKALAM, University of Guelph, Canada
AMIRA GHENAI, Toronto Metropolitan University, Canada

This study examines how music listening is associated with stress levels among individuals with mental health conditions (depression, anxiety, PTSD, bipolar) using large-scale social media data. We analyzed 10,264 Twitter users and applied a propensity score–matched mixed-effects modeling framework to control confounding factors and account for repeated measures. We explored how music genres (Jazz, Hip Hop, Rock, Pop) and audio attributes (tempo, valence, instrumentalness) correlate with stress responses across mental health conditions versus matched controls. Results revealed significant between-group differences and genre-specific effects. For example, depressed users showed 23.5% higher stress than controls 60 minutes after listening to pop music, and PTSD users showed a 33% increase after 30 minutes. Anxiety and bipolar groups did not exhibit significant stress changes with pop. We also found low-valence ("sad") songs led to delayed stress increases (e.g., 14% in depression, 25% in bipolar), while high-valence ("happy") music caused no significant elevation. Finally, using these insights, we demonstrate a proof-of-concept recommendation model that more effectively ranks stress-reducing songs (MRR = 0.354 vs. 0.18 baseline). These findings show that the association between music listening and stress varies by genre and mental health status, highlighting the potential of data-driven music interventions for emotional well-being.

CCS Concepts: • **Social and professional topics** → *User characteristics.*

Additional Key Words and Phrases: Mental Health, Music, Social Media, Stress

ACM Reference Format:
Parya Abadeh, Zeinab Noorian, Fattane Zarrinkalam, and Amira Ghenai. 2026. Music Listening, Mental Health, and Stress: A Computational Framework for Personalized Analysis and Recommendation. 1, 1 (February 2026), 35 pages. https://doi.org/10.1145/nnnnnnn.nnnnnnn

1 Introduction

In recent years, social media has become an integral part of our daily life, serving as a platform where individuals express their thoughts, emotions, and personal experiences. Many users, particularly those facing psychological disorders, implicitly or explicitly discuss their mental health struggles on these platforms as a form of emotional release and self-expression [5, 75]. This trend has spurred a growing interest in leveraging digital traces of social media to assess and understand the mental states of users [19, 37, 105]. Previous research has primarily focused on the analysis of textual content, images, and online behavioral patterns to identify linguistic and psychological markers associated with various mental health disorders [80, 97].

Authors' Contact Information: Parya Abadeh, University of Guelph, Guelph, Canada, pabadeh@uoguelph.ca; Zeinab Noorian, Toronto Metropolitan University, Toronto, Canada, znoorian@torontomu.ca; Fattane Zarrinkalam, University of Guelph, Guelph, Canada, fzarrink@uoguelph.ca; Amira Ghenai, Toronto Metropolitan University, Toronto, Canada, aghenai@torontomu.ca.

ACM XXXX-XXXX/2026/2-ART
https://doi.org/10.1145/nnnnnnn.nnnnnnn

In parallel, music has gained considerable attention as a therapeutic tool for stress management and emotional regulation. Music listening and music therapy have been shown to alleviate stress and anxiety [27, 38]. Researchers have explored multiple dimensions of the relationship between music and mental health, including the influence of different genres of music on emotional responses [66], the connections between musical preferences and psychological conditions [3], and the role of music participation in the management of stress among people with various mental health challenges [62, 74]. Collectively, these findings indicate that music is a promising medium for promoting mental well-being and reducing stress.

However, despite these encouraging insights, existing studies often suffer from methodological limitations. First, most prior research has been conducted in controlled clinical settings or through traditional self-reported surveys, which typically involve small sample sizes and may not capture the complexities of real-world behavior [66]. Second, many studies focus exclusively on healthy individuals or a specific clinical group, often overlooking broader interactions between music listening, stress, and diverse mental health conditions [38, 66]. For example, the work by [66] examines music effects only in a clinical context, leaving open questions about broader populations. Furthermore, there is limited research that integrates multiple dimensions – such as social media activity, behavioral patterns, and music preferences – into a comprehensive framework to assess how music engagement associates with stress levels in naturalistic settings. This gap in the literature underscores the need for a more scalable and data-driven approach to understanding the role of music in stress modulation.

In response to these gaps, our research employs a methodological approach combining propensity score matching (PSM) with generalized linear mixed-effects models (GLMMs) to rigorously analyze the association between music listening, stress, and mental health using large-scale social media data. PSM effectively reduces observational bias by creating matched control groups, whereas GLMMs accommodate individual-level variability and repeated observations (e.g., multiple listening sessions) to offer precise insights into how music attributes relate to stress. This integrated $PSM + GLMM$ framework allows for a more robust analysis of the association of music listening on stress under real-world conditions, addressing the shortcomings of earlier studies.

1.1 Research Objective

To comprehensively address the outlined literature gaps, this study pursues the following explicit research objectives:

- To systematically investigate how music listening behaviors on social media correlate with stress responses among individuals with various mental health conditions (depression, anxiety, PTSD, and bipolar disorder) compared to matched controls without diagnosed conditions.
- To explore variations in stress responses across different music genres and audio features (e.g., tempo, valence, instrumentalness) among individuals with mental health conditions and their matched controls, assessing their associations at different intervals following music listening events.
- To evaluate the applicability and utility of identified music–stress associations in designing personalized music recommendation systems aimed at stress management and emotional support, demonstrating real-world implications of our research findings.

These research objectives are operationalized through the following research questions: **RQ1.** How is music listening associated with stress levels among individuals with mental health conditions compared to their matched control group without disorders?

RQ2. How does the relationship between music listening and stress vary across different genres and time intervals after listening (e.g., in the 30- or 60-minute period post-listening) among individuals with mental health conditions compared to matched controls?

RQ3. How do specific audio features (tempo, valence, instrumentalness) of music relate to stress levels among users with various mental health conditions compared to their matched controls?

RQ4. How can the associations identified in RQ2 and RQ3 be leveraged in a personalized recommendation system to support users' emotional well-being?

In summary, this paper makes the following key contributions:

(1) We employ a novel methodological framework that combines Propensity Score Matching with Mixed-Effects Modeling to enable association analysis of observational social media data. This approach is the first to integrate PSM and GLMMs in the context of music and mental health, allowing us to rigorously examine the relationship between music listening and stress levels while controlling for confounders, and accounting for individual-level variability and repeated measures.

(2) Using the above framework, we provide an extensive empirical analysis of how music listening correlates with stress. We show how this relationship differs between users with mental health disorders and matched controls, and how it varies across music genres and over time after listening. Our findings offer novel insights into genre-specific and temporal effects of music on stress in real-world settings, addressing gaps left by prior small-scale or single-group studies.

(3) We investigate the association of specific musical features (such as tempo, valence, and instrumentalness) on stress levels across different mental health groups. This analysis reveals which audio characteristics of music are most associated with stress reduction or elevation, shedding light on why certain music might be therapeutic for some individuals.

(4) We provide a proof-of-concept illustrating how relationships identified between specific music attributes and stress levels can inform the development of personalized music recommendation systems aimed at supporting mental health. Specifically, we demonstrate how recommendation frameworks could leverage these insights to prioritize music selections associated with stress reduction, highlighting the potential for music-driven interventions to enhance users' emotional well-being.

2 Related Work

Prior research has explored the effects of various interventions on stress, highlighting the role of different activities and therapies in managing stress levels. For instance, the study done in [73] found that practicing yoga during pregnancy is effective in reducing stress and preventing complications. Similarly, [6] examined the impact of meditation on stress reduction among college students, demonstrating its potential as a practical tool for stress management. Another significant area of research focuses on the effects of music on stress. For example, [86] confirmed that music therapy serves as an effective strategy for stress management among medical students. Additionally, [108] showed that a mindfulness-based stress reduction program combined with music therapy not only reduced stress and depressive symptoms but also improved psychological well-being in cancer patients. These findings collectively suggest that stress levels can be influenced by a range of factors, with music playing a particularly significant role.

The relationship between music and mental health has attracted increasing attention in recent years [56]. Music is widely recognized for its therapeutic potential, particularly in stress relief and emotional regulation, making it a valuable tool for individuals with mental health conditions [87]. Numerous studies have demonstrated its benefits; for instance, [93] and [11] found that music can alleviate symptoms of anxiety and depression, while [99] and [35] highlighted its role in enhancing mood and emotional well-being. Specific applications of music interventions have also been explored, such as their positive effects on psychological health and academic performance among students [52] and their ability to reduce perinatal depression [42] or alleviate anxiety in cardiac patients [104].

However, the influence of music on mental health is not universally positive. Some research suggests that, in certain contexts, music may intensify distress rather than alleviate it [70]. For individuals experiencing high levels of emotional turmoil, music might reinforce negative moods rather than provide relief. Additionally, music's ability to evoke memories—while beneficial for conditions like dementia—can be problematic for individuals with PTSD, as it may trigger unwanted recollections [8, 56]. These findings emphasize the complex and context-dependent nature of music's effects on mental health. Beyond its direct emotional and psychological impact, musical engagement—including listening habits, genre preferences, and personal associations—plays a crucial role in shaping individual responses to music. Understanding these nuances is essential for leveraging music as an effective tool for mental health support.

Music therapy also has demonstrated significant efficacy across various mental health conditions [88, 100], with notable improvements in depression symptoms [103], sleep patterns [95], and overall quality of life (QOL) [100]. It has also been effective in reducing Post-traumatic stress disorder (PTSD) symptoms among school adolescents [48] and depression levels among elderly populations [47].

However, the majority of existing mental health studies have been confined to clinical settings or depend on traditional approaches such as questionnaires and self-reports [48], frequently involving limited participant samples [47]. To overcome these constraints, researchers are increasingly leveraging social media data to explore and identify a wider spectrum of mental health conditions, including depression, anxiety, and stress [71, 105]. Social media provides a valuable repository of user-generated content, which can be analyzed through natural language processing, machine learning, and other computational methods to uncover linguistic patterns, behavioral indicators, and emotional states associated with mental health issues [105, 109].

Social media offers an additional perspective on traditional clinical methods by providing real-time, large-scale data on mental health experiences and expressions. Several studies have demonstrated this potential: [97] examined how sharing misinformation affects anxiety levels using Twitter data, while [60] developed a comprehensive framework for measuring user stress levels through social media content, incorporating a benchmark dataset with stress-oriented features. Similarly, [110] proposed a BERT-based model for rapid detection of psychological stress in postgraduate students via social media analysis. Additionally, most recent research focuses exclusively on analyzing user-generated text to identify psychological disorders like depression [91], bipolar disorder [90], and suicide risk [39, 76], often neglecting other factors such as the music users share. Different from clinical studies that investigate the link between mental health and musical engagement, a recent study [3] examined the relationship between users' self-reported mental health issues on social media and their music preferences. The study analyzed Twitter data in detail, looking at the linguistic patterns of music liked by individuals with different mental health disorders and comparing these patterns with those from a control group without such conditions. Furthermore, another study [1] was conducted to explore the relationship between music listening and stress levels among social media users diagnosed with anxiety, depression, and Post-traumatic stress disorder (PTSD), specifically examining how demographic factors like age and education level influence this relationship, and showed that older age groups experienced higher stress levels, particularly for anxiety and depression, while lower education levels were associated with lower stress scores across all investigated mental health conditions.

While prior research has provided valuable insights into the effects of music on stress and mental health, several critical gaps remain:

- Limited Focus on Specific Mental Health Conditions: Most studies focus on general populations or specific clinical groups, often overlooking the nuanced ways in which individuals with different mental health conditions—such as depression, anxiety, bipolar disorder, and PTSD—respond to music [8, 56]. Our study addresses this gap by examining how music listening, genres, and features are associated with stress levels across distinct mental disorder groups, providing tailored insights for each condition.
- Reliance on Controlled Settings and Small Samples: Most of the existing work relies on controlled environments, self-reported data, or small participant samples Bradt et al. [11], Iyendo et al. [48], Toprak et al. [93],

limiting the generalizability of findings to real-world scenarios. We overcome this limitation by leveraging a large-scale dataset of Twitter users, ensuring our findings are representative of real-world behaviors.
- Neglect of Music Features: While some studies have examined the impact of music on stress [86, 108], they often ignore key features of music, such as genres and attributes like tempo and valence. Our study addresses this gap by analyzing the role of music genres and features in stress reduction, uncovering how specific musical attributes influence stress levels across different mental health conditions.
- Limited Consideration of Individual Variability and Repeated Measures: Many prior studies, including those employing traditional randomized control trials (RCTs), do not adequately capture intra-individual variability or repeated observations of user behavior [88, 93, 108]. Such methodologies overlook how an individual's responses to music might vary across multiple listening sessions. To overcome this limitation, we utilize Generalized Linear Mixed Models (GLMMs), which explicitly account for both individual-level differences and repeated measures. This approach allows for a deeper, more personalized analysis of the relationship between music exposure and stress.
- Insufficient Clinical Insights in Music Recommendation Systems: While current emotion-aware music recommender systems [63, 102] effectively consider users' emotional states, they lack explicit integration of clinical knowledge regarding how specific musical features influence stress responses among individuals with varying mental health conditions. Traditional recommendation algorithms typically do not differentiate these nuanced interactions. Our study highlights the importance of integrating clinical insights by providing a proof-of-concept music recommendation system that leverages the relationships identified through our proposed methodological framework. This approach demonstrates how specific music characteristics can be utilized to create personalized recommendations tailored to reduce stress levels across varying mental health conditions.

By addressing these limitations, our study bridges the gap between controlled studies and real-world behaviors, offering practical insights for designing music-based interventions tailored to specific mental health needs.

3 Data Collection

This section describes the construction and characteristics of the dataset used in our research. The dataset consists of six distinct mental health conditions, categorized according to the Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition (DSM-5) [28]: Depression, Anxiety, PTSD, Bipolar Disorder, Borderline Personality Disorder, and Panic Disorder. Additionally, it includes individuals who do not have a diagnosed mental health condition. A defining feature of this dataset is that all users, irrespective of their mental health status, have documented instances of music listening in their tweets. This unique characteristic allows an in-depth investigation of the intricate relationship between music engagement and mental well-being.

3.1 Dataset

This study utilizes a publicly available dataset [2], which includes Twitter users categorized into disorder and non-disorder groups based on self-reported mental health diagnoses. Disorder users were identified by analyzing tweets for explicit statements; high-precision pattern-matching techniques were applied to locate phrases such as "I was diagnosed with X Y," where "X" denotes a specific disorder, and "Y" indicates a date specifier (e.g., today, yesterday, this week). Data collection spanned from January 2020 to February 2022. A set of 17,370 tweets was independently annotated by three students, yielding substantial inter-rater agreement (Fleiss' K = 0.9). After removing users whose annotations did not match, for each identified user, up to 3,200 of their most recent tweets were extracted, resulting in a corpus containing 25,747,205 tweets, from 15,651 users.

For the non-disorder group, Twitter users were randomly sampled from the same period. Each sampled timeline was reviewed using a predefined lexicon to ensure that no mental health diagnoses were mentioned. Following

Table 1. Summary of the control and disorder datasets

User Group	# of users	# of tweets			# of music sessoions			online listening platforms		
		per user	min::max::std	# of tweets	per user	min::max::std	total	Spotify	Apple Music	SoundCloud
control	4483	8476.787865	1::3,250::1,198.73	38,001,440	17.71090787	1::1,026::46.47	79,398	53,979	21,545	3,874
anxiety	583	9154.024014	1::3251::1,154.28	5,336,796	13.48370497	1::886::41.55	7,861	5,748	1,637	476
depression	1270	8763.379528	1::3,251::1,145.1158	11,129,492	13.34566929	1::832::34.22	16,949	13,409	2,760	780
bipolar	277	8147.913357	1::3,252,::1,115.008	2,256,972	17.79422383	1::351::34.59	4,929	3,737	991	200
PTSD	500	11783.292	1::3,352::1,150.8403	5,891,646	16.112	1::777::51.57	8,056	5,680	1,933	443

this procedure, timelines comprising 38,001,440 tweets from 158,586 non-disordered users were obtained. After the disorder and non-disorder groups were identified, every timeline was scanned for links to popular music platforms—Spotify, SoundCloud, and Apple Music—to ensure that each user had at least one music-listening record. This screening removed 163,973 users, leaving 10,264 users across the two groups. Their combined timelines contained 20,910,917 tweets, of which 152,011 were music-related.

3.1.1 Music Genre Extraction. To extend the original dataset and enhance its analytical depth, we incorporated new measures that capture the relationship between music genres, their characteristics, and stress levels. Specifically, we expanded the dataset by integrating music metadata from multiple sources, primarily using APIs from Spotify, Apple Music, and SoundCloud. However, since some tracks were unavailable on these platforms, we supplemented our dataset with information from OneMusicAPI[1], which aggregates metadata from various music databases, covering approximately 8 million albums and 4 million artists. To further improve genre classification accuracy, we introduced a cross-referencing step using Last.fm's API[2], ensuring that only consistently identified genres across both platforms were retained. By extending the dataset in this manner, we significantly enhanced its completeness and reliability, enabling a more comprehensive analysis of music engagement and its potential impact on stress levels.

Refinement of Genre Classification: To further enhance the dataset, we modified and standardized the genre classification process to ensure that each track was assigned a single, distinct genre. The raw genre labels obtained from the APIs often included multiple genre associations (e.g., Hip Hop and R&B for the same track). To address this, we grouped subgenres under broader parent genres using external resources such as Chosic's Music Genre List [3]. For example, K-pop and Indie Pop were categorized under the parent genre Pop.

Building on this, we refined the classifications further by incorporating established taxonomies from prior research [4], mapping subgenres to their corresponding parent categories—for instance, R&B was consolidated under Hip Hop. Tracks associated with multiple conflicting genres that could not be assigned a single category were excluded from the dataset, leading to the removal of 16,382 tracks with multiple genres and 53,468 tracks with missing genre information.

Additionally, some users were excluded from the analysis due to null genre associations or the presence of multiple genres within their listened tracks, ensuring consistency in the dataset. Table 2 illustrates the refinement process applied to genre classification. The overall distribution of users and the volume of music-related content in the dataset is detailed in Table 1.

Our study focused on the four most prevalent genres in the dataset: Jazz, Hip Hop, Rock, Pop. These genres were selected based on their dominance in the dataset. The distribution of music tracks across these genres is detailed in Table 3. By concentrating on these genres, we aimed to provide a clearer and more focused analysis

[1]http://www.onemusicapi.com/
[2]https://www.last.fm/api
[3]https://www.chosic.com/list-of-music-genres/

Table 2. Example of Genres Refinement By Taxonomy [4] and By Website (Chosic's Music Genre List)

Original Genre	Refinement By Website (1)	Refinement By Taxonomy (2)
K Pop, German Pop	Pop	Pop
Contemporary R&B, Soul	R&B	Hip Hop
Metal, Rock	Metal	Rock
Blue, Jazz, Easy Listening	Jazz, Easy Listening	Jazz
Metal, doomgazw	Metal	Rock
New Age, Sound	New Age	New Age
Folk/Acoustic	Rock	Rock

of the relationship between music characteristics and stress levels, while ensuring the dataset's coherence and interpretability.

3.1.2 Music Additional Features. We further extracted additional music features encompassing a wide range of audio characteristics, including tempo, speechiness, energy, danceability, time signature, mode, instrumentalness, liveness, duration, valence, acousticness, and loudness. Detailed descriptions of these features can be found on the Spotify website[4]. For tracks identified on Apple Music or SoundCloud, we retrieved corresponding Spotify links to ensure consistent extraction of these audio features across platforms.

To refine our analysis and avoid redundancy, we calculated the pairwise correlation between these music features using Pearson correlation. Features exhibiting a correlation coefficient greater than 0.5 or less than -0.5 were excluded to reduce multicollinearity and enhance the interpretability of our findings. This process resulted in the selection of three key features: valence (the emotional positivity of a song), tempo (the speed or pace of a track), and instrumentalness (the extent to which a track is instrumental rather than vocal).

3.2 Stress Level Assessment

The emotional and linguistic patterns in social media discourse have been widely used as indicators of stress and anxiety levels [82, 97]. To quantify stress in users' Twitter posts, we employed TensiStrength [92], a validated lexicon-based tool designed to detect stress and relaxation in textual data. TensiStrength assigns stress scores ranging from -1 (no stress) to -5 (extreme stress), leveraging predefined stress-related lexicons while integrating spelling correction, booster words (e.g., very stressed), negation handling (e.g., not happy), punctuation emphasis (e.g., stressed!!!), elongated words (e.g., sooo tired), and emoticons to refine classification accuracy. The tool has demonstrated robust performance in prior studies on stress detection in social media [41, 89].

Prior to applying TensiStrength, tweets were preprocessed and cleaned following best practices outlined in [92], with the analysis restricted to English-language posts. To facilitate interpretability, the stress scores were rescaled to a 0–4 range, where 0 represents minimal stress and 4 indicates maximum stress.

To examine stress levels in relation to music engagement, we computed stress scores all users based on their social media activity following a music-related post. Specifically, we identified instances where users engaged with music and extracted all tweets posted within the subsequent 60 minutes. The final stress score was derived by averaging the stress levels of tweets within two distinct time windows (30 minutes and 60 minutes post-music engagement) for both user groups. The analytical framework underlying this approach is detailed in Section 4.

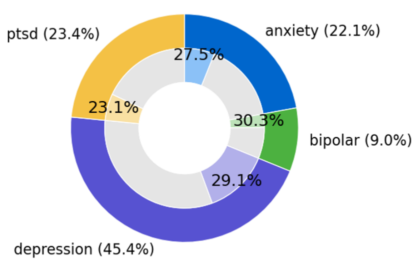
[4]https://developer.spotify.com/documentation/web-api/reference/get-audio-features

Table 3. Distribution of Music Tracks Across Top 4 Genres

Genre Category	Number of Music
Hip Hop	31,592
Pop	28,542
Rock	17,748
Jazz	1,920

Table 4. The distribution of users across mental health groups before and after Propensity Score Matching

Disorder/ Control Group		Before matching	After matching
depression/control	depression	1270	1085
	control	2183	2098
anxiety/control	anxiety	583	502
	control	710	689
PTSD/control	ptsd	500	395
	control	520	498
Bipolar/control	bipolar	532	514
	control	277	239



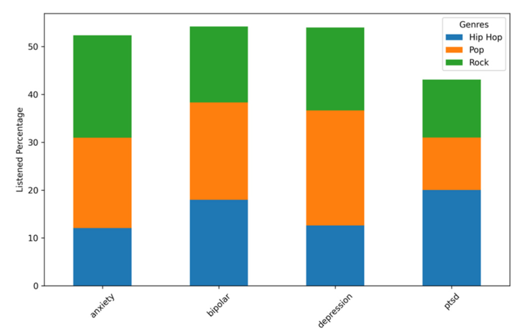


Fig. 1. Population Distribution and Music Sharing by Disorder (left side) (The percentages in the inner ring show what percentage of users with each disorder shared music.) - Distribution of three most frequently listened-to genres (right side)

3.3 Descriptive Statistics on Mental Health Groups

Figure 1 (left side) provides a visual representation of the distribution of mental health conditions and the prevalence of music-sharing behavior across disorder groups. The outer ring of the chart depicts the overall prevalence of each disorder, with depression emerging as the most commonly reported condition. The inner ring illustrates the proportion of users within each disorder category who actively share music on social media. While depression is the most prevalent disorder, the data reveal that individuals with bipolar disorder exhibit the highest rate of music-sharing behavior, suggesting potential variations in music engagement patterns across mental health conditions.

3.3.1 Music Genre Preferences by Mental Health Group. To examine the relationship between mental health conditions and music genre preferences, Figure 1 (right side) presents the distribution of pop, rock, and hip hop listening behaviors across disorder groups. The data indicate that hip hop emerges as the most frequently listened-to genre among individuals with PTSD, accounting for 20.02% of their total listening time. This proportion is notably higher compared to other disorder groups, where hip hop constitutes 12.04% for Anxiety, 17.99% for Bipolar, and 12.59% for Depression. These findings suggest potential genre-specific preferences associated with different mental health conditions.

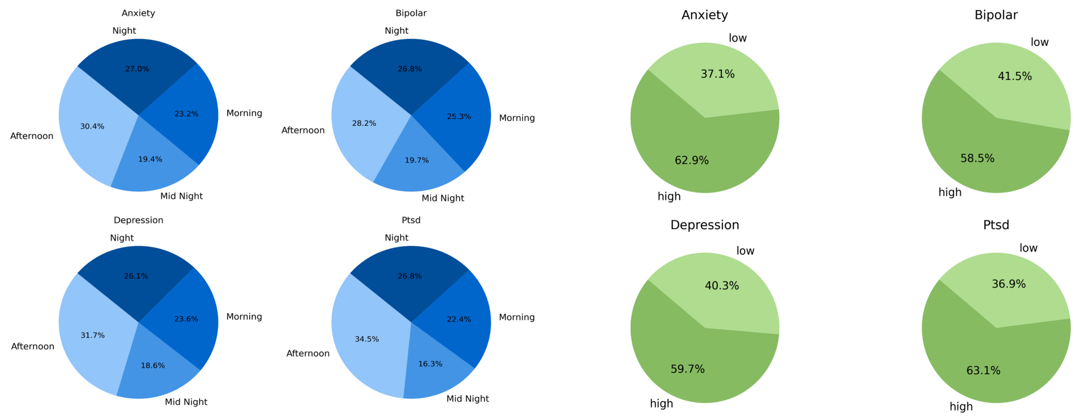


Fig. 2. Distribution of Time of Day for Shared Music (left side) - Distribution of Music Tempo Category by Disorder (right side)

Pop and rock emerge as the dominant music genres across most mental health conditions, with notable variations in preference distributions. Among individuals with Depression, pop accounts for the highest proportion of listening time at 24.05%, followed by Rock at 17.36%. Similarly, in the Bipolar disorder group, pop leads with 20.34%, while rock (15.88%) and hip hop (17.99%) also maintain substantial shares. In contrast, individuals with Anxiety exhibit a stronger preference for rock (21.38%), with pop closely following at 18.93%.

Overall, while pop and rock maintain widespread popularity across disorders, hip hop stands out as a particularly favored genre among individuals with PTSD, reinforcing the distinct listening patterns observed in different mental health groups.

3.3.2 *Temporal and Acoustic Characteristics of Music Sharing.*

Music Sharing Patterns Across Time of Day: Figure 2 (left side) presents an analysis of music-sharing behaviors across different times of the day, categorized into four periods: Midnight (00:00–06:00), Morning (06:00–12:00), Afternoon (12:00–18:00), and Night (18:00–24:00). To ensure consistency, all timestamps were standardized based on the database's default timezone.

Across all mental health groups, the Afternoon emerges as the most prominent period for music sharing, with the highest activity observed among individuals with PTSD (34.5%), followed by those with Depression (31.7%), Anxiety (30.4%), and Bipolar disorder (28.2%). Night-time sharing is relatively stable across groups, ranging from 26.1% (Depression) to 27.0% (Anxiety). In contrast, Morning engagement varies more distinctly, with Bipolar disorder exhibiting the highest share (25.3%) and PTSD the lowest (22.4%).

The Midnight period shows the lowest overall activity, with individuals with PTSD sharing the least music (16.3%), while Anxiety (19.4%) and Bipolar disorder (19.7%) exhibit slightly higher engagement levels. Depression falls in between, with 18.6% of music shared during this timeframe.

Music Tempo Preferences: Figure2 (right side) further examines tempo-based preferences, categorizing tracks into low tempo (below 70 BPM) and high tempo (120 BPM and above). Across all groups, there is a clear preference for high-tempo music, with PTSD (63.1%) and Anxiety (62.9%) displaying the highest inclination toward faster-paced music. Similarly, individuals with Depression (59.7%) and Bipolar disorder (58.5%) also favor high-tempo

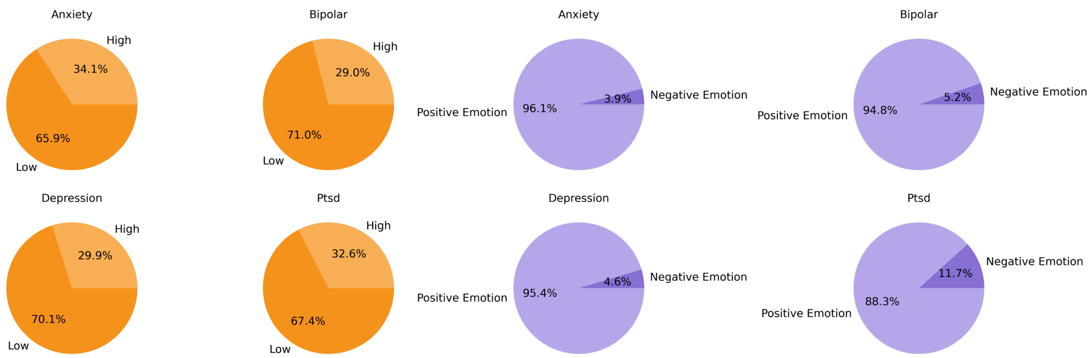


Fig. 3. Distribution of Music by Instrumentalness Category (left side) - Distribution of Music by Valence Category in Each Disorder Group (right side)

tracks, though to a slightly lesser extent. Conversely, low-tempo music is more prevalent among individuals with Bipolar disorder (41.5%) and Depression (40.3%), whereas Anxiety (37.1%) and PTSD (36.9%) exhibit lower engagement with slower music. These results suggest that individuals with PTSD and Anxiety tend to favor more energetic and rhythmic tracks, whereas those with Bipolar disorder and Depression show a relatively more balanced distribution between low and high-tempo selections.

Music Instrumental Preferences: Figure 3 (left side) illustrates instrumentalness preferences, categorizing tracks into low instrumentalness (below 0.5) and high instrumentalness (0.5 and above). Across all groups, there is a clear preference for low-instrumentalness music, meaning tracks with prominent vocals. This preference is most pronounced in individuals with Bipolar disorder (71.0%) and Depression (70.1%), followed closely by those with PTSD (67.4%) and Anxiety (65.9%). Conversely, high-instrumentalness music, characterized by more instrumental or ambient compositions, is selected by a smaller proportion of users across all mental health conditions, ranging from 29.0% to 34.1%. These results indicate that individuals with mental health conditions generally prefer vocal-rich music over instrumental tracks.

Music Valence Preferences: Figure 3 (right side) also illustrates music valence preferences, where valence measures the emotional tone of a track (higher values indicate more positive or uplifting music). The results demonstrate a strong inclination toward positive-valence music across all groups, with 94–96% of selections falling in the positive range (valence > 0.5). In contrast, low-valence music (valence < 0.5), indicative of a somber or more emotionally intense tracks, comprises only 3–6% of selections. This analysis suggests that individuals with mental health conditions generally prefer uplifting and positive-toned music, with limited engagement in negative-valence tracks.

4 Research Methodology

To investigate the relationship between music listening and stress across mental health conditions, we first construct two primary user groups: a treatment group, comprising users who self-reported one of four mental health conditions (depression, anxiety, PTSD, or bipolar disorder), and a control group, consisting of users without any such self-reported diagnoses. All users in both groups had verifiable music-listening activity on their social media timelines, allowing for a consistent basis of comparison.

4.1 An Overview Analytical Framework

This study introduces a novel analytical framework designed to examine the relationship between mental health conditions and stress levels, as reflected in social media posts, in response to music engagement—addressing RQ1, RQ2, and RQ3. The framework integrates Propensity Score Matching (PSM) to construct comparable user groups by balancing covariates, followed by Generalized Linear Mixed Models (GLMMs) with a zero-inflated component to model stress variations. Figure 4 provides an overview of the methodological pipeline and process.

4.2 Propensity Score Matching for Bias Reduction

4.2.1 Design and Rationale. In this study, Propensity Score Matching (PSM) is employed as a statistical approach to improve comparability between groups by reducing the impact of pre-existing differences. Unlike its traditional use in causal inference [43], our application of PSM is intended to enhance the validity of associational analysis by ensuring that treatment and control groups share similar baseline characteristics, allowing for meaningful statistical comparisons. We adopt a one-to-many nearest neighbor matching approach to construct a well-balanced control group, ensuring that users without mental disorders closely resemble those with mental disorders across key covariates, including social media activity, linguistic patterns, and engagement metrics. This methodology enables a more reliable investigation of associations between mental health conditions, music listening, and stress levels using Generalized Linear Mixed Models (GLMMs).

PSM operates by balancing covariates between groups, reducing selection bias and mitigating the influence of confounding variables. By creating matched groups with similar characteristics, PSM facilitates valid comparisons in observational settings, an approach widely adopted in computational social science research [36, 65, 83].

Building on these strengths, our implementation of one-to-many nearest neighbor PSM–rather than alternatives such as Inverse Probability Weighting (IPW) [7, 29] or stratified matching [82, 97]–ensures a well-defined control group while preserving individual-level granularity and statistical power. Unlike IPW, which assigns weights rather than directly matching units, nearest neighbor PSM optimizes comparability by selecting the most similar control users based on propensity scores. This approach enables a more precise analysis of stress modulation in response to music genres across mental health conditions.

By reducing confounding bias and ensuring balanced comparison, PSM enhances the reliability of our mixed-effects model analysis, allowing us to investigate the relationship between mental disorders, stress levels, and music genres with greater confidence. Our approach aligns with recent advances in computational social science, where PSM is increasingly leveraged to improve group comparability and mitigate selection bias in large-scale social media studie[50, 51].

4.2.2 Assignment of placebo dates for control users. In this study, we define the *treatment date* as the time when a user in the treatment group explicitly shared a self-reported post about their mental disorder. However, since users in the control group did not disclose any mental health diagnoses in their timelines, a *placebo date* was assigned to each control user to ensure that their timeline remained comparable to that of the treatment group, mitigating potential temporal confounds that could arise due to differences in the timing of data collection.

To construct a valid set of placebo dates, motivated by research of [97], we non-parametrically simulated placebo dates from the pool distribution of treatment dates. This approach ensures that placebo dates follow the same temporal pattern as treatment dates, preventing potential biases caused by seasonal or temporal trends in user behavior. Once generated, placebo dates were assigned to control users based on criteria ensuring sufficient social media activity and music-listening behavior while maintaining temporal alignment with the treatment group. After assigning placebo dates, we evaluated their comparability to treatment dates using the Kolmogorov-Smirnov (KS) test [51]. The KS test quantifies the maximum difference between the empirical cumulative distribution functions of the treatment and placebo dates. The resulting KS statistic of 0.019 indicates that the placebo dates

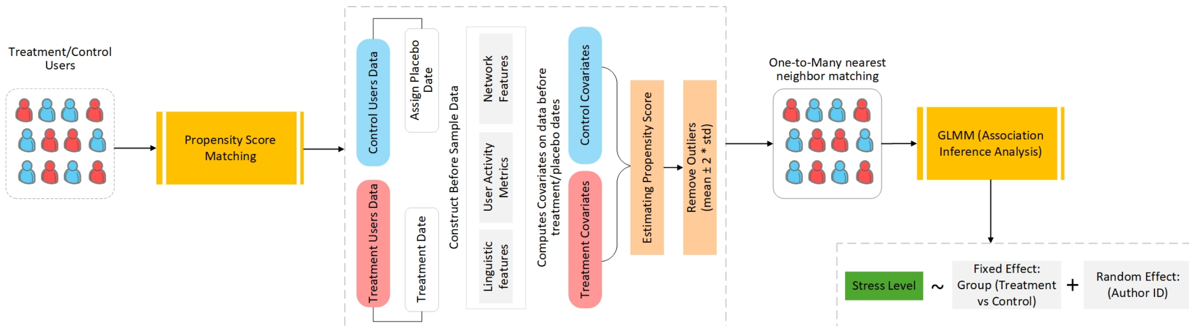


Fig. 4. The workflow of the proposed methodological framework

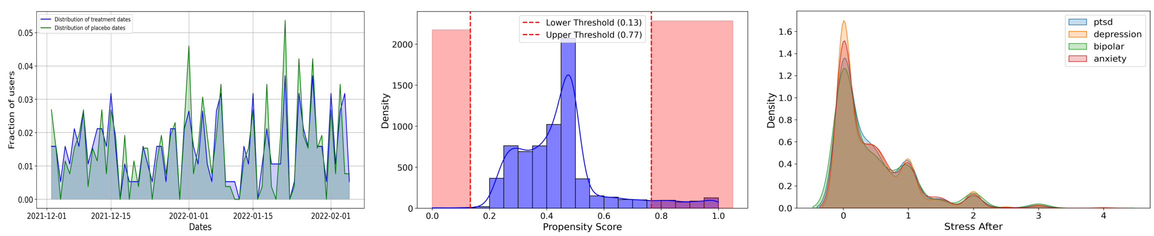


Fig. 5. Distribution of treatment and placebo dates for all disorders. (left side) - Propensity score distribution for anxiety group (Shaded region represents those dropped in our analysis (center side) - Kernel Density Estimation (KDE) plot of stress values across different disorders. The density represents the relative frequency of stress levels within each disorder group, scaled such that the area under each curve sums to 1. A high number of 0 values causes peaks near zero, highlighting its prevalence across disorders. (right side).

effectively mimic the temporal pattern of treatment dates. This alignment is further illustrated in Figure 5 (left side), which presents the overlapping distributions of treatment and placebo dates across all mental disorders.

4.2.3 Matching Covariates. The covariates used in this study are carefully selected to capture the psycholinguistic, behavioral, and social dimensions of user activity on social media. These features are inspired by prior research in mental health studies on social media [97]. To ensure a robust and multidimensional approach to user matching, we categorize our covariates into three main groups: *linguistic features*, *user activity metrics*, and *network engagement indicators*. These variables help reduce potential confounding in our analysis and enhance the validity of our comparisons.

Linguistic feature: We utilize the Linguistic Inquiry and Word Count (LIWC) classification framework [77] for analyzing psychological and linguistic patterns in text. LIWC has been widely used in prior research for the detection of mental health markers, cognitive patterns, and emotional states in social media data [94]. From the 72 LIWC categories, we refine our selection to 12 high-level categories based on their relevance to capturing underlying psycholinguistics and nuances of user behavior. These LIWC features include: *linguistic dimensions*, which capture structural aspects of text, such as word complexity and grammatical diversity; *psychological processes*, reflecting language associated with anxiety, sadness, and self-awareness; *social processes*, which include pronoun usage and references to interpersonal interactions; *cognitive processes*, capturing words related to causation, insight, and problem-solving; *perceptual processes*, encompassing references to sensory experiences such as vision, hearing, and touch; *biological processes*, including language related to health, body, and physiological states; *drives*, which signal motivation and behavior related to power, affiliation, and achievement; *time orientations*, capturing temporal focus through references to the past, present, or future; *relativity*, which includes spatial, temporal, and motion-related expressions; *personal concerns*, reflecting language about work, money, and relationships; *informal language*, encompassing slang, abbreviations, and casual expressions common in social media discourse and *other grammar*, including variations in syntax and informal writing patterns.

User activity metrics: The user activity metrics include: *author tweet count*, which measures the total number of tweets posted by a user; *author music count*, capturing the frequency of tweets containing music-related content; *average retweet*, representing how often a user's tweets are shared by others; *average reply*, indicating the extent of direct interactions through responses; *average like*, measuring the frequency of positive engagement with a user's posts; *average quote*, representing how often a user's content is reshared with added commentary; *average weekly tweets*, capturing a user's posting frequency on a weekly basis; and *average monthly tweets*, providing insights into monthly posting patterns of users.

Network engagement features: The network engagement features include: *followers count*, which represents the number of users following an account; *followings count*, measuring the number of accounts a user follows; and *created date*, marking the account's inception to assess a user's longevity on the platform.

4.2.4 Propensity Score Analysis. Propensity Score Matching (PSM) was employed to adjust for potential confounding variables and estimate the correlation between music listening and stress. To determine the propensity scores, we develop a logistic regression model predicting a user's likelihood of belonging to the treatment group based on their characteristics, including their linguistic features, network and engagement metrics. We eliminate outliers with propensity scores outside ±2 standard deviations from the mean [82]. Figure 5 (center side) shows the distribution of propensity scores in Anxiety group. Using the computed propensity scores, we applied a one-to-many nearest-neighbor matching approach without replacement to pair treatment and control users. This approach enhances the comparability of treatment and control groups, thereby improving the validity of the subsequent association inference analysis. Table 4 presents the distribution of treatment and control groups for each mental disorder following nearest-neighbor PSM.

4.2.5 Quality of Matching. To assess the quality of matching between the treatment and control groups, we utilized the Standardized Mean Difference (SMD) as the primary evaluation metric. SMD is a widely recognized measure for examining covariate balance in observational studies, as it quantifies the difference in means of a covariate between groups while accounting for pooled standard deviation, ensuring consistency across varying group sizes and distributions [36, 82]. In this study, we follow the commonly accepted threshold of SMD values below 0.25, as recommended in prior research [36, 82, 97], indicating sufficient balance between treatment and control group and minimizing potential confoundinf effects. After implementing propensity score matching, we computed SMD values for all covariates included in the analysis, encompassing linguistic features, activity metrics, and network characteristics. The results demonstrate that all covariates meet the specified threshold, confirming confirming a high degree of balance between the different groups of treatment and control users.

4.3 Statistical Analysis

4.3.1 Model Selection and Rationale. The primary objective of this study is to examine the relationship between listening to various music genres and stress levels among individuals diagnosed with different mental disorders, including depression, anxiety, PTSD, and bipolar disorder. Given the nature of our dataset, where each user has multiple sessions of music listening, it is imperative to employ a statistical model that can account for individual differences and the repeated measures present in the data.

To address these requirements, we employed a Generalized Linear Mixed-Effects Model (GLMM) [10], which is well-suited for modeling continuous outcomes like stress levels while accounting for intra-individual correlations due to repeated measurements.

Given the significant number of zero stress scores in our dataset, as depicted in Figure 5 (right side), we utilized a Zero-Inflated GLMM (ZIGLMM) [13]. This model is specifically designed to handle datasets with excessive zeros by decomposing the data into two underlying processes: one that generates zero values and another that generates positive stress scores. By capturing these distinct processes, the ZIGLMM improves the accuracy of parameter estimation and enhances overall model fit [14].

4.3.2 Model Formalization. To investigate RQ1, RQ2, and RQ3, we employ GLMM model with a zero-inflated component. Given the *skewed distribution* of stress scores and the presence of *excess zero values*, we adopt a *zero-inflated gamma model (ZIGLMM)* to appropriately model the data.

The model is estimated using maximum likelihood estimation (MLE) in the glmmTMB package [13], employing a Gamma family with a log link function for the conditional model and a logit link function for the zero-inflation component. Since each research question focuses on different aspects of music listening and stress, we construct separate models, incorporating distinct subsets of user timelines and varying numbers of observations.

For **RQ1**, we compare stress levels between users with mental disorders (e.g., Anxiety, PTSD, Bipolar, Depression) and their PSM-matched control group who also engage in music listening. This comparison is made without considering specific music genres. Regarding **RQ2**, we analyze stress levels across specific music genres (e.g., pop, jazz, hip-hop, rock). For each genre, we restrict the analysis to user sessions where individuals have listened to that genre and compare stress levels between mental disorder groups and their matched control users, considering stress score calculated within 30-minute and 60-minute post-listening intervals. Finally, for **RQ3**, we investigate the relationship between specific music features (e.g., tempo, valence) and stress levels. The analysis is restricted to user sessions where the selected music feature exhibits either high or low values, following classifications from prior research [69]. Only user timelines containing music with the targeted feature values are considered, ensuring a focused examination of how distinct music attributes associate with stress levels.

Below, we present the formal model specification in detail. **Conditional Model (Gamma Distribution for Stress Scores):** The stress level $Y_{i,t,m}$ of user i at time t for genre m follows a Gamma distribution:

$$Y_{i,t,m} \sim \text{Gamma}(\mu_{i,t,m}, \theta) \tag{1}$$

where the expected stress level $\mu_{i,t,m}$ is modeled as:

$$\log(\mu_{i,t,m}) = \beta_0 + \beta_1 G_i + u_i \tag{2}$$

where:

- G_i represents the user group (mental disorder vs. PSM-bases matched control), included as a *fixed effect.*
- β_0 is the intercept, capturing the baseline stress level.
- β_1 quantifies the effect of the user group on stress level.
- $u_i \sim \mathcal{N}(0, \sigma_u^2)$ is the *random effect* capturing individual-level variability.

Zero-Inflation Model (Logistic Regression for Probability of Zero Stress): To account for the excess number of zero-stress values, we model the probability of a zero-stress outcome using a logistic regression:

$$P(Y_{i,t,m} = 0) = \frac{1}{1 + e^{(\gamma_0 + \gamma_1 G_i + v_i)}} \tag{3}$$

where:

- γ_0 is the intercept.
- γ_1 represents the effect of user group on the probability of reporting zero stress.
- $v_i \sim \mathcal{N}(0, \sigma_v^2)$ is the *random effect*, accounting for variability in reporting zero stress.

Model fit Evaluation: To assess the validity of our model, we examined Pearson residuals vs. fitted values to evaluate potential model misfit and check for patterns indicating overdispersion [16]. The residuals appeared randomly scattered around zero, suggesting that the model effectively captured variability in stress scores without systematic bias. Additionally, we assessed zero-inflation diagnostics to ensure that the model appropriately accounted for excess zeros in the data. These validation checks confirm that our statistical model is well-specified and accurately represents the association between music listening and stress.

5 Results and Analysis

In this section, we address the three research questions (**RQ1-RQ3**) by presenting the key findings and highlighting results that achieved statistical significance ($p < 0.05$) or marginal significance ($0.05 \leq p < 0.10$). We report both the main effects on stress levels from the conditional model and, where relevant, the complementary findings from the zero-inflated component of the ZIGLMM (which captures differences in the likelihood of zero stress reports). Table 5 summarizes the number of observations and users across all subgroups and experimental settings, along with the variance and standard deviation of the random effect in both the conditional and zero-inflated components. These metrics reflect the degree of user-level variability within each mental health condition and support the need for mixed-effects modeling to account for nested and heterogeneous user behaviors.

5.1 RQ1: Music Listening and Stress Level Associations Across Mental Health Disorders

We compared the stress responses of users with depression, anxiety, PTSD, and bipolar disorder to those of propensity score-matched control users. As indicated in Table 6, the analysis revealed notable differences in how music listening is associated with stress across these mental health conditions. Users with **depression** showed a significantly higher stress level after listening to music compared to controls: about a 17% increase at 30 minutes post-listening and still around 15% higher at 60 minutes, both ($p < 0.01$). The zero-inflated model also confirmed that depressed users were 29% less likely to report zero stress at 60 minutes post-listening ($p < 0.01$), meaning they had a significantly lower chance of being stress-free than control users at that time. Users with **PTSD** exhibited a similar trend of elevated stress (approximately 15% above controls) with ($p < 0.1$) at 30 and 60 minutes intervals. In the **bipolar** group, no immediate stress change was observed at 30 minutes, but by 60 minutes there was a hint of a stress increase (16% above controls, $p \approx 0.08$). In contrast, the **anxiety** group did not show any significant differences in stress levels compared to controls at either the 30-minute or 60-minute interval.

These results indicate that the impact of music on stress can vary greatly depending on the mental health context. Our findings for the depression group align with prior research suggesting that music listening can be a "double-edged sword" for individuals with depression. While it is shown that music can provide temporary relief, however, certain songs—especially those with negative or ambiguous emotional content— can reinforce rumination and other maladaptive emotion-regulation patterns, ultimately elevating stress [33]. In the case of PTSD, our observation of a trend toward higher stress contrasts with studies in therapeutic settings where music interventions (e.g. group music therapy) significantly reduce PTSD symptom severity [17]. This potentially implies that casual music listening—as observed in social media posts—may not yield the same stress-relieving benefits as those achieved through structured music therapy in PTSD users. Meanwhile, the slight stress increase we observed in the bipolar group is consistent with reports that individuals with bipolar disorder can experience intensified negative emotions in response to music. Choppin et al [22], for example, noted heightened emotional sensitivity to musical stimuli among bipolar patients, even with music that is typically positive or uplifting.

Taken together, these findings suggest a complex role of music in emotion regulation: what serves as a calming stimulus for one group may act as a stressor for another, underscoring the importance of listener-specific factors.

Table 5. The Variance & Std of Observations in Different Settings

Research Question	Disorder	Interval	# Observations	# Authors	Conditional Model		Zero Inflation Model	
					Variance	Std	Variance	Std
RQ 1 (Music Listening)	Depression	30m	3593	809	0.173	0.415	0.972	0.986
		1h	5016	949	0.162	0.403	0.994	0.997
	PTSD	30m	1129	255	0.117	0.342	0.413	0.643
		1h	1634	298	0.175	0.419	0.890	0.943
	Bipolar	1h	1211	266	0.176	0.419	0.991	0.995
RQ 2 (Pop)	Depression	30m	1893	577	0.209	0.457	1.207	1.099
		1h	2644	686	0.197	0.444	1.030	1.015
	PTSD	30m	499	157	0.160	0.400	0.346	0.588
		1h	742	196	0.179	0.423	0.879	0.937
RQ 2 (Rock)	Bipolar	30m	207	88	0.088	0.298	5.93E-09	7.70E-05
		1h	308	105	0.111	0.334	0.654	0.808
RQ 2 (Hip Hop)	PTSD	30m	66	31	0.177	0.421	1.55E-09	3.94E-05
RQ 2 (Jazz)	Depression	1h	132	98	2.25E-01	4.75E-01	8.51E-09	9.22E-05
	Anxiety	30m	52	34	2.37E-09	4.87E-05	1.71E-12	1.31E-06
		1h	71	44	5.65E-09	7.52E-05	5.49E-28	2.34E-14
RQ 3 (Fast Tempo)	Depression	30m	1243	453	0.124	0.353	0.66	0.810
		1h	1725	564	0.091	0.302	0.8694	0.932
	Anxiety	30m	650	194	0.029	0.173	1.046	1.02E+00
		1h	878	238	0.021	0.146	0.565	0.751
	PTSD	1h	594	180	0.099	0.315	0.71	0.843
RQ 3 (Slow Tempo)	Depression	30m	65	57	0.437	0.661	1.268	1.126
	PTSD	1h	39	29	0.416	0.645	3.58E-08	0.0001
	Bipolar	30m	16	13	0.061	0.248	4.91E-09	7.01E-05
RQ 3 (Low Valence)	Bipolar	30m	251	119	0.034	0.184	0.4013	0.633
	Depression	1h	1487	535	0.114	0.338	0.656	0.815
RQ 3 (High Valence)	Bipolar	30m	185	92	0.043	0.209	0.816	0.903
RQ3 (Low Instrumentalness)	Depression	30m	3005	728	0.107	0.327	0.982	0.991
		1h	4204	855	0.097	0.311	1.018	1.009
	Anxiety	1h	2329	361	0.065	0.256	0.947	0.973
RQ3 (High Instrumentalness)	Bipolar	30m	15	12	0.078	0.279	0.023	0.152

5.2 RQ 2: Stress Levels Across Mental Disorder Groups and Genres

For RQ2, we investigated how different music genres might influence stress levels among users with mental health disorders, as compared to controls. We analyzed four major genres (Pop, Rock, Hip Hop, and Jazz) and looked at stress responses within one hour of listening (at 30-minute and 60-minute intervals). Figure 6 summarizes these genre-related stress outcomes, showing the stress level differences (and 95% confidence intervals) for each genre between the disorder groups and their PSM-matched control group. In the following subsections, we highlight the notable genre-specific findings for each mental health group, with statistical results summarized in Tables 7 to 10.

Table 6. Association of Listening to Music with Stress Level

Group	Interval	Conditional Model				Zero Inflated Model			
		Estimate	Std. Error	CI	Pr(>|Z|)	Estimate	Std. Error	CI	Pr(>|Z|)
Deprresion	30m	0.158	0.056	[0.047, 0.270]	0.005 **	-0.185	0.138	[-0.457, 0.08]	0.179
	1h	0.142	0.047	[0.048, 0.236]	0.002 **	-0.349	0.123	[-0.591, -0.107]	0.004 **
PTSD	30m	0.146	0.083	[-0.016, 0.3105]	0.078 .	0.075	0.175	[-0.269, 0.419]	0.669
	1h	0.127	0.077	[-0.023, 0.279]	0.097 .	-0.090	0.185	[-0.454, 0.274]	0.627
Bipolar	1h	0.151	0.086	[-0.018, 0.321]	0.0803 .	0.045	0.213	[-0.372, 0.463]	0.830

5.2.1 Finding 1 (Pop Music Consumption). For users with **depression**, exposure to pop music was associated with a time-dependent increase in stress. After 30 minutes of listening, their stress levels were about 14% higher than those of the control group ($p \approx 0.066$), and by 60 minutes this difference grew to a significant 23.5% increase ($p < 0.001$). Consistent with these elevations, the zero-inflated model indicated that depressed users were 29%-32% less likely to report zero stress at 30 minutes ($p \approx 0.066$) and 60 minutes ($p = 0.011$) after listening to pop music than their matched control group. A similar pattern emerged for users with **PTSD**: 30 minutes after listening to pop music, stress levels in the PTSD group were roughly 33% higher than in controls ($p = 0.021$). This effect slightly diminished by the 60-minute mark (around 19% above controls) with ($p \approx 0.075$). In contrast, neither the **anxiety** nor **bipolar** groups showed any significant change in stress relative to controls after listening to pop music at either interval. The results are presented in Table 7.

Prior studies have highlighted the therapeutic potential of pop music. For example, Huang and Duell [46] demonstrated that carefully curated, relaxing pop songs can significantly reduce anxiety among adolescents. Similarly, Kresovich [58] found that pop songs addressing mental health struggles can foster greater empathy in college students. These positive outcomes align with theories that music (including pop) serves as a tool for mood regulation and emotional support . However, whether pop listening alleviates or aggravates stress appears to depend on context and individual differences [33, 81]. Our finding that pop music listening was linked to heightened stress in the depression and PTSD groups underscores this complexity. It is possible that, for some vulnerable individuals, even familiar or popular music might trigger negative emotions or stress, especially if the lyrical content or personal associations of the music reinforce their current mood. This dual nature of music's impact – potentially beneficial in general but sometimes detrimental for certain listeners – reflects the nuanced role of music as either a healing influence or a stressor [33].

5.2.2 Finding 2 (Rock Music Consumption). In our analysis, rock music exposure had a pronounced association primarily for the **bipolar** group. Individuals with bipolar disorder experienced a sharp rise in stress after listening to rock: at 30 minutes post-listening, their stress levels were about 34–35% higher than those of controls $p \approx 0.07$). Reinforcing this result, the zero-inflated model showed that bipolar users were 46% less likely to report zero stress at the 30-minute mark ($p = 0.032$) when listening to rock. By 60 minutes their stress was approximately 45% higher than controls ($p \approx 0.007$). This suggests a strong and lasting stress response to rock music in this group. By contrast, we observed no significant stress-level changes associated with rock music in the **depression**, **anxiety**, or **PTSD** groups.(see Table 8 for the rock music results.)

Table 7. Association of Pop genre with stress

Group	Interval	Pop - Conditional Model				Pop - Zero Inflated Model			
		Estimate	Std. Error	CI	Pr(>|Z|)	Estimate	Std. Error	CI	Pr(>|Z|)
Depression	30m	0.137	0.074	[-0.008, 0.283]	0.065 .	-0.347	0.189	[-0.719, 0.02]	0.066 .
	1h	0.2107	0.061	[0.08, 0.331]	0.0006 ***	-0.396	0.189	[-0.768, -0.02]	0.011 *
PTSD	30m	0.291	0.126	[0.044, 0.538]	0.020 *	0.269	0.157	[-0.038, 0.578]	0.269
	1h	0.185	0.103	[-0.018, 0.389]	0.074 .	-0.032	0.250	[-0.522, 0.451]	0.897

Table 8. Association of Rock genre with stress

Group	Interval	Rock - Conditional Model				Rock - Zero Inflation Model			
		Estimate	Std. Error	CI	Pr(>|Z|)	Estimate	Std. Error	CI	Pr(>|Z|)
Bipolar	30m	0.297	0.163	[-0.022, 0.617]	0.068 .	-0.614	0.287	[-1.178, -0.051]	0.032 *
	1h	0.370	0.136	[0.103, 0.637]	0.006 **	-0.232	0.332	[-0.881, 0.421]	0.484

Our finding for the bipolar group aligns with a broader literature noting potential negative emotional effects of heavy or intense music genres. Some studies have reported that rock music fans tend to exhibit lower overall mental well-being compared to fans of other genres [9]. Similarly, research on heavy metal music—a genre closely related to rock—has linked it to increases in feelings of sadness and other mood disturbances in listeners [68]. Heavy metal fans have also been observed to report higher levels of depression, anxiety, and anger relative to non-fans [85]. These patterns are consistent with our observation that rock music exposure was associated with heightened stress in individuals with bipolar disorder. They suggest that more intense and aggressive musical styles might act as stressors or triggers for negative emotions, particularly in vulnerable populations.

5.2.3 Finding 3 (Hip Hop Music Consumption). Overall, listening to hip hop music was linked to lower stress levels in our sample, although the observed associations were modest. The clearest evidence of a hip hop benefit was observed in the **PTSD** group: 30 minutes after hip hop exposure, users with PTSD had stress levels about 42% lower than their control counterparts ($p = 0.041$). However, this reduction was short-lived; by the 60-minute mark, the PTSD group's stress was no longer significantly different from the control group. The other mental health groups (**depression**, **anxiety**, and **bipolar**) did not exhibit any statistically significant changes in stress following hip hop music listening (though their estimated effects were in the direction of lower stress). The results are presented in Table 9. Our finding of lower stress levels linked with hip hop music, particularly for PTSD, aligns with emerging research on hip hop as a therapeutic tool. In a trauma-focused context, hip hop-based interventions have been found highly effective for at-risk adolescents (such as refugee youth), helping to address trauma and build resilience over the course of a multi-month therapy program [40]. Another study also demonstrated that engaging in hip hop music creation (writing and performing mixtapes in a group counseling setting) significantly lowered stress, depression, and anxiety among vulnerable youth [59]. These studies support our observation that hip hop listening coincided with reduced stress in the PTSD group.

Table 9. Association of Hip Hop genre with stress

Group	Interval	Hip-Hop Conditional Model				Hip-Hop Zero Inflation Model			
		Estimate	Std. Error	CI	Pr(>|Z|)	Estimate	Std. Error	CI	Pr(>|Z|)
PTSD	30m	-0.547	0.268	[-1.073, -0.022]	0.040 *	0.510	0.522	[-0.513, 1.534]	0.328

Table 10. Association of Jazz genre with stress

Group	Interval	Jazz - Conditional Model				Jazz - Zero Inflation Model			
		Estimate	Std. Error	CI	Pr(>|Z|)	Estimate	Std. Error	CI	Pr(>|Z|)
Depression	1h	0.295	0.1702	[-0.037, 0.629]	0.082 .	-0.755	0.443	[-1.623, 0.113]	0.088 .
Anxiety	30m	-0.552	0.253	[-1.045, -0.054]	0.028 *	1.609	0.602	[0.429, 2.789]	0.007 **
	1h	-0.352	0.199	[-0.74, 0.038]	0.077 .	0.500	0.479	[-0.439, 1.441]	0.297

5.2.4 Finding 4 (Jazz Music Consumption). Jazz music was linked to somewhat divergent outcomes for different groups. In the **depression** group, listening to jazz was associated with an increase in stress at the 60-minute mark (approximately +34% relative to controls), with ($p \approx 0.08$), and no significant change was observed at 30 minutes. Notably, the zero-inflated model suggested that depressed users were also less likely to be stress-free after 60 minutes of jazz listening ($p = 0.088$), complementing our observation in the conditional model. In contrast, the **anxiety** group experienced a decrease in stress after jazz exposure: about 42% lower stress than controls at 30 minutes ($p = 0.029$), also the zero-inflated model suggested that anxious users were more likely to be stress-free after 30 minutes of jazz listening ($p \approx 0.09$), complementing our observation in the conditional model. The observed association with lower stress in the anxiety group weakened by 60 minutes (the stress reduction was around 35% at that point, $p \approx 0.08$). The **PTSD** and **bipolar** groups did not exhibit any significant stress changes in response to jazz music at either interval. Results are present in Table 10.

Recent studies have shown no clear influence of jazz listening on stress or mental health overall [20]. For example, one experiment noted that listening to jazz did not significantly improve physiological stress recovery (such as blood pressure normalization) compared to silence [18]. Another study reported no significant differences in anxiety outcomes between people who listened to jazz and those who did not, implying that jazz on its own may neither help nor harm emotional state in a predictable way [21]. These observations differ somewhat from our findings, where we observed a modest stress-relief effect in anxious individuals and a slight stress increase in those with depression.

5.3 RQ 3: The Association of Music Features with Stress Levels Across Mental Health Groups

Our third research question (RQ3) addressed whether specific musical features—namely tempo, valence, and instrumentalness—are associated with different stress outcomes in the user groups. For this analysis, we categorized each feature into ”low” and ”high” levels to compare their effects. In the case of tempo, we drew on prior definitions [69] by treating songs below a certain beats-per-minute (BPM) threshold as slow tempo and those above a higher BPM threshold as fast tempo (for example, $< 70BPM$ vs. $> 130BPM$). For valence (the emotional positivity of a song) and instrumentalness (the extent to which a track is instrumental rather than vocal), we split the feature values into low and high bands based on their distribution in our dataset [69]. This allowed us to contrast, for instance, songs with very low valence (more negative or sad) versus very high valence (more positive or happy), and tracks with minimal vocals versus those with predominantly vocal content. The following subsections outline the key findings for each feature category, with results summarized in Tables 11 to 13.

5.3.1 Finding 1 (Slow and Fast Tempo). Significant differences emerged when investigating slow and fast music. In the **depression** group, fast-tempo songs are associated with markedly higher stress: about a 24% increase in stress at 30 minutes post-listening $p = 0.004$), and an 18–19% increase still evident at 60 minutes ($p = 0.003$). The result of zero-inflated model also suggest that users with depression were 88% less likely to report zero stress level when listening to slow music ($P = 0.074$) although no significant stress change was observed in the conditional model. The **anxiety** group showed a similar initial pattern, with fast music linked to a short-term stress spike of roughly 23% at 30 minutes ($p = 0.011$); however, this effect was transient, and by 60 minutes the stress levels of the anxiety group were not significantly different from those under fast music. Notably, the zero-inflated model suggested that even at 60 minutes, the anxiety group was somewhat less likely to report being stress-free after fast music ($p \approx 0.05$). The **PTSD** group exhibited a strikingly opposite trend for tempo: rather than finding slow music calming, individuals with PTSD experienced a dramatic stress increase in the slow-tempo condition. After 60 minutes of listening to slow music, their stress was about 147% higher than in the fast music condition ($p \approx 0.007$), indicating an unexpected adverse reaction to slow, presumably relaxing, music. Meanwhile, the **bipolar** group had no immediate reaction to tempo at 30 minutes in the conditional analysis. However, the zero-inflated model revealed that bipolar users were already less likely to report zero stress after 30 minutes of slow music ($p = 0.024$), despite no significant stress changes at that point. but over time, Fast-tempo music correlated with approximately 19% higher stress levels at 60 minutes ($p = 0.029$), indicating a delayed relationship. The results are presented in Table 11.

These findings illustrate that the relationship between musical tempo and stress is highly contingent on the listener's mental health status. Under typical circumstances, slow music is often associated with relaxation – for example, one study noted that low-tempo music significantly reduced heart rate in individuals undergoing high-stress tasks [106]. Our PTSD result, however, runs counter to this expectation, highlighting that what is "relaxing" for many can be agitating for some. On the other hand, some experimental studies have found no overall effect of tempo on stress in the general population. In one such study focusing on immediate mood and stress outcomes reported no significant influence of musical tempo on stress indicators in a general population [67]. In another study, fast vs. slow songs produced brain-wave patterns differ between tempo conditions [25], although no significant differences in physiological stress measures (e.g., heart rate or skin conductance) are observed. These subtle brain changes were interpreted as reflecting factors like increased concentration or mild anxiety, which aligns with the stress increase we observed with fast music in different mental-health groups. The outcomes of these studies underscore the importance of listener context: the lack of an effect in an average group does not mean tempo has no effect on anyone; rather, it suggests that certain populations (like those with mental health conditions) can react very differently [24]. Indeed, our analysis shows that a listener's psychological profile can significantly alter how tempo influences stress – what calms one person may alarm another. Tempo, in sum, does not have a uniform effect on stress levels, but its impact is moderated by who is listening and in what emotional state.

5.3.2 Finding 2 (Low and High Valence). The emotional tone of music (negative vs. positive mood) showed a distinct association across conditions. Low-valence music (songs with negative or sad emotional content) was associated with increased stress in several groups. For instance, in the **depression** group, listening to low-valence music was associated with a delayed rise in stress: about a 14% increase relative to controls, observed at the 60-minute point ($p = 0.034$). Consistent with this finding, the zero-inflated model showed that depressed users were 34% less likely to report zero stress at 60 minutes post-listening to low-valence music ($p = 0.015$), indicating fewer stress-free sessions in the depressed group under sad music. The **bipolar** group also reacted to low-valence

Table 11. Association of Tempo with stress

Group	Interval	Category	Tempo - Conditional Model				Tempo - Zero Inflation Model			
			Estimate	Std. Error	CI	Pr(>|Z|)	Estimate	Std. Error	CI	Pr(>|Z|)
Depression	30m	Fast	0.214	0.074	[0.068, 0.360]	0.003 **	0.067	0.187	[-0.298, 0.434]	0.718
	1h		0.170	0.057	[0.056, 0.283]	0.003 **	-0.278	0.171	[-0.615, 0.058]	0.105
	30m	Slow	-0.150	0.290	[-0.718, 0.418]	0.605	-2.168	1.216	[-4.553, 0.216]	0.074 .
Anxiety	30m	Fast	0.205	0.081	[0.046, 0.364]	0.011 *	0.258	0.273	[-0.278, 0.795]	0.346
	1h		0.094	0.062	[-0.028, 0.217]	0.133	-0.410	0.081	[-0.569, -0.252]	0.052 .
PTSD	1h	Slow	0.902	0.212	[0.487, 1.318]	0.006 **	-0.539	0.652	[-1.818, 0.740]	0.409
Bipolar	30m	Slow	0.436	0.368	[-0.285, 1.158]	0.236	-3.044	1.345	[-5.681, -0.407]	0.023 *
	1h	Fast	0.173	0.079	[0.017, 0.328]	0.029 *	-0.069	0.319	[-0.695, 0.556]	0.828

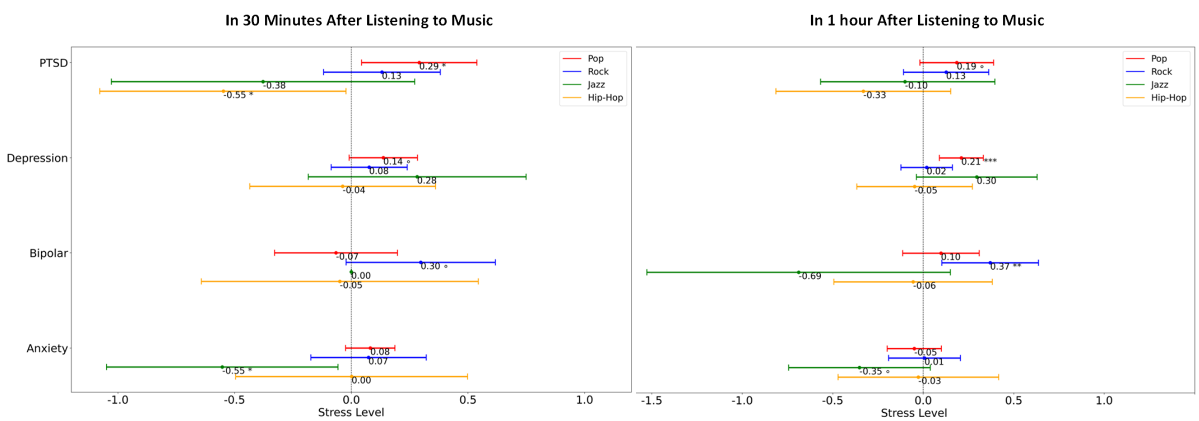


Fig. 6. An overall comparison of stress levels after listening to music for 30 (left) minutes and 1 hour (right). The plot shows the 95% confidence intervals for each genre across the two groups: control and disorders. Although many variables are not statistically significant, there is a noticeable potential association between stress levels and music genres.

music: we observed about a 25–26% stress increase at 30 minutes into listening to sad/negative music ($p = 0.047$). In all of these cases, exposure to emotionally negative themes in music elevated stress relative to controls. By contrast, high-valence (happy or positive) music did not produce any significant stress changes in the depression, PTSD, or bipolar groups at either time interval. (The **anxiety** group likewise showed no strong reactions to music valence in our data.) However, the zero-inflated model did reveal a subtle effect for the bipolar group even with positive music: bipolar users were less likely to report zero stress at 30 minutes when listening to high-valence

Table 12. Association of Valence with stress

Group	Interval	Category	Valence - Conditional Model				Valence - Zero Inflation Model			
			Estimate	Std. Error	CI	Pr(>|Z|)	Estimate	Std. Error	CI	Pr(>|Z|)
Bipolar	30m	Low	0.229	0.115	[0.002, 0.456]	0.047*	0.173	0.307	[-0.428, 0.775]	0.573
		High	0.017	0.136	[-0.249, 0.283]	0.896	-0.903	0.428	[-1.741, -0.06]	0.034 *
Depression	1h	Low	0.130	0.061	[0.009, 0.251]	0.034 *	-0.421	0.173	[-0.759, -0.080]	0.015 *

tracks ($p = 0.034$), meaning that even ”happy” music did not fully eliminate stress responses in the bipolar group. The results are presented in Table 12.

These observations align with prior research on music valence and emotional state. In general, studies have found that the ”happier” the music, the lower the stress response. For instance, one longitudinal study reported that a song's valence is significantly associated with stress and mood over time – with more positive (high-valence) music linked to lower stress levels and improved mood [30]. Similarly, experimental evidence indicates that positive-valence music promotes better recovery from stress than negative-valence music, both subjectively and physiologically [84]. This is in line with our finding that ”happy” music did not exacerbate stress, whereas ”sad” music did. On the other hand, the literature also showed that music with negative emotional content can have detrimental effects for vulnerable listeners. For example, research in [34] found that listening to sad music significantly increased feelings of depression in individuals with a predisposition to depression. This is consistent with our observation that low-valence music heightens stress in users with depression. Furthermore, sad music has been shown to induce changes in emotion-related memory and judgment, though these effects depend on the music's personal relevance to the listener and on individual traits like empathy [98].

5.3.3 Finding 3 (Low and High Instrumentalness). When considering the presence of vocals in music (vocal-heavy and purely instrumental tracks), we observed distinct patterns for different conditions. In the **depression** group, music with vocals (i.e. low instrumentalness) was associated with higher stress levels. After 30 minutes of listening to vocal-heavy songs, depressed users' stress was about 15% higher than that of their control counterparts ($p = 0.007$). Consistent with this finding, the zero-inflated model showed that depressed users were 27% less likely to report zero stress at 30 minutes post-listening to low-instrumentalness music ($p = 0.032$). This relationship grew slightly by 60 minutes, reaching roughly a 16% stress increase ($p < 0.001$); similarly, depressed users were 29% less likely to report zero stress at 60 minutes when listening to music with vocals ($p = 0.009$). By contrast, listening to purely instrumental pieces did not significantly change stress for the depression group at either time. The **bipolar** group showed nearly the opposite pattern. For individuals with bipolar disorder, music that was predominantly instrumental (lacking vocals) was linked to a strong stress response: a 73% increase in stress after 30 minutes of instrumental music compared to controls ($p = 0.048$). In contrast, bipolar listeners did not experience any significant stress change when the music included vocals. In the **anxiety** and **PTSD** groups, we did not find any notable differences in stress between instrumental and vocal music conditions. Nonetheless, the zero-inflated model uncovered that anxiety-group listeners were significantly less likely to report zero stress following vocal music at the 60-minute mark ($p = 0.018$), even though their overall stress levels did not significantly change as shown in the conditional model. This suggests that for anxious individuals, music with lyrics may subtly reduce the likelihood of being completely stress-free (perhaps inducing some stress) compared to instrumental music. The results are presented in Table 13.

Table 13. Association of Instrumentalness with stress

Group	Interval	Category	Instrumentalness - Conditional Model				Instrumentalness - Zero Inflation Model			
			Estimate	Std. Error	CI	Pr(>|Z|)	Estimate	Std. Error	CI	Pr(>|Z|)
Bipolar	30m	High	0.547	0.277	[0.003, 1.091]	0.048 *	-0.922	1.166	[-3.208, 1.363]	0.429
Depression	30m	Low	0.137	0.050	[0.037, 0.237]	0.006 **	-0.321	0.150	[-0.615, -0.026]	0.032 *
	1h	Low	0.151	0.043	[0.067, 0.236]	0.0004 ***	-0.346	0.133	[-0.60, -0.084]	0.009 **
Anxiety	1h	Low	-0.024	0.053	[-0.128, 0.079]	0.649	-0.407	0.172	[-0.744, -0.069]	0.018 *

These patterns can be better understood in light of prior studies on vocal content (instrumentalness) and emotional coping. Some research suggests that vocal-heavy music is often intertwined with emotional coping mechanisms. For instance, listeners may use songs with lyrics to express or process feelings, whereas instrumental pieces might serve more as background mood enhancers. One recent study on music as emotion regulation supports this idea, noting that whether a track with vocals can influence how people use it to cope emotionally [26]. This perspective aligns with our observation that the presence or absence of lyrics is correlated with stress outcomes. In [101]. researchers articulated that that individuals with depression tend to prefer or share music that is higher in instrumentalness, possibly to avoid lyrical content that could intensify their emotions. Literature also shown that the relationship between instrumental music and stress is not uniform across all populations. For example, one study noted that individuals with high anxiety levels are less inclined to listen to instrumental music, favoring songs with vocals [79]. Another study reported that people experiencing higher overall stress tend to prefer instrumental music, which resonates with our observations in the bipolar group (where high instrumentalness was associated with increased stress) [44]. Additionally, music preference research in the context of pain management has found that individuals in pain often choose music that is more engaging – with higher energy, higher danceability, and with vocals (low instrumentalness) – over purely instrumental tracks [44]. This contrasts with our depression-group findings (where vocal music heightened stress), but it highlights how context-dependent these effects are. Listeners in acute physical pain may prefer lively, lyrical music to distract from discomfort, whereas listeners with depression might avoid lyrical content to prevent emotional overload.

6 Applications of Findings in a Music Recommender System

The insights from our analysis of music genres, audio features, and stress levels across different mental health groups provide a foundation for developing personalized music recommendation systems aimed at emotional well-being. Such systems seek to cater to users' musical tastes while also promoting positive mood regulation. Recommendation systems are ubiquitous in domains like e-commerce, social media, and entertainment [107], filtering information to suggest items that match user preferences. Recent advances in natural language processing, especially with large language models (LLMs) like BERT [55], and GPT-3 [15] as well as generative recommender approaches (e.g., FLAN-T5 [23]), have opened new possibilities for enhancing recommenders. Researchers are exploring how LLMs' reasoning and contextual understanding can improve recommendations [45, 61]. Conventional music recommenders, which rely heavily on a user's listening history and immediate preferences [12], often optimize for short-term satisfaction while overlooking long-term mental health effects. For example, matching a user's current sad mood with melancholic songs might provide temporary comfort but could reinforce negative feelings over time. A more mindful strategy is needed—one that balances user preferences with the potential therapeutic effects of music [53, 63]. In line with this goal, and to investigate **RQ4**, we developed an

LLM-based framework that integrates the findings from our previous analyses (RQ2–RQ3) on the relationship between music listening and stress for the music recommendation task.

6.1 Proposed LLM-Based Music Recommendation Framework

We implemented a proof-of-concept recommendation framework that uses an LLM to balance user satisfaction with mental health benefits. The approach can be summarized in four main steps:

(1) **Detect Stress-Reducing Songs**: Identify user timelines where a user's stress level notably decreases after listening to a particular song. We label this track as the "selected song" for that user session, as it is associated with stress reduction.
(2) **Profile Recent Listening History**: Construct the user's listening profile from at least ten prior music sessions before the selected song. For each session, extract key features (e.g., genre, tempo, instrumentalness, valence) that characterize the music the user listened to.
(3) **Generate Candidate Songs**: From the full music database, compile a pool of candidate songs the user has not heard before (excluding any track already in the user's history or the selected song itself). Randomly sample 19 songs from this pool and combine them with the selected song, yielding a set of 20 candidate tracks for recommendation.
(4) **LLM-Based Ranking**: Provide the LLM with a prompt containing the user's profile (from step 2) and the list of 20 candidate songs with their audio feature values. The LLM is asked to rank these candidates in terms of how well each song aligns with the user's historical preferences and its potential to improve the user's mood and reduce stress. The selected song (which is known to have reduced the user's stress) serves as the "correct" item that an ideal recommender should rank highly. Figure 8 illustrates an example of the prompt used for this LLM ranking, contrasting a basic prompt (baseline) with an enriched version incorporating insights from one of the features (genres) from our results.. The detailed steps of the recommendation task are presented in Algorithm 1.

Using the above framework, we first establish a *baseline model* as a point of comparison. The baseline uses the LLM to perform a simple re-ranking of the 20 candidates based only on the user's listening history, without any information about stress-related effects of music features. This provides a foundational recommendation without our new insights. While the baseline approach can capture the user's taste to some extent, it does not incorporate knowledge of how specific musical features (e.g., tempo or genre) might be linked to stress levels.

6.2 Feature Ablation Experiment Design

We progressively integrated our research findings into the LLM's recommendation process and tested how each type of musical feature contributes to stress-aware recommendations. We conducted a series of experiments under different feature conditions (with all other aspects of the system identical to the baseline) to isolate the impact of each feature on the recommender's performance. The experimental conditions were as follows:

(1) **Only Genre:** The LLM receives information only about the music genre of each candidate track, including how that genre tends to affect stress for the user's demographic (e.g., rock music was associated with a 5% stress reduction for similar users).
(2) **Only Tempo:** The LLM is given only the tempo of each song and the insight about tempo's effect on stress (e.g., whether slower or faster tempo correlates with stress changes).
(3) **Only Valence:** The LLM is given only the valence of each song plus the insight about valence's relationship with stress.
(4) **Only Instrumentalness:** The LLM is given only the instrumentalness of each song and the insight regarding how instrumental vs. vocal music influences stress.

In each of these scenarios, the recommendation procedure remained the same as in the baseline framework—steps 1–4 above—changing only the feature-related information included in the LLM's prompt. In a single-feature condition like "Only Genre," the prompt was identical except that it included only the genre information for each song (plus the genre–stress insight). This feature-ablation approach allows us to assess the relative importance of each musical attribute in guiding the LLM's recommendations. By comparing performance across these conditions, we can determine which aspects of music have the strongest impact on reducing user stress and thus should be prioritized in a stress-aware recommender system.

6.3 Ranking Evaluation and Results

To evaluate the recommender's performance under each experimental condition, we used the Mean Reciprocal Rank (MRR) metric [49], a standard measure in information retrieval and recommendation tasks. MRR is well-suited here because it captures how highly the system ranks the "correct" item—in our case, the stress-reducing selected song—among the list of 20 candidates. For each user session (each instance of the experiment), we recorded the rank position of the selected song in the LLM-produced ranking and then computed the reciprocal rank. We then averaged these reciprocal ranks across all user sessions for a given condition to get an overall MRR score for that condition.

We compared the MRR of each feature condition against the baseline to quantify improvement. A higher MRR indicates that the selected stress-reducing song was ranked closer to the top, meaning the recommender more successfully prioritized the song that actually benefited the user's stress. We performed paired t-tests between each experimental condition's MRR values and the baseline's MRR values to assess the statistical significance of improvements. All configurations (except only-valence) integrating our insights showed improvement over the baseline, and these improvements were statistically significant in most cases ($p < 0.05$). Figure 7 depicts the average MRR achieved in each feature setting alongside the significance of the difference from baseline.

The results reveal notable variation in how each feature insight contributed to recommendation quality. Among the single-feature experiments, the "Only Genre" condition achieved the highest MRR (0.354), indicating that genre information alone was very effective in helping the LLM rank stress-reducing tracks at the top. The "Only Instrumentalness" condition yielded the second-highest MRR (0.301), suggesting that whether a song is instrumental or vocal also plays an important role in identifying calming music. The other single-feature conditions ("Only Tempo" and "Only Valence") resulted in MRR scores (approximately 0.187 and 0.175, respectively) that were equal to or slightly below the baseline. These outcomes underscore that not all features contribute equally to improving recommendations for stress reduction—genre (and to a lesser extent instrumentalness) appears to be especially influential, whereas valence and tempo alone offer limited benefit in this context.

Overall, by integrating targeted musical features derived from our findings, the LLM-based system was better at retrieving the song that actually reduced user stress, thereby answering **RQ4** affirmatively.

7 Discussion

7.1 Summary of Findings

This study investigated the relationship between music listening and stress levels among individuals with different mental health conditions, using a large dataset of social media activity. The results reveal a nuanced and condition-specific relationship between music engagement and stress outcomes.

For RQ1, our findings show that music listening is not universally stress-reducing. In fact, individuals with depression and PTSD exhibited significantly higher stress levels after listening to music compared to matched control users. The stress increase was most pronounced for those with depression, while individuals with bipolar disorder showed a delayed stress elevation. Notably, no significant stress changes were observed in the anxiety group, suggesting that music's effects may be less pronounced or more variable in this population.

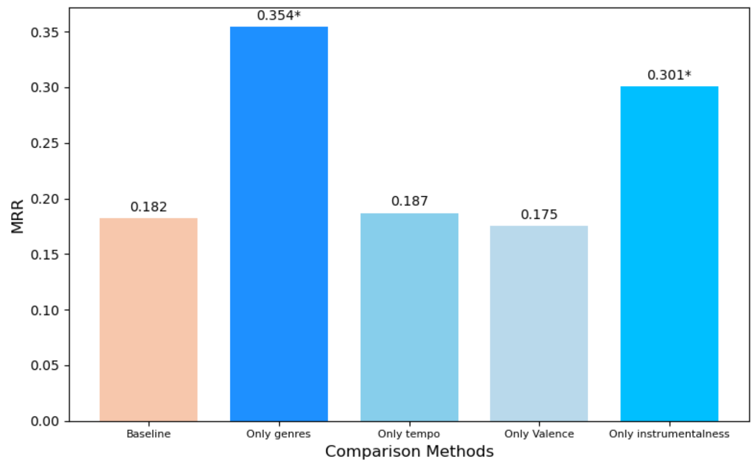


Fig. 7. Comparison of MRRs in different setups. The figure illustrates the average MRR achieved for each condition, highlighting improvements over the baseline. Statistically significant differences from the baseline are indicated ($p < 0.05$ for all conditions), demonstrating the effectiveness of integrating music features in stress-aware recommendation.

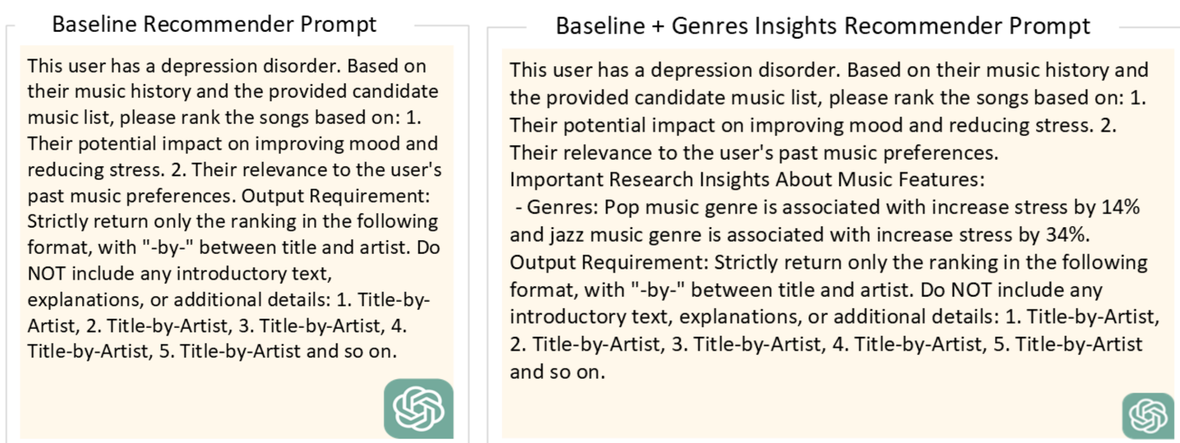


Fig. 8. Example of prompts used in the baseline recommender system (left) and with "All Features" insight (right).

RQ2 examined how different music genres interact with mental health. Pop music was linked to heightened stress in users with depression and PTSD, particularly at 60 minutes post-listening. Rock music exposure produced a marked and sustained stress increase in individuals with bipolar disorder, aligning with prior concerns around intense or emotionally charged music in vulnerable groups. Conversely, hip hop listening showed a short-term stress-reducing effect for users with PTSD, echoing emerging therapeutic literature on the genre's potential for emotional processing. Jazz elicited more mixed responses: it was associated with a modest stress reduction in the anxiety group, but a slight stress increase among individuals with depression.

Algorithm 1 LLM-Based Music Recommendation Framework

Input: User listening history H, music database D
Output: Ranked list of recommended songs R

```
stress_song ← IdentifyStressReducingSong(H)
profile_features ← ExtractFeatures(H, features = ['genre', 'tempo', 'valence', 'instrumentalness'])
candidate_songs ← SampleCandidateSongs(D, H, exclude = stress_song, sample_size = 19)
candidate_pool ← candidate_songs ∪ {stress_song}
feature_settings ← {
    "genre_only": ['genre'],
    "tempo_only": ['tempo'],
    "valence_only": ['valence'],
    "instrumentalness_only": ['instrumentalness'],
    "baseline": [ ]
}
for setting ∈ feature_settings do
    prompt ← GenerateLLMPrompt(profile_features, candidate_pool, feature_settings[setting])
    rankings[setting] ← RankSongsWithLLM(prompt)
    ranks[setting] ← GetRank(stress_song, rankings[setting])
end for
for setting ∈ feature_settings do
    MRR[setting] ← CalculateMRR(ranks[setting])
end for
PerformStatisticalTests(MRR, baseline_setting = "baseline")
```

RQ3 focused on how specific acoustic features—tempo, valence, and instrumentalness—relate to stress outcomes. Tempo effects were particularly striking: while fast-tempo music elevated stress for individuals with depression and anxiety, slow-tempo music unexpectedly heightened stress among those with PTSD. Valence also played a critical role: low-valence (sad or negative) music increased stress in depression and bipolar groups, whereas high-valence (positive) music did not significantly alter stress levels. Instrumentalness revealed a divergent relationship: vocal-heavy tracks increased stress in depressed users, while instrumental tracks elevated stress in the bipolar group.

7.2 Limitations

While this study contributes valuable insights into the interplay between music and stress across mental health conditions, several limitations should be noted. First, the analysis is based on publicly available social media data, which may not capture the full scope of individual emotional experiences or contextual variables such as listening intent, setting, or concurrent life events. Second, stress levels were estimated using a computational model (TensiStrength), which, although validated, may not reflect clinical stress diagnoses or nuanced emotional states.

Additionally, the dataset does not include ground truth about the user's mental health status beyond self-reported diagnoses, which introduces potential for misclassification. The musical features and genre categorizations were inferred from metadata and external APIs, limiting the ability to assess the subjective emotional content or personal significance of specific songs. Finally, causal relationships between music listening and stress cannot be inferred from the observational nature of the data, and unmeasured confounding variables may influence the outcomes.

7.3 Ethical Considerations

The data used in this study were derived from a publicly available dataset built from Twitter posts. All data collection and analysis were conducted in accordance with ethical standards for research involving social media data, including the removal of identifiable user information and the use of aggregate-level analysis to prevent re-identification.

That said, ethical sensitivity remains important when interpreting findings related to mental health, especially when discussing potentially vulnerable populations. While the dataset offers a valuable window into large-scale behavioral patterns, it is crucial not to conflate observed stress indicators with clinical diagnoses or assume uniform emotional responses within mental health groups. Future work should consider engaging with clinical experts and individuals with lived experience to guide interpretation and application of findings.

7.4 Implications: Practical and Theoretical

Our findings offer several important implications at the intersection of music, emotion regulation, and mental health, spanning both practical applications and theoretical contributions.

Practical Implications. Our findings underscore the need for emotionally aware recommender systems in mental health contexts. While music platforms and mental health apps increasingly use AI to personalize content, they rarely account for users' mental health profiles [78, 96]. Evidence from our study—such as increased stress after pop or slow-tempo music in users with depression or PTSD—suggests that generic recommendations may do harm. Systems that incorporate diagnosis and prior emotional responses could offer safer, more supportive suggestions.

For example, such a system might learn to steer a PTSD-affected listener away from ostensibly "calming" music that, for them, yields anxiety, instead favoring content that that user has historically found soothing. Incorporating our insights on features like genre, valence sensitivity, and tempo reactivity into music recommendation algorithms could thus reduce the risk of unintentional emotional harm and improve user wellness [64]. This extends recent research on mood-based music recommenders by adding a mental health-aware dimension to personalization, moving beyond generic mood matching toward therapeutic music curation.

From a public health perspective, leveraging music listening behaviors captured through social media data presents a novel avenue for monitoring and supporting community mental health. While traditional digital profiling methods have utilized passive smartphone data—such as GPS, call logs, and app usage—to assess mental health states like stress, anxiety, and depression [72], these approaches often overlook the rich emotional context provided by music engagement. Recent studies have highlighted that music listening is intricately linked to stress modulation and mood regulation in daily life [57]. However, existing systems rarely integrate music listening patterns into real-time mental health assessments. By analyzing large-scale data from platforms like Twitter, our approach demonstrates how passive digital footprints, specifically music-related behaviors, can reflect psychological stress in real time. This method offers a complementary perspective to traditional digital characterization, potentially enhancing the granularity and immediacy of mental health monitoring. Public health institutions could harness such data to detect when populations exhibit signs of heightened stress—such as shifts in musical preferences correlating with stress-related discourse online—and respond with timely interventions or resources.

Unlike clinical assessments that require active participation, monitoring stress via everyday music engagement is unobtrusive and highly scalable. Our findings, therefore, contribute to the emerging paradigm of digital characterization for mental health, wherein online behaviors serve as proxies for well-being at the population level. Identifying stress through music listening trends could complement traditional surveillance by providing early warnings of societal stress crises or by highlighting at-risk groups based on their unique music response patterns. Notably, recent large-N studies have begun linking everyday music use to stress and mood outcomes in the general population [31], underscoring that our social-media-based approach can enhance ecological validity and reach. In sum, this work suggests practical pathways for both personalized technology design and public health strategy – from therapeutically informed music recommenders to population-level stress monitoring – grounded in people's natural music interactions.

Theoretical Implications. Theoretically, our study advances the understanding of how genre-specific emotional responses to music are modulated by individual mental health conditions, thereby extending prior research on music and emotion in clinical populations. Earlier studies of music's effect on mood were often limited to small samples or single-disorder groups in controlled settings [38, 66], which could not capture the complex, condition-specific patterns we observe at scale. In contrast, our large-scale analysis reveals that the emotional impact of music is highly condition-dependent: a genre that alleviates stress in one group may exacerbate it in another.

For example, while pop or slow-tempo pieces are generally thought to induce relaxation, we found they led to heightened stress in listeners with depression and PTSD, even as listeners without disorders or with anxiety showed little to no calming benefit. This nuanced result challenges the implicit assumption of uniformly calming effects of "relaxing" music and highlights that individuals with different diagnoses engage in distinct emotion-regulation processes through music. Our findings align with prior work showing individual variability in responses to emotional music [32]. Specifically, we extend this literature by demonstrating that such variability is not only present but also systematically patterned according to mental health conditions.

This diagnosis-specific patterning has important theoretical implications. In particular, our results lend support to mood-congruence theories by suggesting that vulnerable individuals may select or react more strongly to music that mirrors their affective state, which can inadvertently reinforce negative emotions. A person with depression, for instance, might gravitate toward somber tunes that resonate with their current mood, only to experience intensified sadness or stress afterward—a phenomenon hinted at in small experiments [32] but now evidenced on a much broader scale. Specifically, We contribute to theory by integrating psychopathology into models of music emotion: the differential effects of tempo, lyrical content, and genre across disorders add a new layer to frameworks of stimulus appraisal and affective reactivity [32, 54]. In essence, our study suggests that models of music-induced emotion and mood regulation must account for the listener's mental health context – an insight that bridges music psychology and clinical understanding of emotional reactivity.

Methodologically, this work introduces an innovative analytical framework by combining Propensity Score Matching with Generalized Linear Mixed Models (PSM + GLMM) to examine music and stress relationships. This approach represents a significant advancement over prior mental health studies that often relied on simple group comparisons or cross-sectional designs. By using PSM to match individuals with and without mental health conditions on relevant confounders, we approximated a quasi-experimental design that strengthens causal inference from observational social media data. At the same time, applying GLMM allowed us to model within-subject variability – capturing how repeated listening events affected stress for each user over time – and to include random effects for individuals. Together, this PSM+GLMM strategy enabled us to disentangle the music–stress relationship with a level of rigor and granularity unattainable in most earlier works.

Previous small-scale studies typically lacked either the sample size or the methodological flexibility to do this: for example, clinical trials or lab studies with dozens of patients could monitor within-person responses but not generalize broadly [66], whereas large surveys could find correlations but not isolate cause and could miss dynamic effects. Our integrated approach fills this gap by achieving both scalability and internal validity. The result is a more rigorously constructed dataset that enables analysis of both within-person changes in stress following music exposure and between-group differences across mental health conditions, all based on real-world behavioral data. This methodological contribution not only strengthens the validity of our findings but also provides a practical framework for future research connecting digital trace data with mental health. By demonstrating that careful matching and mixed-effects modeling can yield meaningful insights from noisy, in-the-wild data, we enhance the ecological validity of mental health research.

In summary, the study's approach and outcomes substantially extend existing research – moving beyond the constraints of prior small or clinic-bound studies – and illustrate how new data-driven methods can deepen theoretical understanding and inform practical interventions in the mental health domain.

8 Conclusion and Future Works

This study explored the complex relationship between music listening and stress levels across different mental health conditions using a large-scale social media dataset. By integrating Propensity Score Matching (PSM) and Generalized Linear Mixed Models (GLMMs), we addressed key research questions about the association of music genres and features with stress responses in individuals with depression, anxiety, PTSD, and bipolar disorder, compared to matched controls.

Our findings revealed that music's relationship with stress is highly context-dependent and varies significantly across mental health conditions. For instance, while pop and rock music were associated with increased stress in depression and bipolar groups, hip hop showed short-term stress-reducing effects for individuals with PTSD. Similarly, acoustic features such as tempo and instrumentalness exhibited distinct influences on stress, with fast-tempo and low-valence music often exacerbating stress in the studied group. These results underscore the importance of personalized approaches to music-based interventions, as generic recommendations may inadvertently heighten stress for some individuals.

The development of an LLM-based music recommendation framework demonstrated the practical application of these insights, highlighting how genre-specific and feature-aware recommendations can better align with therapeutic outcomes. By prioritizing stress-reducing music selections, such systems have the potential to enhance emotional well-being while respecting individual preferences.

To deepen our understanding of the music-stress relationship, future research should extend the current observational framework by incorporating longitudinal designs to establish causal relationships between music listening and stress responses. Expanding the dataset to include underrepresented mental health conditions (e.g., schizophrenia, OCD) and diverse demographic groups would improve the generalizability of findings. Additionally, researchers could broaden data collection beyond Twitter to incorporate other social media platforms such as Reddit, Threads, and Instagram, while also leveraging multimodal data including voice recordings, images, and videos to provide a more comprehensive analysis of stress indicators. Collaboration with mental health professionals would be valuable for translating research findings into clinical practice, particularly in developing tailored music therapy protocols for specific disorders. Furthermore, investigating how cultural differences in music preferences and emotional expression influence the music-stress relationship would help create globally applicable recommendations. By addressing these research directions, future studies can further bridge the gap between computational social science and clinical applications, ultimately advancing the development of more effective, scalable, and personalized music interventions for mental health support.

9 Acknowledgement

This research was funded by the Natural Sciences and Engineering Research Council of Canada (NSERC) through Discovery Grants RGPIN-2022-03809, RGPIN-2022-04197, and RGPIN-2022-05193.

References

[1] Mazin Abdalla, Parya Abadeh, Zeinab Noorian, Amira Ghenai, Fattane Zarrinkalam, and Soroush Zamani Alavijeh. 2025. The Impact of Listening to Music on Stress Level for Anxiety, Depression, and PTSD: Mixed-Effect Models and Propensity Score Analysis. *IEEE*

Transactions on Computational Social Systems (2025), 1–15. doi:10.1109/TCSS.2025.3561073
[2] Soroush Zamani Alavijeh, Xingwei Yang, Zeinab Noorian, Amira Ghenai, and Fattane Zarrinkalam. 2025. Twitter-MusicPD: melody of minds-navigating user-level data on multiple mental health disorders and music preferences. *EPJ Data Science* 14, 1 (2025), 1–24.
[3] Soroush Zamani Alavijeh, Fattane Zarrinkalam, Zeinab Noorian, Anahita Mehrpour, and Kobra Etminani. 2023. What users' musical preference on Twitter reveals about psychological disorders. *Information Processing & Management* (2023).
[4] Jayme Garcia sArnal Barbedo and Amauri Lopes. 2006. Automatic genre classification of musical signals. *Journal on Advances in Signal Processing* 2007 (2006), 1–12.
[5] Krishna C Bathina, Marijn ten Thij, Lorenzo Lorenzo-Luaces, Lauren A Rutter, and Johan Bollen. 2020. Depressed individuals express more distorted thinking on social media. *arXiv preprint arXiv:2002.02800* (2020).
[6] Juliana Batubara, Roma Sinta, and Firta Panjaitan. 2025. The Effect of Meditation on Stress Reduction among College Students: A Quantitative Study. *Acta Psychologia* 3, 4 (2025), 162–170.
[7] François Bettega, Monique Mendelson, Clémence Leyrat, and Sébastien Bailly. 2024. Use and reporting of inverse-probability-of-treatment weighting (IPTW) for multi-category treatments in medical research: a systematic review. *Journal of Clinical Epidemiology* (2024), 111338.
[8] Sayali Bhandarkar, Bhagyashree V Salvi, and Pravin Shende. 2024. Current scenario and potential of music therapy in the management of diseases. *Behavioural Brain Research* 458 (2024), 114750.
[9] Tom T Bogt, William W Hale III, and Andrik Becht. 2021. "wild years": rock music, problem behaviors and mental well-being in adolescence and Young adulthood. *Journal of youth and adolescence* 50, 12 (2021), 2487–2500.
[10] Benjamin M Bolker, Mollie E Brooks, Connie J Clark, Shane W Geange, John R Poulsen, M Henry H Stevens, and Jada-Simone S White. 2009. Generalized linear mixed models: a practical guide for ecology and evolution. *Trends in ecology & evolution* (2009).
[11] Joke Bradt, Noah Potvin, Amy Kesslick, Minjung Shim, Donna Radl, Emily Schriver, Edward J Gracely, and Lydia T Komarnicky-Kocher. 2015. The impact of music therapy versus music medicine on psychological outcomes and pain in cancer patients: a mixed methods study. *Supportive Care in Cancer* 23 (2015), 1261–1271.
[12] John Breese, David Heckerman, and Carl Kadie. 2013. Empirical analysis of predictive algorithms for collaborative filtering. *preprint arXiv:1301.7363* (2013).
[13] Mollie E Brooks, Kasper Kristensen, Koen J Van Benthem, Arni Magnusson, Casper W Berg, Anders Nielsen, Hans J Skaug, Martin Machler, and Benjamin M Bolker. 2017. glmmTMB balances speed and flexibility among packages for zero-inflated generalized linear mixed modeling. *The R journal* (2017).
[14] Mollie E Brooks, Kasper Kristensen, Koen J van Benthem, Arni Magnusson, Casper W Berg, Anders Nielsen, Hans J Skaug, Martin Maechler, and Benjamin M Bolker. 2017. Modeling zero-inflated count data with glmmTMB. *BioRxiv* (2017), 132753.
[15] Tom B Brown. 2020. Language models are few-shot learners. *arXiv preprint arXiv:2005.14165* (2020).
[16] Katherine Burak and Adam Kashlak. 2024. Bootstrapping generalized linear models to accommodate overdispersed count data. *Statistical Papers* (2024), 1–20.
[17] Catherine Carr, Patricia d'Ardenne, Ann Sloboda, Carleen Scott, Duolao Wang, and Stefan Priebe. 2012. Group music therapy for patients with persistent post-traumatic stress disorder–an exploratory randomized controlled trial with mixed methods evaluation. *Psychology and Psychotherapy: Theory, Research and Practice* 85, 2 (2012), 179–202.
[18] Sky Chafin, Michael Roy, William Gerin, and Nicholas Christenfeld. 2004. Music can facilitate blood pressure recovery from stress. *British journal of health psychology* 9, 3 (2004), 393–403.
[19] Stevie Chancellor and Munmun De Choudhury. 2020. Methods in predictive techniques for mental health status on social media: a critical review. *NPJ digital medicine* (2020).
[20] Angela Chen. 2024. Music, Mental Health, and Mood Regulation: A Data-Driven Approach to Understanding the Role of Music in Emotional Well-being. (2024).
[21] Lori Elizabeth Chester. 1995. *The effects of precategorized jazz and classical music and silence on state anxiety in mental health professionals.* Texas Woman's University.
[22] Sabine Choppin, Wiebke Trost, Thibaut Dondaine, Bruno Millet, Dominique Drapier, Marc Vérin, Gabriel Robert, and Didier Grandjean. 2016. Alteration of complex negative emotions induced by music in euthymic patients with bipolar disorder. *Journal of affective disorders* 191 (2016), 15–23.
[23] Zeyu Cui, Jianxin Ma, Chang Zhou, Jingren Zhou, and Hongxia Yang. 2022. M6-rec: Generative pretrained language models are open-ended recommender systems. *arXiv preprint arXiv:2205.08084* (2022).
[24] Martina De Witte, Anouk Spruit, Susan Van Hooren, Xavier Moonen, and Geert-Jan Stams. 2020. Effects of music interventions on stress-related outcomes: a systematic review and two meta-analyses. *Health psychology review* 14, 2 (2020), 294–324.
[25] Derrell Dias, Theresa Odmark, Kierstyn Payne, and Megan Shanahan. 2018. Effects of Music Tempo on the Stress Response to Short Term Memory and Problem Solving Tests. (2018).
[26] Deniz Duman, Pedro Neto, Anastasios Mavrolampados, Petri Toiviainen, and Geoff Luck. 2022. Music we move to: Spotify audio features and reasons for listening. *Plos one* 17, 9 (2022), e0275228.

[27] Pichaya Dungsirisangthong, Treetip Boonyam, Natee Chiengchana, and Ampai Buranaprapuk. 2025. The Effect of Online Music Therapy Interventions on Reducing Stress Levels in Small Hotel Managers: A Mixed-Methods Pilot Study. *Music Therapy Perspectives* 43, 1 (2025), miae018.

[28] Fifth Edition et al. 2013. Diagnostic and statistical manual of mental disorders. *Am Psychiatric Assoc* 21, 21 (2013), 591–643.

[29] ÖNCÜ Emine, OKAY Elif, Yağmur ÇELİK, and UÇAR Kadir. 2024. Youth at Work: Exploring the Relationship Between Social Status, a History of Child Labour, and Health. *Social Science & Medicine* (2024), 117579.

[30] Anja C Feneberg, Ana Stijovic, Paul AG Forbes, Claus Lamm, Giulio Piperno, Ekaterina Pronizius, Giorgia Silani, and Urs M Nater. 2023. Perceptions of stress and mood associated with listening to music in daily life during the COVID-19 lockdown. *JAMA network open* 6, 1 (2023), e2250382–e2250382.

[31] Anja C. Feneberg, Ana Stijovic, Paul A. G. Forbes, Claus Lamm, Giulio Piperno, Ekaterina Pronizius, Giorgia Silani, and Urs M. Nater. 2023. Perceptions of Stress and Mood Associated With Listening to Music in Daily Life During the COVID-19 Lockdown. *JAMA Network Open* 6, 1 (2023), e2250382. doi:10.1001/jamanetworkopen.2022.50382

[32] Sandra Garrido and Emery Schubert. 2011. Individual differences in the enjoyment of negative emotion in music: A literature review and experiment. *Music Perception* 28, 3 (2011), 279–296. doi:10.1525/mp.2011.28.3.279

[33] Sandra Garrido and Emery Schubert. 2013. Adaptive and maladaptive attraction to negative emotions in music. *Musicae Scientiae* 17, 2 (2013), 147–166.

[34] Sandra Garrido and Emery Schubert. 2015. Music and people with tendencies to depression. *Music Perception: An Interdisciplinary Journal* 32, 4 (2015), 313–321.

[35] Lucia Gassner, Monika Geretsegger, and Julia Mayer-Ferbas. 2022. Effectiveness of music therapy for autism spectrum disorder, dementia, depression, insomnia and schizophrenia: update of systematic reviews. *European journal of public health* 32, 1 (2022), 27–34.

[36] Amira Ghenai, Zeinab Noorian, Hadiseh Moradisani, Parya Abadeh, Caroline Erentzen, and Fattane Zarrinkalam. 2025. Exploring hate speech dynamics: The emotional, linguistic, and thematic impact on social media users. *Information Processing & Management* 62, 3 (2025), 104079. doi:10.1016/j.ipm.2025.104079

[37] Shreya Ghosh and Tarique Anwar. 2021. Depression intensity estimation via social media: A deep learning approach. *IEEE Transactions on Computational Social Systems* (2021).

[38] Elizabeth Gillen, Francis Biley, and Davina Allen. 2008. Effects of music listening on adult patients' pre-procedural state anxiety in hospital. *International Journal of Evidence-Based Healthcare* 6, 1 (2008), 24–49.

[39] Joy Gorai and Dilip Kumar Shaw. 2024. A BERT-encoded ensembled CNN model for suicide risk identification in social media posts. *Neural Computing and Applications* 36, 18 (2024), 10955–10970.

[40] Salih Gulbay. 2021. Exploring the use of hip hop-based music therapy to address trauma in asylum seeker and unaccompanied minor migrant youth. In *Voices: A World Forum for Music Therapy*, Vol. 21.

[41] Chhaya Gupta, Sangeeta, Nasib Singh Gill, and Preeti Gulia. 2022. A Deep Learning-Based Framework for Analyzing Stress Factors Among Working Women. In *International Conference on Innovations in Data Analytics*. Springer.

[42] Jingjing Han, Shengnan Cong, Xiaoqing Sun, Hongyan Xie, Shiqian Ni, Yaxuan Wu, Mengyao Wang, and Aixia Zhang. 2024. The effect of music intervention on depression in perinatal women: A systematic review. *Heliyon* 10, 19 (2024).

[43] Jie Hou, Shaofei Shen, Jing Han, Siqi Xu, and Yijing Liu. 2022. Propensity Score Matching on Discrete Treatment: Beijing Pm2. 5 Case Study. In *Proceedings of the 2022 5th International Conference on Machine Vision and Applications*. 93–98.

[44] Claire Howlin and Brendan Rooney. 2021. Patients choose music with high energy, danceability, and lyrics in analgesic music listening interventions. *Psychology of Music* 49, 4 (2021), 931–944.

[45] Xu Huang, Jianxun Lian, Yuxuan Lei, Jing Yao, Defu Lian, and Xing Xie. 2023. Recommender ai agent: Integrating large language models for interactive recommendations. *ACM Transactions on Information Systems* (2023).

[46] Zhixing Huang and Natasha Duell. 2020. A novel music therapy intervention utilizing pop music to reduce adolescent anxiety and depression. *Journal of Student Research* 9, 2 (2020).

[47] Muhammad Iqbal and Rastini Olivia. 2024. Examining the Efficacy of Music Therapy in Alleviating Depression Among the Elderly. *Jurnal Ilmu Pendidikan dan Humaniora* 13, 2 (2024), 121–131.

[48] Timothy Onosahwo Iyendo, Patrick Chukwemeka Uwajeh, Bernard Nmamdi Oseke, Lei Dong, Aderemi Adejumo, Ibrahim Umar, Oberiri Destiny Apuke, and Celestine Verlumun Gever. 2024. Effect of Audio-Visual-Based Art and Music Therapy in Reducing Adolescents Post Traumatic Stress Disorder. *Journal of Adolescent Health* 74, 2 (2024), 283–291.

[49] Aryan Jadon and Avinash Patil. 2024. A comprehensive survey of evaluation techniques for recommendation systems. In *International Conference on Computation of Artificial Intelligence & Machine Learning*. Springer, 281–304.

[50] Hitkul Jangra, Rajiv Shah, and Ponnurangam Kumaraguru. 2023. Effect of feedback on drug consumption disclosures on social media. In *Proceedings of the International AAAI Conference on Web and Social Media*, Vol. 17. 435–446.

[51] Shagun Jhaver, Himanshu Rathi, and Koustuv Saha. 2024. Bystanders of Online Moderation: Examining the Effects of Witnessing Post-Removal Explanations. In *Proceedings of the CHI Conference on Human Factors in Computing Systems*. 1–9.

[52] Jing Jiang. 2024. Impact of music learning on students' psychological development with mediating role of self-efficacy and self-esteem. *PloS one* (2024).
[53] Erkang Jing, Yezheng Liu, Yidong Chai, Shuo Yu, Longshun Liu, Yuanchun Jiang, and Yang Wang. 2025. Emotion-aware Personalized Music Recommendation with a Heterogeneity-aware Deep Bayesian Network. *ACM Transactions on Information Systems* (2025).
[54] Patrik N Juslin and Daniel Västfjäll. 2008. Emotional responses to music: The need to consider underlying mechanisms. *Behavioral and Brain Sciences* (2008).
[55] Jacob Devlin Ming-Wei Chang Kenton and Lee Kristina Toutanova. 2019. Bert: Pre-training of deep bidirectional transformers for language understanding. In *Proceedings of naacL-HLT*, Vol. 1. Minneapolis, Minnesota, 2.
[56] Naba Raj Koirala, Ajita Paudel, Shikha Upadhyay, and Akshika Koirala. 2024. Music and mental health. *Journal of Psychiatrists' Association of Nepal* 13, 1 (2024), 37–43.
[57] Amanda E. Krause, James A. Dimmock, Amanda L. Rebar, and Ben Jackson. 2022. Music listening predicted improved life satisfaction in university students during early stages of the COVID-19 pandemic. *Frontiers in Psychology* 13 (2022), 820543. doi:10.3389/fpsyg.2022.820543
[58] Alex Kresovich. 2022. The influence of pop songs referencing anxiety, depression, and suicidal ideation on college students' mental health empathy, stigma, and behavioral intentions. *Health communication* 37, 5 (2022), 617–627.
[59] Ian Levy and Raphael Travis. 2020. The critical cycle of mixtape creation: Reducing stress via three different group counseling styles. *The Journal for Specialists in Group Work* 45, 4 (2020), 307–330.
[60] Huijie Lin, Jia Jia, Liqiang Nie, Guangyao Shen, and Tat-Seng Chua. 2016. What Does Social Media Say about Your Stress?.. In *IJCAI*. 3775–3781.
[61] Jianghao Lin, Xinyi Dai, Yunjia Xi, Weiwen Liu, Bo Chen, Hao Zhang, Yong Liu, Chuhan Wu, Xiangyang Li, Chenxu Zhu, et al. 2025. How can recommender systems benefit from large language models: A survey. *ACM Transactions on Information Systems* 43, 2 (2025), 1–47.
[62] Mei-Feng Lin, Ya-Ju Hsieh, Yu-Yun Hsu, Susan Fetzer, and Mei-Chi Hsu. 2011. A randomised controlled trial of the effect of music therapy and verbal relaxation on chemotherapy-induced anxiety. *Journal of clinical nursing* (2011).
[63] Zhiyuan Liu, Wei Xu, Wenping Zhang, and Qiqi Jiang. 2023. An emotion-based personalized music recommendation framework for emotion improvement. *Information Processing & Management* 60, 3 (2023), 103256.
[64] Zhiyuan Liu, Wei Xu, Wenping Zhang, and Qiqi Jiang. 2023. An Emotion-based Personalized Music Recommendation Framework for Emotion Improvement. *Information Processing & Management* 60, 3 (2023), 103256. doi:10.1016/j.ipm.2022.103256
[65] Junqi Ma. 2025. The causal effect of Internet use on rural middle-aged and older adults' depression: A propensity score matching analysis. *Digital Health* 11 (2025), 20552076241310041.
[66] Alireza Malakoutikhah, Mahlagha Dehghan, Asma Ghonchehpour, and Mohammad Ali Afshar, Peimanand Zakeri. 2023. A randomized controlled trial on the effects of different music genres on physiologic parameters and emotion. *Irish Journal of Medical Science(1971-)* (2023).
[67] Sharon Matthews. 2012. The immediate effect of musical tempo on stress, mood and self-efficacy. (2012).
[68] Katrina Skewes McFerran, Sandra Garrido, Lucy O'Grady, Denise Grocke, and Susan M Sawyer. 2015. Examining the relationship between self-reported mood management and music preferences of Australian teenagers. *Nordic Journal of Music Therapy* 24, 3 (2015), 187–203.
[69] Jarad S Miller and Donna J Terbizan. 2017. Clinical outcomes of different tempos of music during exercise in cardiac rehabilitation patients. *International Journal of Exercise Science* 10, 5 (2017), 681.
[70] Reed M Morgan and Brett Marroquín. 2024. Music listening as emotion regulation: Associations with other emotion regulation strategies and symptoms of depression and anxiety. *Musicae Scientiae* (2024).
[71] Usman Naseem, Matloob Khushi, Jinman Kim, and Adam G Dunn. 2022. Hybrid text representation for explainable suicide risk identification on social media. *IEEE transactions on computational social systems* (2022).
[72] Jukka-Pekka Onnela and Scott L. Rauch. 2016. Harnessing smartphone-based digital phenotyping to enhance behavioral and mental health. *Neuropsychopharmacology* 41, 7 (2016), 1691–1696. doi:10.1038/npp.2016.7
[73] Maria Pais, Muralidhar Pai, Asha Kamath, Leslie Lewis, and Jayaraman Nambiar. 2025. Effects of Yoga on Stress in Pregnant Women. *Holistic nursing practice* 39, 2 (2025), 86–95.
[74] Jong-In Park, In-Ho Lee, Seung-Jea Lee, Ryeo-Won Kwon, Eon-Ah Choo, Hyun-Woo Nam, and Jeong-Beom Lee. 2023. 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 medicine and therapies* (2023).
[75] Minsu Park, Chiyoung Cha, and Meeyoung Cha. 2012. Depressive moods of users portrayed in Twitter. (2012).
[76] Stefan Pasch and Jannic Cutura. 2024. Text Classification with Limited Training Data: Suicide Risk Detection on Social Media. In *2024 IEEE International Conference on Big Data (BigData)*. IEEE, 8560–8565.
[77] James W Pennebaker, Ryan L Boyd, Kayla Jordan, and Kate Blackburn. 2015. The development and psychometric properties of LIWC2015. (2015).

[78] Sabine Pieritz, Meraj Ahmad Khwaja, and Aldo A. Faisal. 2021. Personalised Recommendations in Mental Health Apps: The Impact of Autonomy and Data Sharing. *arXiv preprint arXiv:2101.08375* (2021). https://arxiv.org/abs/2101.08375
[79] Miran Pyun, Donghun Kim, Chaeyun Lim, Eunbyul Lee, Jihey Kwon, and Sangyup Lee. 2020. Examining the relationship between songs and psychological characteristics. In *HCI International 2020–Late Breaking Papers: Cognition, Learning and Games: 22nd HCI International Conference, HCII 2020, Copenhagen, Denmark, July 19–24, 2020, Proceedings 22*. Springer, 105–115.
[80] Andrew G Reece and Christopher M Danforth. 2017. Instagram photos reveal predictive markers of depression. *EPJ Data Science* 6, 1 (2017), 15.
[81] Suvi Saarikallio and Jaakko Erkkilä. 2007. The role of music in adolescents' mood regulation. *Psychology of music* 35, 1 (2007), 88–109.
[82] Koustuv Saha, Benjamin Sugar, John Torous, Bruno Abrahao, Emre Kıcıman, and Munmun De Choudhury. 2019. A social media study on the effects of psychiatric medication use. In *Proceedings of the International AAAI Conference on Web and Social Media*.
[83] Koustuv Saha, John Torous, Emre Kiciman, Munmun De Choudhury, et al. 2021. Understanding side effects of antidepressants: large-scale longitudinal study on social media data. *JMIR mental health* 8, 3 (2021), e26589.
[84] Gillian M Sandstrom and Frank A Russo. 2010. Music hath charms: The effects of valence and arousal on recovery following an acute stressor. *Music and Medicine* 2, 3 (2010), 137–143.
[85] GR Shafron and MP Karno. 2013. Heavy metal music and emotional dysphoria among listeners. Psychology of Popular Media Culture, 2 (2), 74–85.
[86] Jenny Ria Sihombing, Juliaster Marbun, Gabriella Maria Christine Sipahutar, Junita Batubara, Windos Marbun, et al. 2024. The Effect of Music Therapy on the Stress Levels of Medical Students at Universitas HKBP Nommensen. *Bulletin of Science Education* 4, 3 (2024), 245–256.
[87] Min Tan, Xinyu Zhou, Lin Shen, Yonghui Li, and Xijing Chen. 2024. Music's Dual Role in Emotion Regulation: Network Analysis of Music Use, Emotion Regulation Self-Efficacy, Alexithymia, Anxiety, and Depression. *Depression and Anxiety* 2024, 1 (2024), 1790168.
[88] Qishou Tang, Zhaohui Huang, Huan Zhou, and Peijie Ye. 2020. Effects of music therapy on depression: A meta-analysis of randomized controlled trials. *PloS one* (2020).
[89] Xiaohui Tao, Aaron Shannon-Honson, Patrick Delaney, Christopher Dann, Haoran Xie, Yan Li, and Shirley O'Neill. 2023. Towards an understanding of the engagement and emotional behaviour of MOOC students using sentiment and semantic features. *Computers and Education: Artificial Intelligence* (2023).
[90] Syauki A Thamrin and Arbee LP Chen. 2024. Detection of Bipolar Disorder on Social Media Data Utilizing Biomedical, Clinical and Mental Health Domain Fine-Tuned Word Embeddings. In *2024 IEEE 12th International Conference on Healthcare Informatics (ICHI)*. IEEE, 612–619.
[91] Joel Philip Thekkekara, Sira Yongchareon, and Veronica Liesaputra. 2024. An attention-based CNN-BiLSTM model for depression detection on social media text. *Expert systems with applications* 249 (2024), 123834.
[92] Mike Thelwall. 2017. TensiStrength: Stress and relaxation magnitude detection for social media texts. *Information Processing & Management* (2017).
[93] Filiz Ünal Toprak, Neşe Uysal, Fatih Göksel, and Yeter Soylu. 2024. The effect of music on anxiety, pain levels, and physiological parameters in women undergoing brachytherapy: a randomized controlled trial. In *Seminars in Oncology Nursing*, Vol. 40. Elsevier, 151575.
[94] Ana-Sabina Uban, Berta Chulvi, and Paolo Rosso. 2021. An emotion and cognitive based analysis of mental health disorders from social media data. *Future Generation Computer Systems* 124 (2021), 480–494.
[95] Meltem Ugurlu, Meryem Vural Şahin, and Oznur Hayat Oktem. 2024. The effect of music on menopausal symptoms, sleep quality, and depression: a randomized controlled trial. *Revista da Associação Médica Brasileira* (2024).
[96] Lee Valentine, Simon D'Alfonso, and Reeva Lederman. 2022. Recommender systems for mental health apps: advantages and ethical challenges. *AI & Society* 37 (2022), 635–646. doi:10.1007/s00146-021-01322-w
[97] Gaurav Verma, Ankur Bhardwaj, Talayeh Aledavood, Munmun De Choudhury, and Srijan Kumar. 2022. Examining the impact of sharing COVID-19 misinformation online on mental health. *Scientific Reports* 12, 1 (2022), 8045.
[98] Jonna K Vuoskoski and Tuomas Eerola. 2012. Can sad music really make you sad? Indirect measures of affective states induced by music and autobiographical memories. *Psychology of Aesthetics, Creativity, and the Arts* 6, 3 (2012), 204.
[99] Jie Wang and Kyung Hyun Suh. 2024. Music and subjective well-being in older Chinese adults: Perceived stress moderates the mediating effects of emotional adaptability. *Current Psychology* 43, 8 (2024), 7275–7283.
[100] Ling Wang, Panpan Liu, and Xin He. 2024. Personalized Music Therapy for Elderly Patients with Chronic Renal Failure to Improve their Quality of Life and Mental Health: A Retrospective Study. *Noise and Health* (2024).
[101] Qihan Wang, Anique Tahir, Zeyad Alghamdi, and Huan Liu. 2023. Exploring Musical, Lyrical, and Network Dimensions of Music Sharing Among Depression Individuals. *arXiv preprint arXiv:2310.11557* (2023).
[102] Shu Wang, Ding Austin Shijun Xu, Chonghuan, and Zhongyun Tang. 2021. A novel emotion-aware hybrid music recommendation method using deep neural network. *Electronics* (2021).

[103] Xiaodan Wang, Zheng Xie, and Guiping Du. 2024. Research on the Intervention Effect of Vibroacoustic Therapy in the Treatment of Patients with Depression. *International Journal of Mental Health Promotion* 26, 2 (2024).
[104] Yu-Ru Wang, Yan-Ye Hung, and Shiou-Fang Lu. 2024. Effect of music-based interventions on anxiety and stress-related vital signs in patients undergoing cardiac catheterization: A systematic review and meta-analysis. *Complementary therapies in medicine* (2024), 103078.
[105] Xuhai Xu, Bingsheng Yao, Yuanzhe Dong, Saadia Gabriel, Hong Yu, James Hendler, Marzyeh Ghassemi, Anind K Dey, and Dakuo Wang. 2024. Mental-llm: Leveraging large language models for mental health prediction via online text data. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* (2024).
[106] Masahira Yamamoto, Shinobu Naga, and Jun Shimizu. 2007. Positive musical effects on two types of negative stressful conditions. *Psychology of Music* 35, 2 (2007), 249–275.
[107] Fan Yang, Zheng Chen, Ziyan Jiang, Eunah Cho, Xiaojiang Huang, and Yanbin Lu. 2023. Palr: Personalization aware llms for recommendation. *arXiv preprint arXiv:2305.07622* (2023).
[108] Dilek Yildirim, Cennet Çiriş Yildiz, Ferda Akyuz Ozdemir, Merve Harman Özdoğan, and Gulbeyaz Can. 2024. Effects of a mindfulness-based stress reduction program on stress, depression, and psychological well-being in patients with cancer: a single-blinded randomized controlled trial. *Cancer Nursing* 47, 2 (2024), E84–E92.
[109] Dongsong Zhang, Lina Zhou, Jie Tao, Tingshao Zhu, and Guodong Gao. 2024. KETCH: a knowledge-enhanced transformer-based approach to suicidal ideation detection from social media content. *Information Systems Research* (2024).
[110] Muni Zhuang, Dongsheng Cheng, Xin Lu, and Xu Tan. 2024. Postgraduate psychological stress detection from social media using BERT-Fused model. *PloS one* (2024).