Journal of Neurodevelopmental Cognition

Journal of Neurodevelopmental Cognition 6(2) (2025) 40- 64 10.48308/jncog.2025.240607.1019

ISSN: 2645-565X

http://www.jncog.sbu.ac.ir

# Role of Acoustic Features of Music on Emotion Processing of Images

Shayan Mousaee[1], Reza Khosrowabadi[2*]

[1] *Institute of Medical Science and Technology, Shahid Beheshti University, Tehran, Iran*

[2] *Institute for Cognitive and Brain Science, Shahid Beheshti University, Tehran, Iran*

*Communicated by Hamidreza Mahdiani*

**Abstract**

Music significantly influences emotional and cognitive functions, and its direct impact on emotion processing is well-documented. The proliferation of music on social media platforms highlights the need to consider music as a context during the processing of other stimuli. This aimed to understand how specific acoustic features of widely consumed music influence human emotion processing. Manipulated acoustic cues—pitch, tempo, and timbre—in popular instrumental music from social media were used to influence perceived emotion of image stimuli. An innovative online experimental platform was developed and used to collect data from 109 participants (aged 15–62 years). Participants were exposed to music excerpts in which one of the three acoustic components was altered (high, low, or normal state) while viewing images from the International Affective Picture System and rated their emotional responses using the Self-Assessment Manikin questionnaire. Statistical analysis revealed that tempo was the strongest acoustic predictor of emotional arousal, with distinct patterns across age and gender. Increased tempo was also associated with higher valence ratings. Pitch manipulations had a stronger impact on the arousal levels of older participants, while timbre primarily influenced valence perception in younger listeners. Significant gender differences were also observed: male participants were more sensitive to changes in pitch and timbre, while female participants were more responsive to tempo variations. These findings contribute to a deeper understanding of how specific acoustic characteristics of contemporary, widely disseminated music shape emotional perception and highlight the complex interplay of acoustic cues with individual differences such as age and gender.

**Keywords**: Emotional Perception, Pitch, Tempo, Timbre, Arousal and Valence

* Corresponding author

*Email addresses:* r_khosroabadi@sbu.ac.ir



*Received: April 2025 Accept: May 2025*

## 1. Introduction

Music's capacity to evoke and modulate emotion is a well-recognized phenomenon, with its pervasive influence now significantly amplified in the digital age through social media platforms like YouTube, TikTok, and Instagram. These platforms have reshaped music consumption into rapid, dynamic, and socially embedded experiences, where viral music tracks, gaining swift popularity, increasingly orchestrate emotional engagement within these digital ecosystems. The profound impact of auditory stimuli on emotional cognition, memory, mood, and even social interaction and identity formation is extensively documented (Gabrielsson & Lindström, 2001; Hargreaves & North, 1999; Juslin & Sloboda, 2010; Sloboda & Juslin, 2001). To quantify and understand emotional responses to music, the dimensional model of emotion is widely employed in research. This model conceptualizes emotions along continuous axes, with arousal and valence being the most prominent dimensions (Bradley & Lang, 1994; McAdams et al., 2017; Nineuil et al., 2020; Russell, 1980; Zaatar et al., 2024).

**Arousal** refers to the intensity of the emotional experience, ranging from calm or sleepy to excited or agitated. **Valence** describes the pleasantness or unpleasantness of the emotion, ranging from positive (e.g., happy, joyful) to negative (e.g., sad, angry). This two-dimensional framework provides a robust and parsimonious way to capture the core affective quality of music-induced emotions and is particularly useful for comparing responses across different stimuli and individuals. While the valence-arousal model is foundational, it is acknowledged that the full spectrum of music-evoked emotions can be richer, encompassing more discrete states like joy, sadness, fear, or even complex emotions like nostalgia and wonder (Eerola & Saari, 2025). Nevertheless, the valence-arousal model offers a standardized and widely validated approach for assessing fundamental emotional reactions, which is adopted in the present study.

The emotional impact of music is conveyed through the manipulation and interaction of various acoustic cues. Among these, specific acoustic features such as pitch, tempo, and timbre are pivotal in shaping the perceived emotional landscape of a musical piece (Pring et al., 2024). **Pitch** encompasses both the perceived highness or lowness of a sound (pitch height) and its organization into scales and modes (e.g., major and minor in Western music). **Tempo**, or the speed of music, is perhaps one of the most powerful cues for emotional arousal. **Timbre**, often described as the "color" or quality of a sound, distinguishes different instruments or voices even when they produce the same pitch and loudness. Faster tempos, higher pitches, and brighter timbres typically evoke happiness, aligning with higher arousal and more positive valence perceptions (Armitage et al., 2024; Pring et al., 2024; Wang et al., 2021; Zentner et al., 2008). For instance, rapid tempi can heighten alertness and engagement, a principle leveraged in advertising and sports (Brattico et al., 2016; North et al., 1999), potentially through physiological entrainment where the nervous system synchronizes with external rhythms, while slow tempi are often associated with low arousal states like calmness, sadness, or tenderness. Similarly, high-pitched melodies are often perceived as energetic and joyful, while lower pitches may convey sadness or gravity (Zentner et al., 2008). Timbre refines these affective judgments, enabling listeners to discern emotions like sadness or happiness based on an instrument's tonal quality alone, engaging both auditory and limbic systems (Hailstone et al., 2009).

However, much of the existing evidence stems from studies using controlled laboratory music excerpts. Consequently, there is a lack of understanding of how subtle modifications to these acoustic features within the context of real-world social media music content—often characterized by brevity, repetitiveness, and emotional priming—impact listeners' emotional experiences, particularly across diverse age groups and genders. This question is increasingly pertinent as immersive short-form videos accompanied by background music are widely consumed by a broad demographic, potentially influencing emotional and cognitive development in both youth and adults. Theoretical frameworks, such as the multicomponent model proposed by (Juslin & Västfjäll, 2008), suggest that music-induced emotions arise from various mechanisms, including brain stem reflexes, emotional contagion, and episodic memory, with emotions themselves being constructed experiences shaped by cognitive schemas and context (Barrett, 2017). The interplay of these processes with acoustic cues allows listeners to imbue abstract sound sequences with personal meaning.

Developmental factors and individual listener characteristics are known to moderate the emotional processing of music. Recent research suggests that age can influence emotional ratings of music; for instance, older adults tend to provide more positive valence ratings overall and may respond differently to tempo and mood cues compared to younger listeners (Hofbauer & Rodriguez, 2023; Stephenson et al., 2016). This heightened sensitivity in older adults could be due to more established exposure to genre-specific tonal patterns (Ferrer et al., 2013; Juslin & Västfjäll, 2008). Similarly, gender differences in music-evoked emotions have been reported, though findings are not always consistent (Fuentes-Sánchez et al., 2025). It has been suggested that women and men might attend to distinct musical cues or differ in arousal reactivity, with some research indicating women may be more responsive to rhythmic and tempo variations, while men might show greater sensitivity to tonal and timbral attributes (Fuentes-Sánchez et al., 2025). Such developmental sensitivities are observed early, with studies showing children can manipulate pitch and tempo to express emotion and reliably match musical cues to perceived emotional content (Adachi & Trehub, 1998; Dalla Bella et al., 2001). Furthermore, cross-cultural studies suggest a degree of biological grounding in human responses to musical elements (Egermann et al., 2015), and neuroimaging research highlights music's activation of brain areas involved in emotional regulation and autobiographical memory (Ferreri et al., 2015).

To address the existing gaps concerning contemporary media, the present study focused on how controlled changes in pitch (musical key), tempo (speed), and timbre (tone color) within popular social-media music clips influence two core emotion dimensions – valence (pleasure) and arousal (intensity) – as perceived by listeners. We further explored whether these effects differ by age group and gender, considering potential neurodevelopmental and psychosocial distinctions in auditory emotion processing. For this research, we developed a custom online platform to deliver an interactive, multimedia listening experience and to facilitate data gathering and interpretation within a cognitive task design. A total of 109 participants, spanning adolescence to late adulthood (ages 15–62), were recruited, with an effort to ensure roughly equal representation across five age brackets. Each participant engaged with emotionally evocative music stimuli derived from viral social media posts. These stimuli were manipulated in either pitch, tempo, or timbre, while keeping the core musical content constant. Participants' emotional responses were recorded in real-time using Self-Assessment Manikin (SAM) scales for valence and arousal (Lang, 1980), synchronized with International Affective Picture System (IAPS) images to enhance immersion and ecological validity.

In the current study, we hypothesized that (a) tempo changes would most strongly affect arousal, given tempo's established link to energy and activation (Droit-Volet & Meck, 2007); (b) timbre changes would primarily influence valence, as timbral brightness can convey pleasantness (Pring et al., 2024); and (c) pitch shifts might have subtler effects, potentially varying with listener age (Hofbauer & Rodriguez, 2023). Furthermore, we anticipated distinct sensitivity patterns by gender, based on prior indications of sex-specific emotional processing of music (Fuentes-Sánchez et al., 2025). The following brief communication presents our methodology and key findings, which aim to reveal differential impacts of these auditory features on emotion perception across various demographic groups, contributing to a more nuanced understanding of music's emotional power in the context of pervasive social media engagement.

## 2. Method

### Participants

A total of 109 participants (64 female, 43 male, 2 other; Mean = 27.16 years, SD = 10.23, range = 15-62 years) successfully completed the online experiment. Participants were recruited through various online channels, directing them to the custom-built experimental platform[1]. Demographic information is summarized in Table 1. No participant reported hearing impairments or neurological conditions. Each was compensated with a small honorarium or course credit.

Before the main experiment, participants underwent a screening process. This included providing demographic details and completing the Depression, Anxiety, and Stress Scale-21 Items (DASS-21) on a third-party website[2]. The DASS-21 is a self-report instrument designed to measure the severity of a range of symptoms common to depression, anxiety, and stress. Participants were eligible to continue only if their scores on all three subscales (Depression, Anxiety, Stress) fell within the "normal" to "moderate" range, as defined by Lovibond and Lovibond (1995). This was implemented to minimize the influence of pre-existing severe negative affective states on the emotional ratings of the musical stimuli. The age of participants was also restricted to individuals between 15 and 65 years. All participants provided digital informed consent prior to the start of the study. The research was conducted under the ethical standards set by Shahid Beheshti University.

**Table 1** *Participant Demographics (N=109)*

[1] The custom online experiment platform (developed by the authors) is accessible at survey.mousaee.ir, demonstrating the study interface and stimulus delivery system.

[2] https://www.teblink.com/assessment/DASS-21.cshtml

| Characteristic | Category | N | Percentage |
|---|---|---|---|
| **Age Group (Years)** | 15–21 | 26 | 23.9% |
| | 22–25 | 37 | 33.9% |
| | 26–30 | 22 | 20.2% |
| | 31–40 | 15 | 13.8% |
| | 41+ | 9 | 8.3% |
| **Gender** | Female | 64 | 58.7% |
| | Male | 43 | 39.4% |
| | Other | 2 | 1.8% |
| **Education Level** | Below Diploma | 6 | 5.5% |
| | Diploma | 19 | 17.4% |
| | Associate Degree | 7 | 6.4% |
| | Bachelor's Degree | 39 | 35.8% |
| | Master's Degree | 37 | 33.9% |
| | Doctorate | 1 | 0.9% |
| **Dominant Hand** | Right-handed | 96 | 88.1% |
| | Left-handed | 13 | 11.9% |

The use of DASS-21 for pre-screening is a notable feature of the methodology, aiming to enhance the internal validity of the emotional assessments by controlling for potentially confounding baseline mood states, a common challenge in online emotion research (Egermann et al., 2006).

**Materials and Stimuli**

**Musical Stimuli:** Six instrumental music pieces, reported as being popular on social media platforms, were selected for the study. For each of the three target acoustic components (pitch, tempo, timbre), two distinct musical pieces were chosen. These pieces were then manipulated to create three versions for each original piece: a "normal" (original) version, a "low" version (e.g. lower pitch, slower tempo, or softer timbre), and a "high" version (higher pitch, faster tempo, or brighter timbre) of the targeted acoustic component. For each audio component, one was altered while the others remained constant: for pitch, the musical key of the excerpt was transposed up and down by ±200 cents—a unit of measurement in music that describes the pitch difference between two notes (approximately one semitone)—without otherwise altering the melody or harmonic structure; for tempo, the playback speed (beats per minute) was increased or decreased by about 10–13 BPM from the original tempo, a subtle change chosen to be noticeable while preserving musical rhythm and style; for timbre, the tonal balance was modified by equalization – specifically, boosting versus cutting a mid-high frequency band (around 1–6 kHz) by ~8 dB – to produce a brighter, harsher tone in one version and a warmer, smoother tone in the other. The specific manipulations were performed using FL Studio software version 21, as detailed in Table 2, and were designed to be clearly perceptible to individuals with normal hearing while having the potential to elicit differential emotional responses. The choice of popular, and thus potentially familiar, instrumental music aimed to increase ecological validity, reflecting common music listening experiences on social media, while the controlled manipulation allowed for systematic investigation of individual acoustic cues.

**Table 2** *Musical Stimuli and Acoustic Manipulations*

| Acoustic Component | Original Music Piece | State | Modification Details |
|---|---|---|---|
| **Pitch** | Yiruma – River Flows in You | Normal | A Major |
| | | Low | G Major (200 cents decrease) |
| | | High | B Major (200 cents increase) |
| | Pachelbel – Canon in D (Piano) | Normal | D Major |
| | | Low | C Major (200 cents decrease) |
| | | High | E Major (200 cents increase) |
| **Tempo** | Merry Go Round of Life – Howl's Moving Castle | Normal | 82 BPM |
| | | Low | 69 BPM (13 BPM decrease) |
| | | High | 95 BPM (13 BPM increase) |
| | The Verve – Bitter Sweet Symphony (Instrumental) | Normal | 85 BPM |
| | | Low | 75 BPM (10 BPM decrease) |
| | | High | 95 BPM (10 BPM increase) |

| Acoustic Component | Original Music Piece | State | Modification Details |
|---|---|---|---|
| **Timbre** | Joe Hisaishi – Summer | Normal | No change |
| | | Low | 2 kHz frequency range, 8 dB amplitude reduction |
| | | High | 2 kHz frequency range, 8 dB amplitude increase |
| | T'Mia & Sebastian's Theme (Piano) | Normal | No change |
| | | Low | 1–6 kHz frequency range, 8 dB amplitude reduction |
| | | High | 1–6 kHz frequency range, 8 dB amplitude increase |

**Visual Stimuli:** Simultaneously with the musical stimuli, participants viewed 30 specific images selected from the International Affective Picture System (IAPS) database. The IAPS is a standardized set of color photographs with normative ratings for valence, arousal, and dominance, widely used in emotion research to elicit affective responses. In each phase, an IAPS image was displayed full-screen for 4 seconds concurrently with the playback of a music excerpt. The inclusion of IAPS images, while primarily focusing on the music's emotional impact, provides a controlled visual context during the emotional rating task.

**Figure 1** *One of the 30 IAPS images used as the visual stimuli*

**Experimental Platform and Procedure**

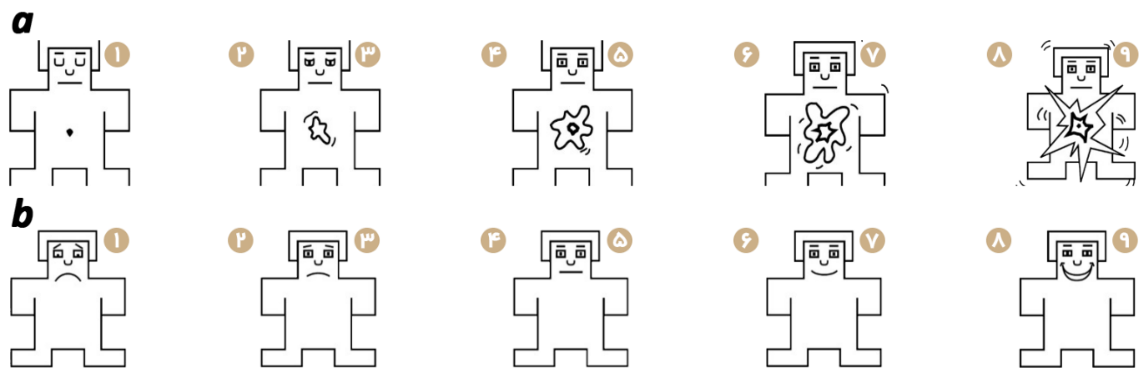
The experiment was conducted **entirely online** using a custom-developed web-based platform hosted at survey.mousaee.ir. The platform was built using HTML, CSS, PHP, and JavaScript. Key features included:

- **Participant Verification:** A phone number and SMS-based code verification system to ensure unique participant entries.
- **Initial Screening:** Collection of demographic data and DASS-21 scores.
- **Auditory Calibration:** A brief hearing and device volume calibration test, where participants listened to an audio phrase and typed it into a text box to ensure they could adequately perceive the auditory stimuli.
- **Designing cognitive tasks:** A platform for designing cognitive tasks, performing them, and collecting real-time data.

Participants were randomly assigned to one of the three acoustic component conditions (pitch, tempo, or timbre) by the platform. Within their assigned condition, they were randomly presented with one of the two unique music pieces associated with that component. The experiment consisted of three phases for each participant, corresponding to the "low," "normal," and "high" states of the manipulated acoustic component. The order of these three phases was randomized for each participant to counterbalance potential order effects.

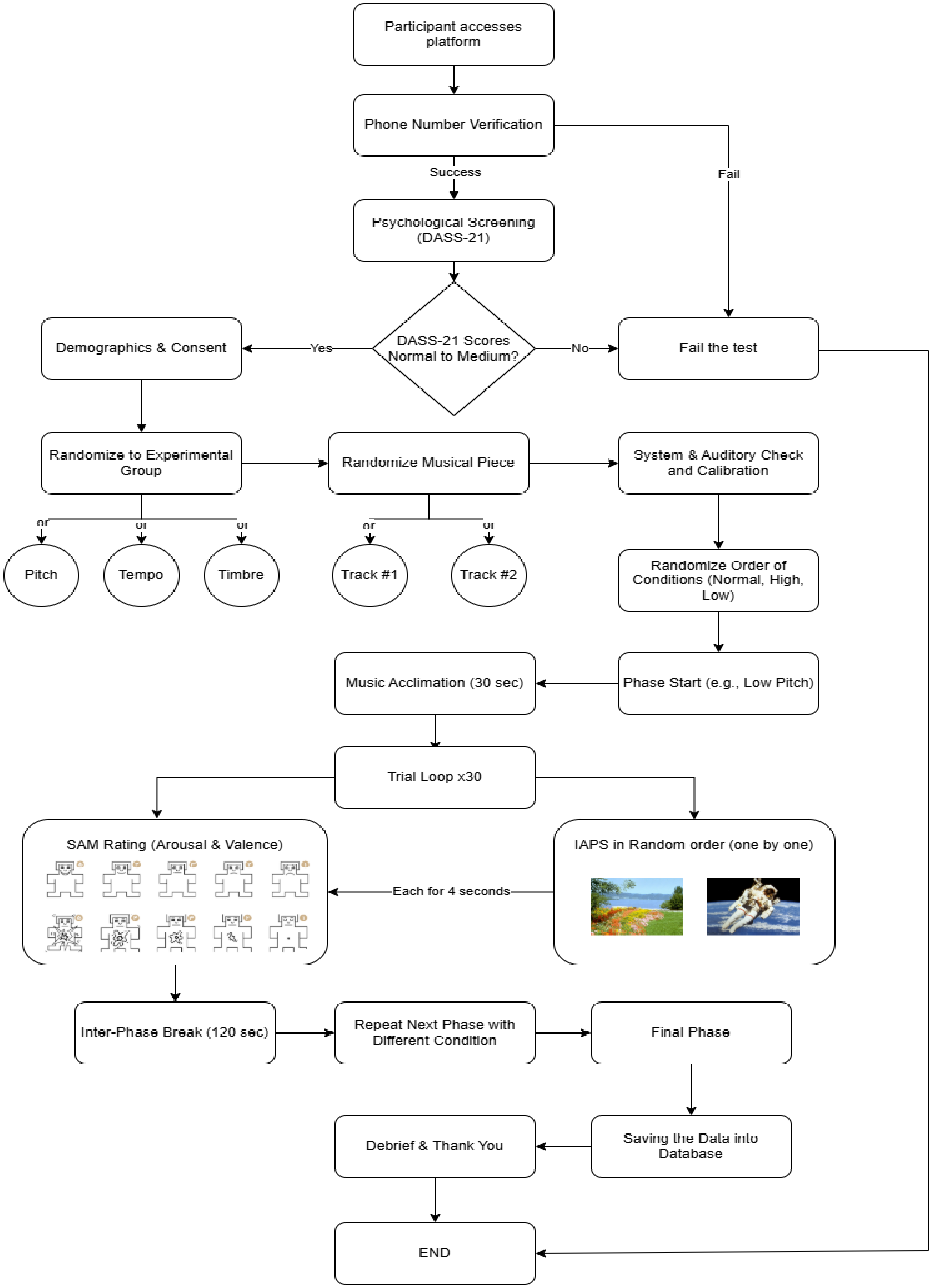
In each trial within a phase, 30 IAPS images were displayed one by one, each for 4 seconds, while the corresponding manipulated music piece played. Immediately following the image-music presentation, participants rated their current emotional state on two dimensions: valence (pleasantness/unpleasantness) and arousal (intensity/activation), using the Self-Assessment Manikin (SAM) scale. The SAM is a non-verbal pictorial scale where participants select one of nine figures representing varying degrees of valence and arousal. This scale is well-suited for online studies due to its simplicity and reduced reliance on language.

**Figure 2** ***a)*** *Arousal Mannikins in the platform's SAM scale (the left one has the lowest arousal value, and the right one the highest arousal value),* ***b)*** *Valence Mannikins in the platform's SAM scale (the left one has the lowest valence value, and the right one the highest valence value)*




Before each of the three main experimental phases, participants listened to the music (with the specific manipulation for that phase) for 30 seconds to acclimate to the sound. A 120-second silent break was enforced between phases to minimize emotional carry-over effects. To ensure continued engagement and focused attention, a 60-second time limit was imposed for responding after each IAPS image presentation; exceeding this limit would result in preventing the participant from continuing the task and asking them to restart the experiment. The total duration for completing all three phases of the experiment was approximately 20 to 30 minutes. The platform's design, incorporating these screening, calibration, and procedural controls, aimed to enhance data quality and methodological rigor within the online experimental context.

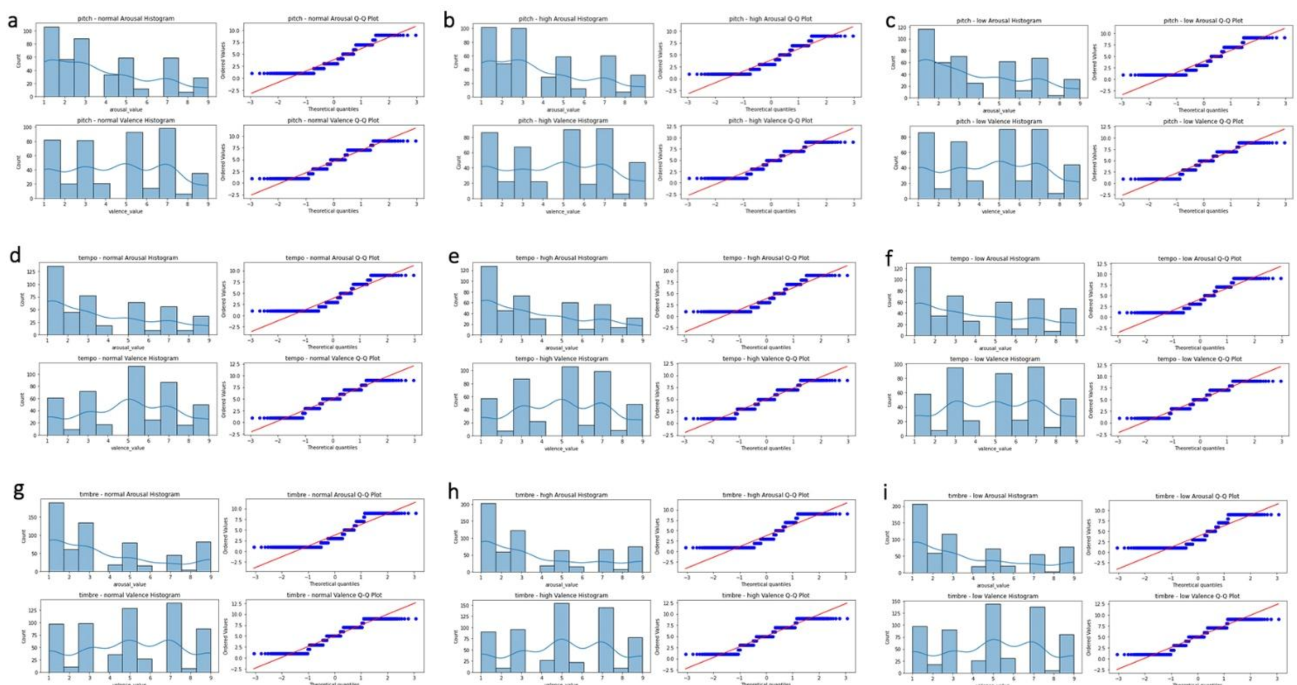
**Figure 3** *Schematic overview of the experimental task design.*



**Data Analysis**

The collected data on valence and arousal ratings were analyzed to determine the impact of the manipulated acoustic components. Initial checks for normality of the data using the Shapiro-Wilk test, along with visual inspection of histograms and Q-Q plots (Figure 4), indicated that the data were not normally distributed across the various conditions. Consequently, non-parametric statistical tests were employed for the main analyses.

**Figure 4** *Histogram (left) and Q-Q plots (right) for arousal and valence distribution for all audio components and their variations (a top: Normal Pitch Arousal, a bottom: Normal Pitch Valence, b top: High Pitch Arousal, b bottom: High Pitch Valence, c top: Low Pitch Arousal, c bottom: Low Pitch Valence, d top: Normal Tempo Arousal, d bottom: Normal Tempo Valence, e top: High Tempo Arousal, e bottom: High Tempo Valence, f top: Low Tempo Arousal, f bottom: Low Tempo Valence, g top: Normal Timbre Arousal, g bottom: Normal Timbre Valence, h top: High Timbre Arousal, h bottom: High Timbre Valence, I top: Low Timbre Arousal, I bottom: Low Timbre Valence)*



The **Friedman test** (non-parametric repeated-measures ANOVA) was used to assess whether there were statistically significant differences in median arousal and valence scores across the three states (low, normal, high) of each manipulated acoustic component (pitch, tempo, and timbre). This test is appropriate for related samples when the assumption of normality is violated. Friedman's chi-square statistic (with $df = 2$) and associated p-value were obtained for each test. Where the Friedman test indicated a significant overall effect, post-hoc pairwise comparisons were conducted using the **Wilcoxon signed-rank test** to identify which specific conditions differed significantly from each other (Low vs Normal, Low vs High, and Normal vs High). A **Bonferroni correction** was applied to the p-values obtained from the Wilcoxon tests to control for the increased risk of Type I errors due to multiple comparisons ($\alpha = \frac{0.05}{3}$ per comparison).

For all significant effects, appropriate effect sizes were planned for reporting to indicate the magnitude of the observed differences. For the Friedman test, **Kendall's W** is a suitable effect size, calculated as $W = \frac{X^2}{N\times(k-1)}$, where $X^2$ is the Friedman test statistic, $N$ is the sample size, and $k$ is the number of related groups (conditions). Kendall's W ranges from 0 (no agreement) to 1 (perfect agreement), with interpretation guidelines often suggesting 0.1–<0.3 as small, 0.3–<0.5 as moderate, and $\geq 0.5$ as a large effect. For the Wilcoxon signed-rank test, the **rank-biserial correlation (r)** is a common effect size, calculated as $r = Z/\sqrt{N_{pairs}}$, where $Z$ is the standardized test statistic and $N_{pairs}$ is the number of pairs. The interpretation guidelines for | $r$ | are typically 0.1 for a small effect, 0.3 for a medium effect, and 0.5 for a large effect.

Analyses were conducted for **the overall dataset** and then separately for subgroups based on **gender** and **age** to explore moderating effects. All statistical analyses were performed with a significance level ($\alpha$) set at.05.

## 3. Results

This section details the statistical outcomes of the study, examining the impact of manipulated acoustic components (pitch, tempo, and timbre) on perceived emotional arousal and valence. Analyses were conducted on the full dataset and subsequently for demographic subgroups based on gender and age.

**Participant Characteristics**

The final dataset consisted of 109 participants who successfully completed all phases of the online experiment. Demographic details, including age group distribution, gender, education level, and dominant hand, are summarized in Table 1. All participants included in the analysis passed the DASS-21 screening, indicating normal to moderate levels of depression, anxiety, and stress.

**Overall Impact of Acoustic Components on Arousal and Valence**

The Friedman test was employed to assess the overall effect of each acoustic component (pitch, tempo, and timbre) across its three manipulated states (low, normal, high) on perceived arousal and valence for the entire sample. The results are presented in Table 3.

Tempo is a significant factor influencing both arousal ($X^2(2) = 20.27, p < .001, Kendall's\, W = 0.27$) and, marginally, valence ($X^2(2) = 5.77, p = .056$). Timbre also showed a significant overall effect on perceived valence ($X^2(2) = 7.90, p = .019, Kendall's\, W = 0.11$). Pitch manipulations did not have a significant overall effect on either arousal ($X^2(2) = 3.74, p = .154$) or valence ($X^2(2) = 0.26, p = .876$) for the entire dataset.

**Table 3** *Friedman Test Results for the Overall Effects of Acoustic Components on Perceived Arousal and Valence (Ns per component: Pitch=35, Tempo=37, Timbre=37)*

| Acoustic Component | Emotional Dimension | Friedman $\chi^2$ | df | p-value |
|---|---|---|---|---|
| **Pitch** | Arousal | 3.74 | 2 | .154 |
| | Valence | 0.26 | 2 | .876 |

| Acoustic Component | Emotional Dimension | Friedman $\chi^2$ | df | p-value |
|---|---|---|---|---|
| **Tempo** | Arousal | 20.27 | 2 | **<.001** |
| | Valence | 5.77 | 2 | .056† |
| **Timbre** | Arousal | 1.75 | 2 | .417 |
| | Valence | 7.90 | 2 | **.019** |

*†: Marginally significant. Bold values indicate $p < .05$.*

**Specific Effects of Tempo Manipulations**

Given the significant overall effect of tempo on arousal and its marginal effect on valence, Wilcoxon signed-rank tests with Bonferroni correction were conducted for pairwise comparisons between the tempo conditions (normal, low, high). The results are detailed in Table 4.

For **arousal**, perceived arousal was significantly higher in the **low tempo** condition compared to both the **normal tempo** condition ($Z \approx -5.36, p_{corr} < .001, r = -.88$) and the **high tempo** condition ($Z = -4.40, p_{corr} < .001, r = -.72$). There was no significant difference in arousal between the normal and high tempo conditions ($Z \approx -0.66, p_{corr} = 1.000$). These findings, illustrated in Figure 5A, indicate that decreasing the tempo of viral music significantly increased participants' perceived arousal.

For **valence**, perceived valence was significantly higher in the **high tempo** condition compared to the **normal tempo** condition ($Z \approx -2.40, p_{corr} = .016, r = -.39$). No other pairwise comparisons for valence reached statistical significance after Bonferroni correction (Normal vs. Low: $Z = -0.86, p_{corr} = 1.000$; High vs. Low: $Z = -1.16, p_{corr} = 1.000$). This suggests that increasing the tempo of these music pieces led to them being perceived as more pleasant (Figure 5B).

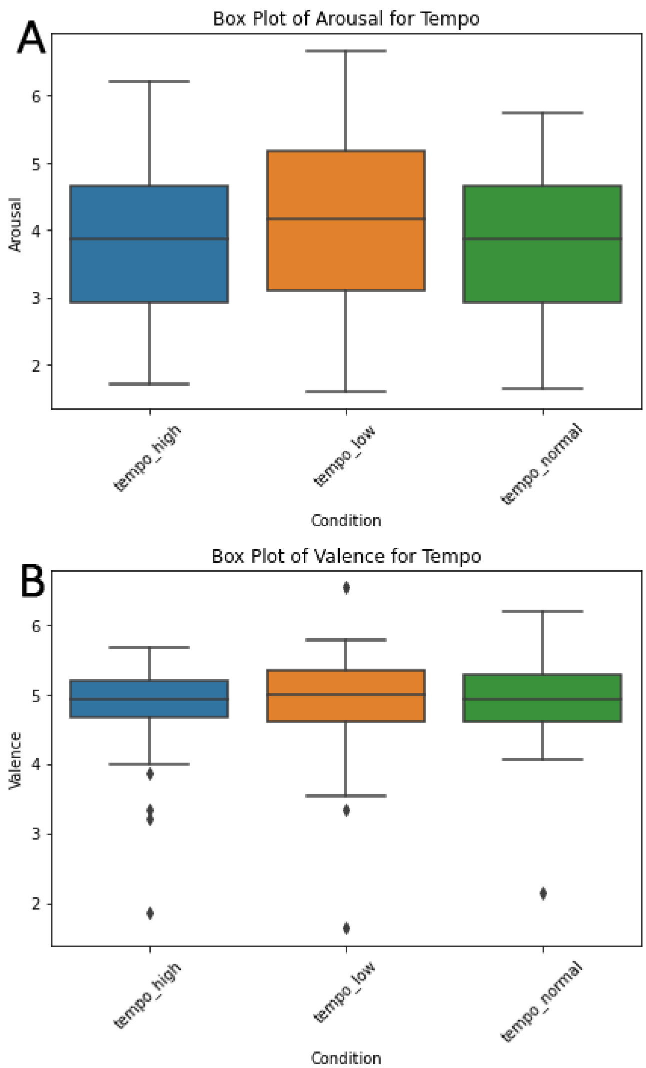
**Table 4** *Wilcoxon Signed-Rank Test Results for Pairwise Comparisons of Tempo Conditions on Arousal and Valence (N=37)*

| Comparison | Dimension | Wilcoxon Statistic (W) | Z-value (approx.) | $p_{uncorr}$ | $p_{corr}$ (Bonferroni) |
|---|---|---|---|---|---|
| Normal vs. High Tempo | Arousal | 93,968.00 | -0.66 | .511 | 1.000 |
| | Valence | 56,530.50 | -2.40 | .005 | **.016** |
| Normal vs. Low Tempo | Arousal | 71,883.50 | -5.36 | <.001 | **<.001** |
| | Valence | 66,350.50 | -0.86 | .388 | 1.000 |
| High vs. Low Tempo | Arousal | 72,653.50 | -4.40 | <.001 | **<.001** |

| Comparison | Dimension | Wilcoxon Statistic (W) | Z-value (approx.) | $p_{uncorr}$ | $p_{corr}$ (Bonferroni) |
|---|---|---|---|---|---|
| | Valence | 58,002.50 | -1.16 | .247 | 1.000 |

*Note: Z-values are approximated based on typical Python statistical analysis output, where W is the sum of ranks for the smaller group, and Z is provided for significance testing.* $p_{corr}$ *is the Bonferroni-corrected p-value. Bold values indicate* $p_{corr} < .05$.

**Figure 5** *Box plots of (A) Arousal and (B) Valence ratings for different Tempo conditions. The central line represents the median, the box edges represent the interquartile range (IQR), and whiskers extend to 1.5 times the IQR.*




**Specific Effects of Pitch Manipulations**

The Friedman test did not reveal a significant overall effect of pitch manipulations on either arousal or valence for the entire sample (see Table 3). Therefore, post-hoc pairwise comparisons for the overall dataset were not conducted for pitch.

**Specific Effects of Timbre Manipulations**

The Friedman test indicated a significant overall effect of timbre manipulations on perceived valence ($X^2(2) = 7.90, p = .019$) but not on arousal (see Table 3). Subsequent Wilcoxon signed-rank tests with Bonferroni correction for pairwise comparisons of timbre conditions (normal, low, high) on valence are presented in Table 5.

After Bonferroni correction, none of the pairwise comparisons for timbre on valence reached statistical significance (Normal vs. High: $Z = -1.66,\ p_{corr} = .290$; Normal vs. Low: $Z = -0.34$, $p_{corr} = 1.000$; High vs. Low: $Z = -2.11,\ p_{corr} = .105$). This suggests that while there was an overall trend, the specific differences between timbre conditions were not robust enough to remain significant after correcting for multiple comparisons in the full dataset.

**Table 5** *Wilcoxon Signed-Rank Test Results for Pairwise Comparisons of Timbre Conditions on Valence (N=37)*

| Comparison | Dimension | Wilcoxon Statistic (W) | Z-value (approx.) | $p_{uncorr}$ | $p_{corr}$ (Bonferroni) |
|---|---|---|---|---|---|
| Normal vs. High Timbre | Valence | 38,798.00 | -1.66 | .097 | .290 |
| Normal vs. Low Timbre | Valence | 39,927.50 | -0.34 | .735 | 1.000 |
| High vs. Low Timbre | Valence | 44,037.50 | -2.11 | .035 | .105 |

**Moderating Effects of Gender**

To investigate gender differences, Friedman tests were conducted separately for male and female participants as well. Results are summarized in Table 6.

**For male participants (n=43):**

- **Pitch:** Significantly influenced both arousal ($X^2(2) = 8.03, p = .018, Kendall's\ W = 0.29$) and valence ($X^2(2) = 6.06, p = .048, Kendall's\ W = 0.22$).
- **Tempo:** Significantly influenced arousal $X^2(2) = 8.63, p = .013, Kendall's\ W = 0.29$), but not valence ($X^2(2) = 5.85, p = .054$).
- **Timbre:** Significantly influenced valence ($X^2(2) = 10.02, p = .007, Kendall's\ W = 0.36$), but not arousal ($X^2(2) = 2.90, p = .234$).

**For female participants (n=64):**

- **Tempo:** Significantly influenced both arousal ($X^2(2) = 11.69, p = .003, Kendall's\ W = 0.27$) and valence ($X^2(2) = 6.05, p = .049, Kendall's\ W = 0.14$).
- **Pitch and Timbre:** Did not show significant overall effects on either arousal or valence for female participants.

**Table 6** *Friedman Test Results for Effects of Acoustic Components by Gender: Male (n per component group: Pitch ≈14, Tempo ≈15, Timbre ≈14), Female (n per component group: Pitch ≈21, Tempo ≈22, Timbre ≈21)*

| Gender | Acoustic Component | Emotional Dimension | Friedman $\chi^2$ | df | p-value |
|---|---|---|---|---|---|
| **Male** | Pitch | Arousal | 8.03 | 2 | **.018** |
| | | Valence | 6.06 | 2 | **.048** |
| | Tempo | Arousal | 8.63 | 2 | **.013** |
| | | Valence | 5.85 | 2 | .054† |
| | Timbre | Arousal | 2.90 | 2 | .234 |
| | | Valence | 10.02 | 2 | **.007** |
| **Female** | Pitch | Arousal | 0.93 | 2 | .629 |
| | | Valence | 4.57 | 2 | .102 |
| | Tempo | Arousal | 11.69 | 2 | **.003** |
| | | Valence | 6.05 | 2 | **.049** |
| | Timbre | Arousal | 2.13 | 2 | .344 |
| | | Valence | 3.07 | 2 | .215 |

Significant Wilcoxon signed-rank test results (Bonferroni-corrected) for gender subgroups are presented in Table 7. For **male group**, low tempo significantly increased arousal compared to normal tempo ($Z \approx -3.15, p_{corr} = .001, r = -.81$) and high tempo ($Z \approx -2.86, p_{corr} = .004, r = -.74$). High tempo significantly increased valence compared to normal tempo ($Z \approx -2.81, p_{corr} = .025, r = -.73$) and low tempo ($Z \approx -2.97, p_{corr} = .009, r = -.77$). For timbre, high timbre was rated as significantly more pleasant than low timbre ($Z \approx -2.84, p_{corr} = .023, r = -.76$). Pitch manipulations did not yield significant pairwise differences for males after correction.

For **female group**, low tempo significantly increased arousal compared to normal tempo ($Z \approx -4.21, p_{corr} < .001, r = -.90$) and high tempo ($Z \approx -3.57, p_{corr} = .001, r = -.76$). No significant pairwise differences were found for valence in females regarding tempo, nor for pitch or timbre on either dimension.

**Table 7** *Summary of Significant Wilcoxon Test Results by Gender (Corrected p-values)*

| Gender | Component | Dimension | Comparison | Z-value (approx.) | $p_{corr}$ | Direction of Effect |
|---|---|---|---|---|---|---|
| **Male** | Tempo | Arousal | Normal vs. Low | -3.15 | **.001** | Low Tempo > Normal Tempo Arousal |
| | Tempo | Arousal | High vs. Low | -2.86 | **.004** | Low Tempo > High Tempo Arousal |
| | Tempo | Valence | Normal vs. High | -2.81 | **.025** | High Tempo > Normal Tempo Valence |
| | Tempo | Valence | High vs. Low | -2.97 | **.009** | High Tempo > Low Tempo Valence |
| | Timbre | Valence | High vs. Low | -2.84 | **.023** | High Timbre > Low Timbre Valence |
| **Female** | Tempo | Arousal | Normal vs. Low | -4.21 | **<.001** | Low Tempo > Normal Tempo Arousal |
| | Tempo | Arousal | High vs. Low | -3.57 | **.001** | Low Tempo > High Tempo Arousal |

**Moderating Effects of Age**

Friedman tests were also conducted for five distinct age groups (15-21, 22-25, 26-30, 31-40, 41+ years). The results showed varying patterns of sensitivity:

- **Group 1 (15-21 years):** Pitch significantly affected arousal ($X^2(2) = 8.75, p = .013, Kendall's\ W = 0.17$). Tempo significantly affected arousal ($X^2(2) = 14.04, p < .001, Kendall's\ W = 0.27$) and valence ($X^2(2) = 9.09, p = .011, Kendall's\ W = 0.17$). Timbre had a significant effect on valence ($X^2(2) = 14.71, p < .001, Kendall's\ W = 0.28$). These results indicate that younger participants are broadly sensitive to different acoustic properties in both emotional dimensions.
    - *Wilcoxon:* Pitch (High vs. Low, $p_{corr} = .047, Z \approx -2.41, r = -.47$ for arousal). Tempo (Normal vs. Low, $p_{corr} < .001, Z \approx -4.01, r = -.79$; High vs. Low, $p_{corr} = .011, Z \approx -2.95, r = -.58$ for arousal; Normal vs. High, $p_{corr} = .013, Z \approx -2.85, r = -.56$; High vs. Low, $p_{corr} < .001, Z \approx -3.81, r = -.71$ for valence). Timbre (Normal vs. High, $p_{corr} = .037, Z \approx -2.5, r = -.49$ for valence).
- **Group 2 (22-25 years):** No acoustic component showed a significant overall effect on arousal or valence. This age group appears to represent a transitional phase, possibly reflecting more heterogeneous music perception profiles.

- **Group 3 (26-30 years):** Tempo significantly affected arousal ($X^2(2) = 6.66, p = .036, Kendall's\ W = 0.15$). Timbre significantly affected valence ($X^2(2) = 12.38, p = .002, Kendall's\ W = 0.28$). This suggests maintained sensitivity to temporal and spectral cues, particularly in emotional intensity and pleasantness, respectively.
    - *Wilcoxon:* Tempo (High vs. Low, $p_{corr} = .003, Z \approx -3.19, r = -.68$ for arousal). Timbre (High vs. Low, $p_{corr} = .023, Z \approx -2.43, r = -.52$ for valence).
- **Group 4 (31-40 years):** Pitch significantly affected arousal ($X^2(2) = 6.89, p = .032, Kendall's\ W = 0.23$). Tempo significantly affected arousal ($X^2(2) = 6.68, p = .035, Kendall's\ W = 0.22$) and valence ($X^2(2) = 11.85, p = .003, Kendall's\ W = 0.40$). The stronger effect of tempo on valence in this group may reflect increased reliance on rhythmic structures for affective appraisal during early midlife.
    - *Wilcoxon:* Tempo (Normal vs. Low, $p_{corr} = .013, Z \approx -3.17, r = -.82$ for valence). Other pitch/tempo effects on arousal were not significant after correction.
- **Group 5 (41+ years):** Pitch significantly affected arousal ($X^2(2) = 12.52, p = .002, Kendall's\ W = 0.69$), indicating a robust and targeted sensitivity to tonal height or key among older listeners. No other acoustic feature significantly influenced emotion in this group.
    - *Wilcoxon:* Pitch (Normal vs. High, $p_{corr} = .003, Z \approx -3.09, r = -1.03$; Normal vs. Low, $p_{corr} = .017, Z \approx -2.75, r = -.92$ for arousal). The high effect size value for this age group is likely due to the very small N=9. Therefore, interpretations regarding pitch sensitivity in older adults should be considered preliminary and warrant replication in larger samples.

These findings suggest that younger participants (15-21 years) were sensitive to manipulations across all three acoustic components. Sensitivity to tempo (for arousal) and timbre (for valence) persisted in the 26-30 age group. Older participants (31-40 and 41+ years) showed particular sensitivity to pitch manipulations affecting arousal, and the 31-40 group also responded to tempo for both arousal and valence. The 22-25 age group showed no significant overall effects for any component.

**Consistency Across Individual Audio Files**

A brief summary of the analysis of individual audio files indicates that the observed effects of tempo on arousal and valence were largely consistent across the two different musical pieces used for tempo manipulation. This suggests that the tempo effect is robust and not solely dependent on the idiosyncratic characteristics of a single musical stimulus. In contrast, the effects of pitch and timbre manipulations showed more variability depending on the specific musical piece, indicating that the inherent qualities of the original track might interact more strongly with these particular acoustic manipulations. This consistency for tempo further strengthens the conclusions regarding its primary role in modulating emotional responses.

## 4. Discussion

This study aimed to investigate the impact of manipulated acoustic cues (pitch, tempo, and timbre) in viral social media music on perceived emotional arousal and valence, utilizing an innovative online experimental platform. The findings reveal a complex interplay between these acoustic features and listeners' emotional experiences, significantly moderated by age and gender.

**Tempo and Emotional Arousal**

**Tempo consistently emerged as the most influential acoustic component** affecting emotional perception in this study. The finding that **decreased tempo significantly increased perceived arousal** is particularly noteworthy as it contrasts with some established literature that typically associates faster tempi with higher arousal (Liu et al., 2018). While some studies report a V-shaped relationship where medium tempi elicit the lowest arousal (Yang et al., 2025), the current results suggest a more direct inverse relationship for arousal within the tempo range and musical context investigated. The "low" tempos used (69/75 BPM) might have moved listeners away from an optimal low-arousal zone associated with medium tempos (around 85-120 BPM, as suggested by the "Medium Tempo Attraction Theory" (Lin et al., 2023)), thereby increasing arousal. Several psychological mechanisms could contribute to this observed increase in arousal with decreased tempo, especially within the context of "viral social media music," which is often inherently energetic or familiar in its original, typically faster form as mentioned below. These results, however, may reflect context-specific effects, such as genre familiarity or viral social media music expectations, and should not be generalized across all musical contexts.

- **Expectation Violation and Suspense:** Viral music often establishes strong rhythmic and metric expectations. A significant slowing of tempo could violate these ingrained expectations, creating a sense of surprise, unfamiliarity, or even unease, which can manifest as heightened physiological and psychological arousal (Van der Zwaag et al., 2011; Yang et al., 2025). The deviation from the "known" tempo might induce anticipatory tension as the listener attempts to reconcile the altered version with their memory of the original or typical exemplars of the genre.
- **Increased Cognitive Load or Attentional Demand:** Processing an unusually slow version of music that is typically faster, or music that deviates from expected genre conventions for tempo, might require greater cognitive effort to maintain engagement, track melodic and rhythmic development, or form coherent perceptual units(Lin et al., 2023; Van der Zwaag et al., 2011; Yang et al., 2025). This increased cognitive load or focused attention could contribute to higher reported arousal.
- **Altered Emotional Intensity Development:** Slower tempos can allow for a more drawn-out development of musical phrases and emotional expression. This protracted unfolding might lead to a build-up of emotional intensity over the listening period, perceived as higher arousal, particularly if the underlying harmonic or melodic content carries inherent emotional weight that is amplified by the slower pace (Van der Zwaag et al., 2011; Yang et al., 2025).

The finding that increased tempo led to higher valence ratings aligns more closely with general expectations, where faster music is often perceived as happier or more positive (Liu et al., 2018). This could be due to associations of fast tempi with energetic and joyful activities or expressions.

The study also revealed that the influence of pitch and timbre on emotional perception was not uniform across all participants but varied significantly with age and, for timbre, with its impact on valence in younger groups.

**Pitch Sensitivity and Age-Related Differences**

The influence of pitch, particularly mode, on valence is also a consistent theme in the literature (Carraturo et al., 2024; Droit-Volet et al., 2013), although this study focused on pitch height/key changes rather than major/minor mode explicitly for all stimuli. Recent research continues to explore the neural correlates of pitch perception and its emotional impact (Sayal et al., 2024). The age-specific effect of pitch on arousal found here adds a nuanced perspective. Pitch manipulations significantly impacted arousal levels in older participants (specifically the 41+ age group, and to some extent the 31-40 group). This heightened sensitivity in older adults could stem from several factors. Age-related changes in auditory processing might make pitch variations more salient or effortful to process, potentially leading to increased arousal (Arehart et al., 2014). Alternatively, a lifetime of exposure to musical conventions could make older listeners more attuned to subtle pitch deviations from expected norms, leading to a stronger affective response. Hofbauer & Rodriguez (2023) found that higher age predicted more *positive valence* ratings for musical stimuli, suggesting a complex relationship between age and emotional responses to music that may differ across emotional dimensions. The current study's findings specifically link pitch to arousal in older adults, which warrants further investigation.

**Timbre's Influence on Valence**

Timbre manipulations were found to significantly influence valence perception, particularly in younger age groups (15-21 and 26-30 years). Younger listeners may be more sensitive to or exploratory of novel timbral experiences, and their genre preferences might be more heavily influenced by timbral characteristics that define the appeal and emotional tone of contemporary music, including that found on social media (Armitage et al., 2024; Hailstone et al., 2009; Wang et al., 2021; Wu et al., 2014). The modifications to timbre in this study (altering frequency ranges and amplitudes) could have created sounds perceived as either richer/harsher or softer/clearer, directly impacting their pleasantness for these age groups.

**Gender-Based Differences in Emotional Responses**

The study identified distinct patterns of emotional sensitivity to acoustic cues based on **gender** as well. Male participants demonstrated greater sensitivity to changes in pitch (for both arousal and valence) and timbre (for valence). In contrast, female participants were more responsive to tempo manipulations, which significantly influenced both their arousal and valence ratings.

These findings contribute to a growing body of literature exploring gender differences in music perception and emotion (Jaquet et al., 2014) . For instance, Maidhof et al. (2023) reported gender differences in stress and mood responses to music and in the use of emotion regulation strategies, with women tending towards reappraisal and men towards suppression (Danhe, 2024). It is plausible that such underlying differences in emotional processing or regulation strategies could interact with how specific acoustic cues are perceived and affectively experienced. For example, the stronger arousal response to tempo changes in women might relate to a greater attunement to rhythmic entrainment or its impact on physiological states, which they might more actively engage with for mood modulation. Men's heightened sensitivity to pitch and timbre variations for valence might reflect different aesthetic criteria or a different way these features contribute to their emotional appraisal of music. The finding by Gabrielsson and Lindström (2001) that gender can moderate the pitch-arousal link is also relevant in this context.

**Broader Implications and Cognitive-Affective Models**

The current study contributes to the cognitive science of music by demonstrating how fundamental acoustic parameters, manipulated within ecologically relevant musical stimuli (viral social media music), systematically affect core emotional dimensions. The findings underscore that emotional responses to music are not monolithic but are shaped by the specific acoustic structure of the music and the individual characteristics of the listener. The use of music prevalent on social media is particularly relevant, as this context involves frequent, often passive, exposure to short musical segments intertwined with other media, potentially shaping emotional responses in ways distinct from dedicated music listening. Understanding these effects is crucial in an era where digitally mediated music is a constant emotional backdrop for many.

The results also highlight the importance of considering individual differences. The observed age and gender variations suggest that models of music emotion perception must account for demographic factors to achieve greater predictive accuracy and ecological validity. For instance, the differential sensitivity to tempo versus pitch/timbre between genders, or the varying impact of pitch on arousal across age groups, points to potentially different cognitive or affective processing strategies or sensitivities that develop or change over the lifespan or differ between sexes.

**Methodological Contributions and the Online Platform**

A significant contribution of this research is the development and successful implementation of a custom-built online experimental platform (survey.mousaee.ir) for data collection. This platform incorporated several features to enhance methodological rigor in an online setting, including SMS verification for participant uniqueness, DASS-21 screening to control for baseline emotional states, and a hearing/volume calibration test to ensure adequate stimulus perception. Such features address common challenges in online research, such as participant attentiveness and variability in equipment and listening environments (Warrenburg, 2020; Egermann et al., 2006). The ability to recruit a relatively large and demographically diverse sample (N=109) demonstrates the utility of online platforms for music perception research. While online studies offer increased ecological validity by allowing participants to use their own equipment in familiar settings, this also introduces variability. The platform's design attempted to mitigate some of this variability, striking a balance between ecological relevance and experimental control. The use of the SAM scale, a non-verbal pictorial tool, is also well-suited for online research with diverse participants (Bradley & Lang, 1994).

**Limitations of the Study and Directions for Future Research**

Despite the contributions, this study has several limitations that highlight the need for future research. First, while the overall sample size was adequate, some demographic subgroups, particularly the oldest age group (41+ years, n=9), were relatively small, which may limit the generalizability of findings specific to these groups. Future studies should aim for larger and more evenly distributed samples across age and gender categories.

Second, the reliance on self-report measures (SAM scale) for emotional responses, while standard, is subject to individual interpretation and potential biases. Incorporating objective physiological measures, such as heart rate variability, skin conductance, or even EEG in online settings where feasible, could provide convergent evidence and a more comprehensive understanding of emotional responses.

Third, the online nature of the experiment means less control over the participants' listening environment (e.g., ambient noise, quality of headphones/speakers) compared to laboratory studies. Although a hearing calibration step was included, variations likely remained.

Fourth, cultural background, a known modulator of music perception and emotion (Wang et al., 2021), was not systematically investigated. The stimuli were popular in a specific culture, and responses might differ in other cultural groups. Future cross-cultural research using the online platform could address this.

Fifth, the study focused on immediate emotional responses. Longitudinal studies are needed to explore how emotional responses to social media music evolve with repeated exposure and how the specific context of social media (e.g., accompanying visual content, social interactions) further modulates these responses.

Future research should also delve deeper into the "decreased tempo, increased arousal" phenomenon, exploring whether it is specific to certain genres (like viral pop instrumentals), dependent on familiarity and prior tempo expectations, or related to specific ranges of tempo change. Investigating the interactive effects of combined acoustic cue manipulations (e.g., tempo and pitch simultaneously) would also provide a more holistic understanding of music's emotional impact. Finally, exploring the influence of musical training, which was not a primary focus here, would add another important dimension to the understanding of individual differences.

## 5. Conclusion

The present study provides valuable insights into the nuanced ways specific acoustic components of viral social media music—pitch, tempo, and timbre—influence perceived emotional arousal and valence. The findings clearly demonstrate that tempo is a primary driver of emotional arousal, with the notable observation that decreased tempo led to increased arousal in the context of these stimuli. Furthermore, increased tempo was associated with more positive valence. The emotional impact of pitch and timbre was found to be significantly moderated by listener age and gender, with older adults more sensitive to pitch for arousal, younger adults showing broader sensitivity across components, and distinct patterns of sensitivity observed between male and female participants.

The research underscores the importance of considering individual differences in the study of music and emotion, highlighting the complex and multifaceted nature of how acoustic cues are decoded and experienced. The successful deployment of an innovative online experimental platform with built-in screening and calibration measures also signifies a valuable methodological advancement for conducting rigorous music perception research in ecologically relevant settings.

Ultimately, this study contributes to a deeper understanding of the intricate relationship between the structural elements of music prevalent in contemporary digital environments and the emotional responses they evoke. The findings have implications for theoretical models of music emotion, as well as practical applications in fields such as music therapy, media production, and personalized content recommendation systems that aim to leverage music for emotional impact. Continued research, particularly exploring interactive cue effects, cultural variations, and objective physiological correlates, will further illuminate the decisive role of music in shaping human emotional experience.

## References

Adachi, M., & Trehub, S. E. (1998). Children's expression of emotion in song. *Psychology of music*, *26*(2), 133-153.

Arehart, K. H., Croghan, N. B., & Muralimanohar, R. K. (2014). Effects of age on melody and timbre perception in simulations of electro-acoustic and cochlear-implant hearing. *Ear and hearing*, *35*(2), 195-202.

Armitage, J., Eerola, T., & Halpern, A. R. (2024). Play it again, but more sadly: Influence of timbre, mode, and musical experience in melody processing. *Memory & Cognition*, 1-12.

Barrett, L. F. (2017). The theory of constructed emotion: an active inference account of interoception and categorization. *Social cognitive and affective neuroscience*, *12*(1), 1-23.

Bradley, M. M., & Lang, P. J. (1994). Measuring emotion: the self-assessment manikin and the semantic differential. *Journal of behavior therapy and experimental psychiatry*, *25*(1), 49-59.

Brattico, E., Bogert, B., Alluri, V., Tervaniemi, M., Eerola, T., & Jacobsen, T. (2016). It's sad but I like it: The neural dissociation between musical emotions and liking in experts and laypersons. *Frontiers in human neuroscience*, *9*, 676.

Carraturo, G., Pando-Naude, V., Costa, M., Vuust, P., Bonetti, L., & Brattico, E. (2024). The major-minor mode dichotomy in music perception. *Physics of Life Reviews*.

Dalla Bella, S., Peretz, I., Rousseau, L., & Gosselin, N. (2001). A developmental study of the affective value of tempo and mode in music. *Cognition*, *80*(3), B1-B10.

Danhe, Z. (2024). Gender differences in the function of Music for emotion regulation development in everyday life an experience sampling method study.

Droit-Volet, S., & Meck, W. H. (2007). How emotions colour our perception of time. *Trends in* *cognitive sciences*, *11*(12), 504-513.

Droit-Volet, S., Ramos, D., Bueno, J. L., & Bigand, E. (2013). Music, emotion, and time perception: the influence of subjective emotional valence and arousal? *Frontiers in Psychology*, *4*, 417.

Eerola, T., & Saari, P. (2025). What emotions does music express? Structure of affect terms in music using iterative crowdsourcing paradigm. *PloS one*, *20*(1), e0313502.

Egermann, H., Fernando, N., Chuen, L., & McAdams, S. (2015). Music induces universal emotion-related psychophysiological responses: comparing Canadian listeners to Congolese Pygmies. *Frontiers in Psychology*, *5*, 116059.

Egermann, H., Nagel, F., Kopiez, R., & Altenmüller, E. (2006). Online measurement of emotional musical experiences using internet-based methods-An exploratory approach. Proceedings of the 9th International Conference of Music Perception and Cognition (ICMPC). Bologna, Italy,

Ferrer, R., Eerola, T., & Vuoskoski, J. K. (2013). Enhancing genre-based measures of music preference by user-defined liking and social tags. *Psychology of music*, *41*(4), 499-518.

Ferreri, L., Bigand, E., Bard, P., & Bugaiska, A. (2015). The influence of music on prefrontal cortex during episodic encoding and retrieval of verbal information: A multichannel fNIRS study. *Behavioural Neurology*, *2015*(1), 707625.

Fuentes-Sánchez, N., García-Fernández, M., Escrig, M. A., Eerola, T., & Pastor, M. C. (2025). The role of gender in emotional reactions elicited by music: Autonomic reactivity, facial expression, and self-reports. *Musicae Scientiae*, *29*(1), 131-148.

Gabrielsson, A., & Lindström, E. (2001). The influence of musical structure on emotional expression.

Hailstone, J. C., Omar, R., Henley, S. M., Frost, C., Kenward, M. G., & Warren, J. D. (2009). It's not what you play, it's how you play it: Timbre affects perception of emotion in music. *Quarterly journal of experimental psychology*, *62*(11), 2141-2155.

Hargreaves, D. J., & North, A. C. (1999). The functions of music in everyday life: Redefining the social in music psychology. *Psychology of music*, *27*(1), 71-83.

Hofbauer, L. M., & Rodriguez, F. S. (2023). Emotional valence perception in music and subjective arousal: Experimental validation of stimuli. *International journal of psychology*, *58*(5), 465-475.

Jaquet, L., Danuser, B., & Gomez, P. (2014). Music and felt emotions: How systematic pitch level variations affect the experience of pleasantness and arousal. *Psychology of music*, *42*(1), 51-70.

Juslin, P. N., & Sloboda, J. A. (2010). The past, present, and future of music and emotion research.

Juslin, P. N., & Västfjäll, D. (2008). Emotional responses to music: The need to consider underlying mechanisms. *Behavioral and brain sciences*, *31*(5), 559-575.

Lang, P. (1980). Behavioral treatment and bio-behavioral assessment: Computer applications. *Technology in mental health care delivery systems*, 119-137.

Lin, H.-M., Kuo, S.-H., & Mai, T. P. (2023). Slower tempo makes worse performance? The effect of musical tempo on cognitive processing speed. *Frontiers in Psychology*, *14*, 998460.

Liu, Y., Liu, G., Wei, D., Li, Q., Yuan, G., Wu, S., Wang, G., & Zhao, X. (2018). Effects of musical tempo on musicians’ and non-musicians’ emotional experience when listening to music. *Frontiers in Psychology*, *9*, 2118.

Lovibond, P. F., & Lovibond, S. H. (1995). Depression anxiety and stress scales. *Behaviour Research and Therapy*.

Maidhof, R. M., Kappert, M. B., Wuttke, A., Schwerdtfeger, A. R., Kreutz, G., & Nater, U. M. (2023). Effects of participant-selected versus researcher-selected music on stress and mood–the role of gender. *Psychoneuroendocrinology*, *158*, 106381.

McAdams, S., Douglas, C., & Vempala, N. N. (2017). Perception and modeling of affective qualities of musical instrument sounds across pitch registers. *Frontiers in Psychology*, *8*, 153.

Nineuil, C., Dellacherie, D., & Samson, S. (2020). The impact of emotion on musical long-term memory. *Frontiers in Psychology*, *11*, 2110.

North, A. C., Hargreaves, D. J., & McKendrick, J. (1999). The influence of in-store music on wine selections. *Journal of Applied psychology*, *84*(2), 271.

Pring, E. X., Olsen, K. N., Mobbs, A. E., & Thompson, W. F. (2024). Music communicates social emotions: Evidence from 750 music excerpts. *Scientific Reports*, *14*(1), 27766.

Russell, J. A. (1980). A circumplex model of affect. *Journal of personality and social psychology*, *39*(6), 1161.

Sayal, A., Guedes, A. G., Almeida, I., Pereira, D. J., Lima, C., Panda, R., Paiva, R. P., Sousa, T., Castelo-Branco, M., & Bernardino, I. (2024). Decoding Music-Evoked Valence and Arousal: Unraveling the Neural Correlates of Naturalistic Music Characteristics through fMRI. *bioRxiv*, 2024.2002. 2027.582309.

Sloboda, J. A., & Juslin, P. N. (2001). Psychological perspectives on music and emotion. *Music and emotion: Theory and research*, 71-104.

Stephenson, K., Quintin, E., & South, M. (2016). Age-related differences in response to music-evoked emotion among children and adolescents with autism spectrum disorders. *Journal of autism and developmental disorders*, *46*(4), 1142-1151.

Van der Zwaag, M. D., Westerink, J. H., & Van den Broek, E. L. (2011). Emotional and psychophysiological responses to tempo, mode, and percussiveness. *Musicae Scientiae*, *15*(2), 250-269.

Wang, X., Wei, Y., Heng, L., & McAdams, S. (2021). A cross-cultural analysis of the influence of timbre on affect perception in western classical music and chinese music traditions. *Frontiers in Psychology*, *12*, 732865.

Wu, B., Horner, A., & Lee, C. (2014). Musical timbre and emotion: The identification of salient timbral features in sustained musical instrument tones equalized in attack time and spectral centroid. ICMC,

Yang, Z., Su, Q., Xie, J., Su, H., Huang, T., Han, C., Zhang, S., Zhang, K., & Xu, G. (2025). Music tempo modulates emotional states as revealed through EEG insights. *Scientific Reports*, *15*(1), 8276.

Zaatar, M. T., Alhakim, K., Enayeh, M., & Tamer, R. (2024). The transformative power of music: Insights into neuroplasticity, health, and disease. *Brain, behavior, & immunity-health*, *35*, 100716.

Zentner, M., Grandjean, D., & Scherer, K. R. (2008). Emotions evoked by the sound of music: characterization, classification, and measurement. *Emotion*, *8*(4), 494.